JMIR Serious Games

Journal Impact Factor (JIF) (2023): 3.8
Volume 12 (2024) ISSN 2291-9279 Editor-in-Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Reviews

Viewpoints

Corrigenda and Addendas

Letter to the Editor

Research Letters

Tutorial

Effects of Virtual Reality Therapy for Patients With Breast Cancer During Chemotherapy: Randomized Controlled Trial

Mengdan Li[1], MPH; Zhifu Yu[1], MPH; Hui Li[1], MSN; Li Cao[1], BSN; Huihui Yu[1], MPH; Ning Deng[1], MD; Yunyong Liu[2], MM

1

2

Corresponding Author:
Yunyong Liu, MM

Related Article:

This is a corrected version. See correction statement: https://games.jmir.org/2024/1/e67718

Abstract

Background: Patients with breast cancer endure high levels of psychological and physical pain. Virtual reality (VR) may be an acceptable, safe intervention to alleviate the negative emotions and pain of patients with cancer.

Objective: We aimed to test the long-term effects of VR on psychological distress and quality of life (QOL) with traditional care in Chinese patients with breast cancer. We also explored the intervention mechanism and the acceptability of VR.

Methods: A total of 327 eligible participants were randomly assigned to a VR intervention group or a control group. The Distress Thermometer, QLQ-C30 (Quality of Life Questionnaire version 3.0), and Virtual Reality Symptom Questionnaire were assessed at baseline, postintervention (3 mo), and follow-up (6 mo). Analysis followed the intention-to-treat (ITT) principle. The generalized estimating equations model was used to analyze the longitudinal data, and the PROCESS macro was used to analyze the mediating effect.

Results: Compared with the control group, patients with breast cancer in the VR group had lower distress scores (*P*=.007), and higher health-related QOL scores (physical, role, emotional, cognitive, and social functioning) after 6 months (*P*<.05). Psychological distress had mediating effects on the longitudinal association between VR and the health-related QOL (indirect effect=4.572 - 6.672, all *P*<.05).

Conclusions: VR intervention technology may help reduce distress and improve QOL for patients with breast cancer over time. By incorporating a mediating analysis, we showed that the QOL benefits of VR intervention was manifested through positive effects on psychological distress risk factors.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000035049; https://www.chictr.org.cn/showproj.html?proj=53648

(JMIR Serious Games 2024;12:e53825) doi:10.2196/53825

KEYWORDS
virtual reality; breast neoplasms; quality of life; psychological distress; longitudinal studies

Introduction

Breast cancer is a common malignant tumor throughout the world [1]. With the progress of detection and treatment technology, the survival rate of patients with breast cancer has been greatly improved. However, increasing numbers of studies have also paid attention to the adverse psychological and physiological sequelae caused by breast cancer surgery or chemotherapy [2,3]. Patients with cancer are at several times the risk of psychological disorders due to lack of normal life, communication, and interaction [4-6]. Other symptoms related to chemotherapy such as nausea, vomiting, and anorexia are frequently reported [7]. These may affect treatment compliance and quality of life (QOL). From 2020, COVID-19 has been spreading around the world. In the process of fighting the pandemic, the normal operations of the hospitals were affected [8]. The postponement of re-examinations for some patients with cancer, interregional hospitalization, and self-isolation have all increased patients' emotional distress [9]. In addition, patients have less access to the outside world and have lost normal social interactions. Psychological needs are even less satisfied, with considerably increased loneliness [10]. Without timely intervention and treatment, a poor mental state might cause tumor progression and deterioration [11].

Virtual reality (VR) is a 3D virtual scene that simulates reality, generated by a computer. It immerses the user in a virtual environment by using specific human-machine interfaces, such as a head-mounted display, a set of wired gloves, a position tracker or other controllers to experience a sense of presence or immersion [12]. Specifically, it is an artificial environment that is experienced through sensory stimuli such as images and sounds provided by a computer. In this environment, our behavior can partially determine our feelings [13]. In recent years, especially in the health system, the emergence of this technology provided patients with a safe environment for intervention and treatment [14-17]. Its effectiveness stemmed from the fact that patients could focus on pleasant or interesting stimuli rather than unpleasant symptoms [18]. These techniques were generally categorized as distraction interventions. By using head-mounted devices, VR immersed patients in computer-generated views while engaging various senses, providing a comprehensive stimulus that helped isolate patients from the hospital environment [19].

Previous studies have confirmed that the VR interventions could alleviate chemotherapy pain and enhance the QOL for patients with cancer [20]. However, research on the mediating mechanisms between VR and QOL remains limited. Among the potential mediating mechanisms, the most widely accepted hypothesis was that VR improves emotional states. Patients who experienced less anxiety, more fun, and more positive emotional valence during VR distraction were more likely to report subjective pain reduction [21]. Specifically, when patients experienced more enjoyment, painful treatments became more tolerable [22], potentially enhancing their self-efficacy and, consequently, improving their QOL [22,23]. These studies suggested that the relationship between VR and the QOL in patients with cancer might have been mediated by emotional regulation. Therefore, it is necessary to verify the effectiveness and obtain the best clinical practice evidence.

Some studies on VR had some limitations, for example, most of the trials lacked a control group [23,24], and the sample size was small (n<50) [25-27]. Owing to the lack of repeated exposure and long-term follow-up results [28-30], it has not been adequately established whether the effects of VR could endure for longer timespans beyond the usage of the device. Our main objective was to explore the long-term effects of VR in reducing distress and some chemotherapy symptoms, as well as improving the QOL for patients with breast cancer in China. Moreover, we further explored whether or not psychological distress mediated the effect of VR interventions on the health-related QOL.

Methods

Study Design and Setting

This study was a single-blinded randomized controlled trial, and was registered in the Chinese Clinical Trial Registry (registration ChiCTR2000035049). Patients were randomly assigned (1:1) to the VR group and control group by block randomization with varying block sizes of 4. The random sequence was generated using the Random Allocation Software (version 23; IBM Corp) by a graduate student who was not involved in the intervention or data collection. Sequentially numbered, opaque, sealed envelopes were prepared by the student. Before the first chemotherapy (baseline, T0), all the participants completed the informed consent form and a baseline questionnaire. Afterward, the second and third questionnaires were completed in the third month (post intervention, T1) and the sixth month (follow-up, T2), respectively. The content of the measurement includes both physical symptoms and psychological and emotional aspects.

Participants

This study was conducted in the Cancer Hospital of China Medical University, which is a public cancer treatment center in northeastern China, from April 2020 to March 2021. The selection criteria were as follows: (1) a confirmed diagnosis and treatment of breast cancer for the first time, and a plan to undergo chemotherapy; (2) aged 18 to 79 years; and (3) complete cognition, normal reading ability, and barrier-free communication. The exclusion criteria were as follows: (1) doctors assessed life expectancy to be less than 6 months, (2) patients with emotional illness, and (3) patients who were receiving other forms of psychotherapy. All participants provided informed consent.

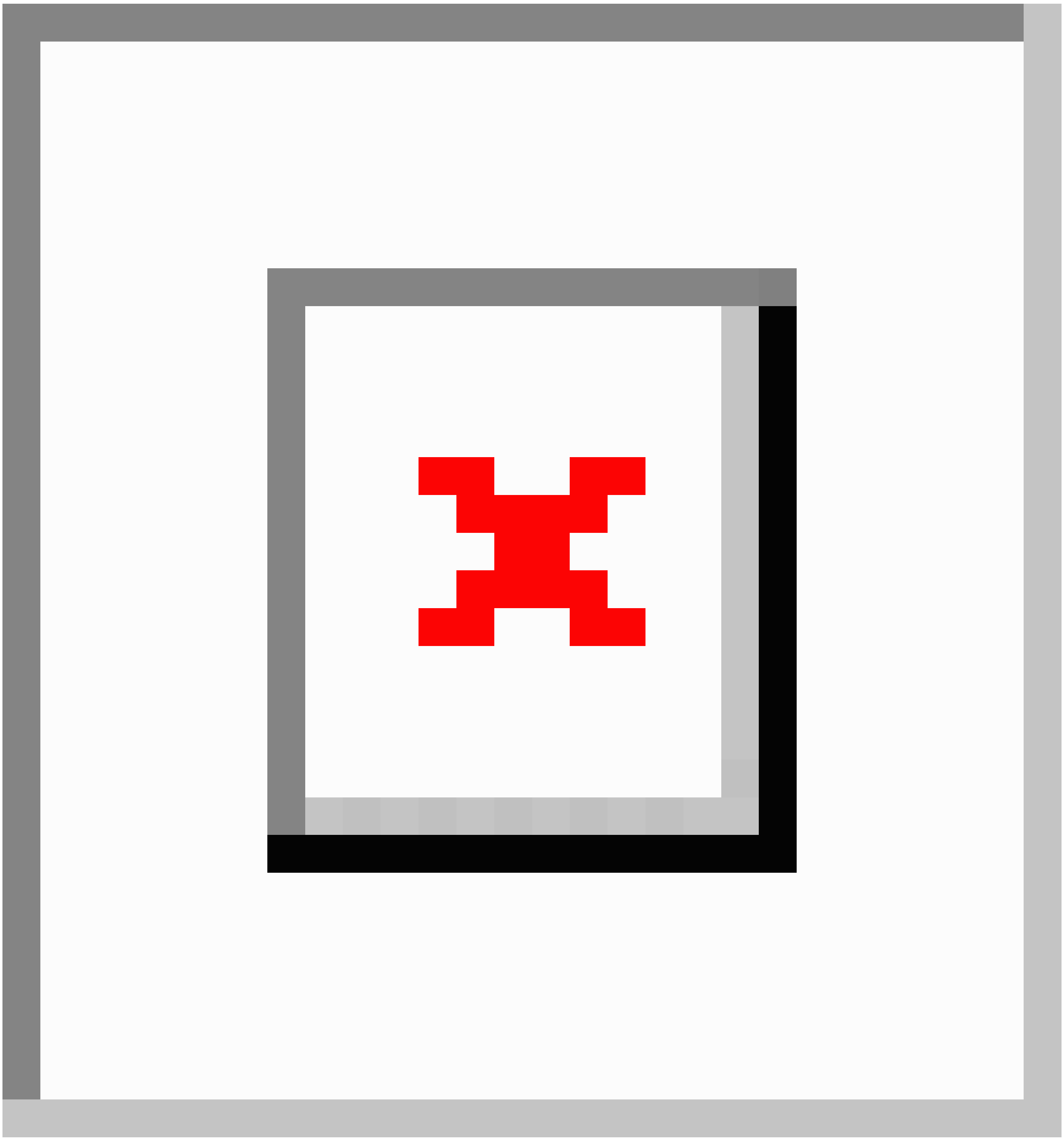
Sample Size

The sample size was calculated using the G*Power (version 3.1; Heinrich-Heine-Universität Düsseldorf) program. Based on a 2-sided significance level of .05, an effect size of 0.6 for distress outcome [31], the 2-arm parallel trial power of 0.80, and a 35% attrition rate, a minimum of 130 participants per group was needed.

Intervention Content and Frequency

The VR device consists of 2 parts: a headset (VIVES110) and a hand controller. Our team developed 3 virtual environments in total for participants to choose freely, and embedded a voice guidance system in the device to help participants relax. Patients could complete explorations such as walking through the forest, walking on the beach, and sightseeing at sea, or relaxation training according to the voice guidance (Figure 1). The intervention trial was a 12-week VR training. The nurses guided participants to wear the device during chemotherapy intervals or during breaks, and to participate for 15 - 20 minutes 1 to 2 times a week in the hospital. All participants had to have at least 12 interventions before T1. After each VR intervention, participants reported whether they had possible adverse symptoms. The participants in the control group only underwent chemotherapy as per the treatment plan in the hospital, and refrained from beginning any VR treatment.

Figure 1. (**A**) shows how a virtual reality (VR) device was used during the intervention; (**B**) shows one of the screen captures from the VR experiences.



Measurements

Demographic and Clinical Information

Participants completed a standardized self-report in the first round of questionnaires, including age, residence, marital status, education level, employment status, and disease stage.

Distress

The Distress Thermometer (DT) [32] is a single-item, self-report measure of psychological distress. The DT has an 11-point range from 0 (no distress) to 10 (extreme distress). Patients were asked to choose the number that best describes how distressed they have been in the past week.

Health-Related QOL

The QOL of the patients with cancer was measured using the Chinese-language translation of the validated European Organization for the Research and Treatment of Cancer QLQ-C30 (Quality of Life Questionnaire version 3.0) [33]. The QLQ-C30 includes a global health status or QOL scale [QL], 5 functional scales (physical [PF], role [RF], emotional [EF], cognitive [CF], and social [SF]), 3 symptom scales (fatigue, pain, and nausea and vomiting), and 6 single-item symptom scores. All scale scores range from 0 to 100, with a high score on the functional scales indicating a high level of functioning and a high score on the symptom scales indicating greater severity of individual symptoms.

Motion Sickness

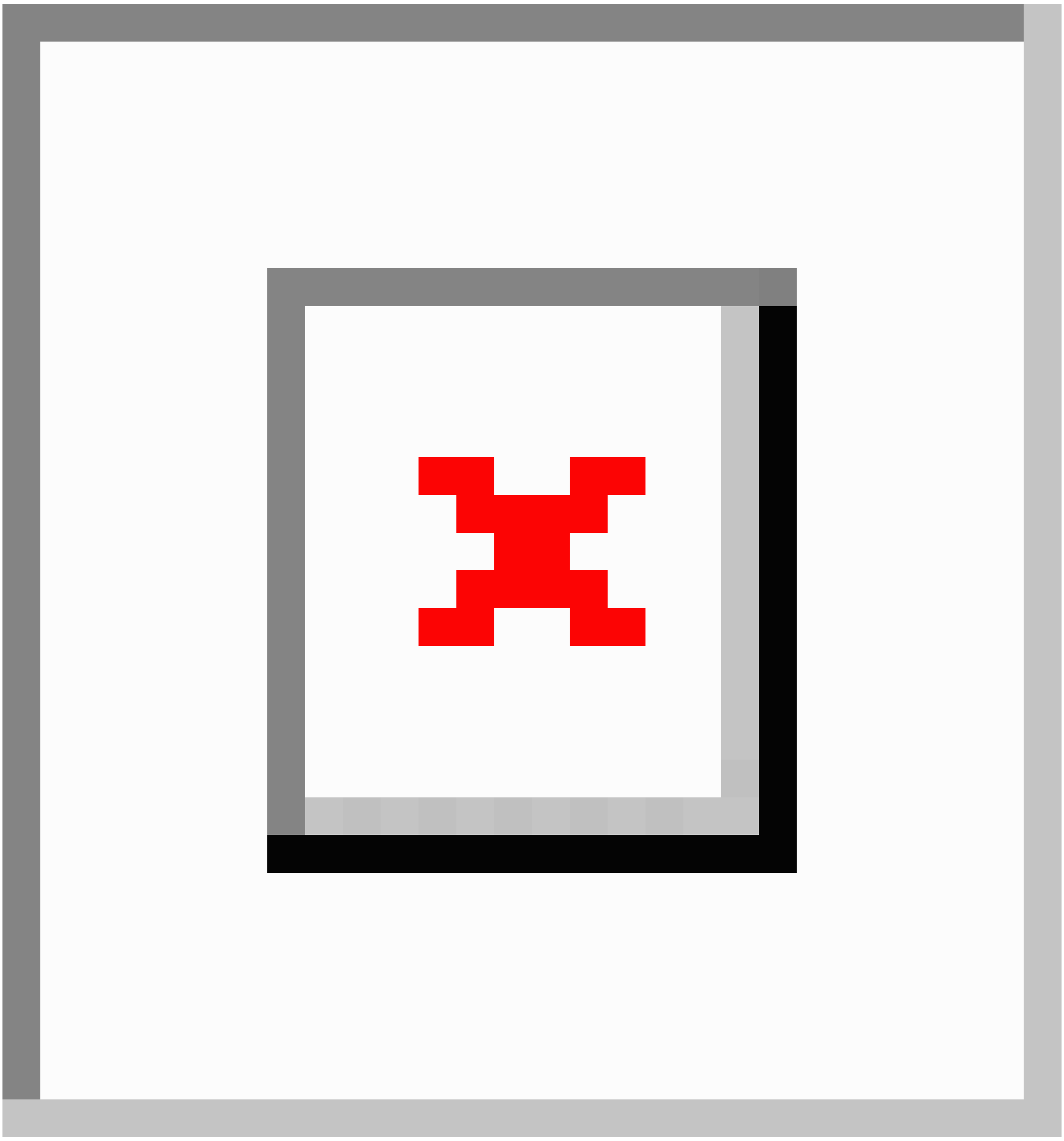
The Virtual Reality Symptom Questionnaire (VRSQ), which was developed by Ames et al [34] in 2005, was used. After the VR intervention, the VRSQ was completed to evaluate the possible symptoms of motion sickness that may occur in a VR environment. The questionnaire assessed 8 general physical side effects (general discomfort, fatigue, boredom, drowsiness, headache, dizziness, concentration difficulties, and nausea) and 5 visual effects (eye fatigue, eye pain, blurred vision, and difficulty focusing). The score ranges from 0 to 6, indicating a degree of symptoms from none to very severe.

Data Analysis

The analysis for this study was performed using IBM SPSS Statistics software (version 23). Appropriate descriptive statistics were used to summarize the characteristics and outcome variables of the participants. The *t* test, chi-square test, and the Wilcoxon signed-rank test were used to evaluate between-group differences. The intention-to-treat (ITT) principle was adopted for outcome analysis [35]. A generalized estimating equations (GEE) model is an extension of the quasi-likelihood method and is often used to analyze longitudinal data [36]. The characteristic of this data is the correlation between the multiple observation points for the same individual, which cannot be processed by the general linear model. The GEE model can solve the correlations between longitudinal data, as well as missing data. In this study, the GEE model was used to evaluate the effects of VR intervention on the changes of QOL and distress scores of patients with breast cancer at various time points (T0, T1, and T2). All statistical tests involved were 2-sided with the level of significance set at 0.05.

We used mediation analysis to investigate the relationship between the VR intervention at T0 and the health-related QOL at T2, with the psychological distress at T2 as a mediator (Figure 2). A mediator variable was defined as a third variable that changed the association between an independent and dependent variable [37]. It provided additional insight into information about the causal links between 2 strongly associated variables. The PROCESS macro with model 4 was used and, to ensure the stability of the path coefficient estimates, the analysis for the mediation model was supplemented with 5000 bootstrap replications. The significance of the indirect effect was examined by the bias-corrected 95% CI after bootstrapping. An indirect effect was considered statistically significant if the 95% CI did not include 0.

Figure 2. Hypothesized mediation model (a, b, and c′=unstandardized regression coefficients): indirect effect=a×b; direct effect=c′, total effect c=sum of indirect and direct effects=a×b+c′.



Ethical Considerations

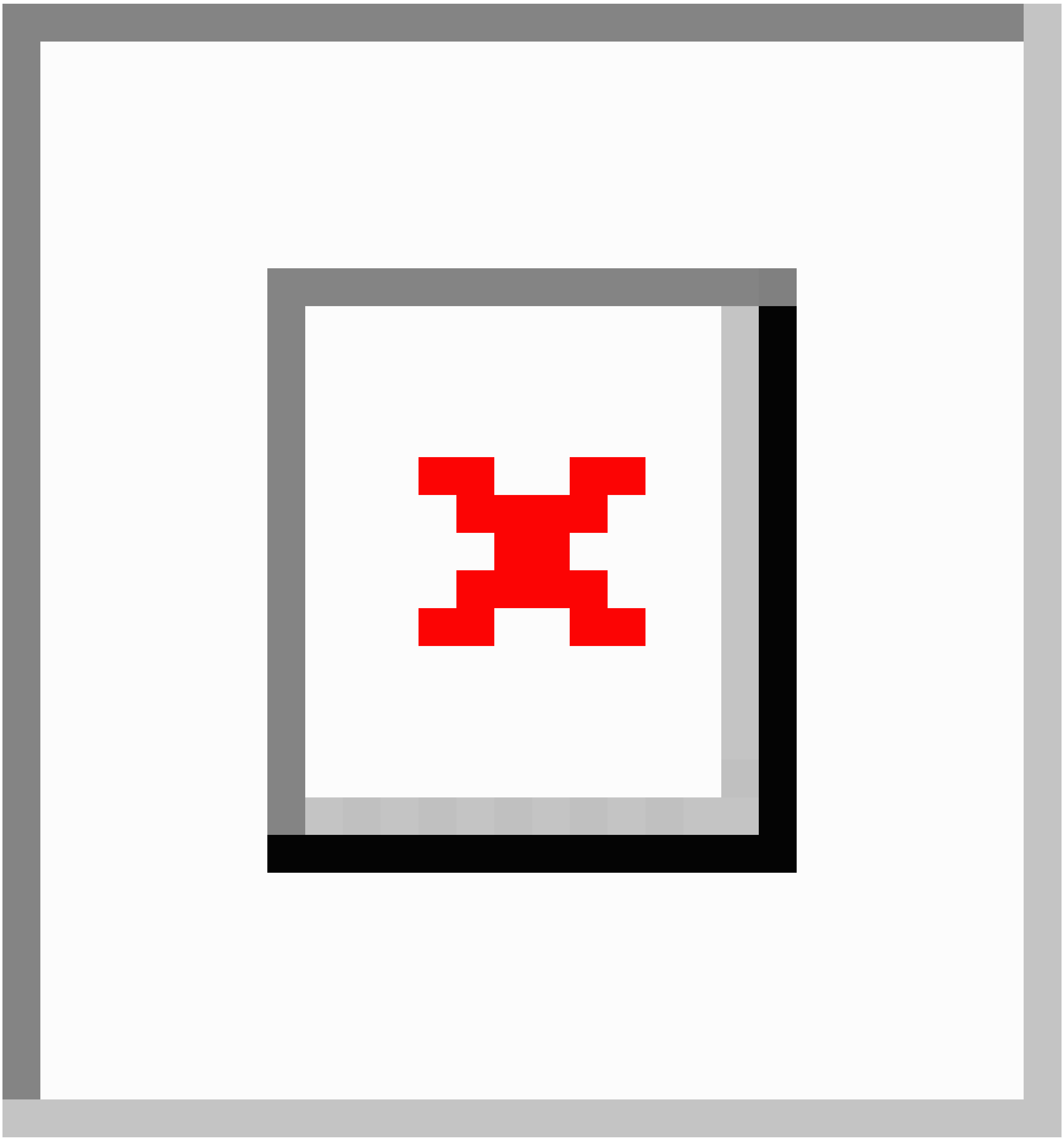
This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of the Liaoning Cancer Hospital & Institute (reference 20200301-2). Each participant provided written informed consent to take part in the study. For the images with a participant' face and body (Figure 1), the author had obtained permission to publish in the study.

Results

Recruitment

As shown in the CONSORT (Consolidated Standards of Reporting Trials) flow diagram (Figure 3; see checklist in Multimedia Appendix 1), 500 participants were screened for eligibility. Among these patients, 120 were excluded because they did not meet the criteria, and 53 declined to participate because of lack of time or interest, or for no specific reasons. Ultimately, 327 participants were recruited, although 97 participants failed to complete the reassessment at T1 or T2, and 18 participants did not have enough interventions. The T1 and T2 attrition rates were 25% and 14%, respectively, and the overall attrition rate was within our acceptable range. We compared the baseline characteristics between those who completed the study and those who dropped out from the study, and there were no significant differences (Multimedia Appendix 2).

Figure 3. CONSORT diagram. CONSORT: Consolidated Standards of Reporting Trials; VR: virtual reality.



Demographics and Clinical Characteristics

The participants' average age at diagnosis was 54.3 (SD 9.58) years, and of whom 216 (66.1%) were residents of a city, 198 (60.6%) had below middle school–level education, 267 (81.7%) were married, and 294 (89.9%) had no family history of cancer. More than half of the participants had an annual personal income of less than CN ¥20,000 (US $2838.61). In total, 263 (80.4%) participants had cancer at stage I or II. At the baseline, there were no differences between the 2 groups in either sociodemographic or clinical variables (Table 1).

Table . Demographic and clinical characteristics of participants by group at baseline.

Variables		Total (N=327)	VR[a] (n=163)	Control (n=164)	Statistics	
					t test or chi-square (df)	*P* value
Age (years) at diagnosis, mean (SD)		54.3 (9.6)	53.6 (9.4)	55.2 (9.7)	0.852[b] (325)	.09
Residence, n (%)					0.659[c] (1)	.16
	City	216 (66.1)	114 (69.9)	102 (62.2)		
	Country	111 (33.9)	49 (30.1)	62 (37.5)		
Education, n (%)					1.001[c] (2)	.99
	Below middle school	198 (60.6)	99 (60.7)	99 (60.4)		
	High school	79 (24.2)	39 (23.9)	40 (24.4)		
	University or above	50 (15.3)	25 (15.3)	25 (15.2)		
Marital status, n (%)					0.130[c] (1)	.98
	Married	267 (81.7)	133 (81.6)	134 (81.7)		
	Single, widowed, or divorced	60 (18.3)	30 (18.4)	30 (18.3)		
Work, n (%)					0.349[c] (2)	.84
	Yes	170 (52.0)	78 (47.9)	92 (56.1)		
	Retired	108 (33)	54 (33.1)	54 (32.9)		
	No	49 (15)	31 (19)	18 (11)		
Menopause, n (%)					1.963[c] (1)	.18
	Yes	197 (60.2)	92 (56.4)	105 (64.0)		
	No	130 (39.8)	71 (43.6)	59 (36)		
Family history of cancer, n (%)					1.699[c] (1)	.20
	Yes	33 (10.1)	20 (12.3)	13 (7.9)		
	No	294 (89.9)	143 (87.7)	151 (92.1)		
Annual personal income (1 CN ￥=0.14 US $)					2.422[c] (2)	.30
	<20,000	181 (55.4)	85 (52.1)	96 (58.5)		
	20,000 - 50,000	121 (37)	67 (41.1)	54 (32.9)		
	>50,000	25 (7.6)	11 (6.7)	14 (8.5)		
BMI (kg/m^2), mean (SD)		25.6 (10.4)	24.4 (3.4)	24.9 (3.3)	1.366[b] (325)	.17
Cancer stage, n (%)					3.869[c] (3)	.51
	I	100 (30.6)	46 (28.2)	54 (32.9)		
	II	163 (49.8)	84 (51.5)	79 (48.2)		
	III	48 (14.7)	25 (15.3)	23 (14)		
	IV	16 (4.9)	8 (4.9)	8 (4.9)		

[a]VR: virtual reality.

[b]*t* test.

[c]Chi-square.

Outcome Variables at Baseline

Overall, there was no significant difference between the 2 groups in each scale score at baseline (*P*<.05). In total, 245 (74.9%) patients scored ≥4 on the DT and had psychological distress. At the same time, participants in both groups reported low QOL at T0. Detailed scores are shown in Table 2.

Table . Mean scores of the outcome variables for intervention and control groups across study time points and the baseline comparisons.

Outcome variables and group		T0 (baseline), mean (SD)	T1 (3 months), mean (SD)	T2 (6 months), mean (SD)	Comparison of groups at T0	
					t test (*df*)	*P* value
Participants (n)						
	Control	164	125	111		
	VR[a]	163	121	101		
DT [b]					1.306 (325)	.19
	Control	5.15 (2.35)	4.63 (1.91)	4.27 (2.21)		
	VR	4.83 (2.05)	3.27 (1.81)	3.01 (1.96)		
QL [c]					−0.843 (325)	.40
	Control	46.87 (24.28)	55.12 (19.58)	54.12 (18.15)		
	VR	48.93 (19.49)	60.56 (17.95)	62.87 (22.10)		
PF [d]					−0.392 (325)	.70
	Control	63.54 (23.08)	65.54 (22.57)	65.93 (20.71)		
	VR	64.54 (23.22)	72.63 (19.01)	80.20 (18.70)		
RF [e]					−0.271 (325)	.79
	Control	62.91 (30.74)	66.33 (26.35)	68.31 (22.51)		
	VR	63.80 (29.14)	70.31 (27.35)	79.70 (24.56)		
EF [f]					−1.108 (325)	.27
	Control	63.92 (19.14)	58.66 (24.03)	62.91 (26.07)		
	VR	66.36 (20.60)	74.76 (18.58)	79.62 (19.49)		
CF [g]					−1.281 (325)	.20
	Control	71.85 (21.57)	65.18 (24.51)	67.40 (26.52)		
	VR	74.95 (22.17)	81.11 (17.91)	81.02 (20.42)		
SF [h]					−1.776 (325)	.08
	Control	57.42 (27.27)	54.13 (26.18)	59.34 (21.55)		
	VR	62.88 (28.35)	63.97 (25.12)	72.94 (21.72)		

[a]VR: virtual reality.
[b]DT: Distress Thermometer.
[c]QL: global health status or quality of life scale.
[d]PF: physical functioning.
[e]RF: role functioning.
[f]EF: emotional functioning.
[g]CF: cognitive functioning.
[h]SF: social functioning.

Effects of VR Interventions

Effects on Distress

As shown in Figure 4 and Table 3, compared with T0, DT scores in the VR group gradually decreased at T1 and T2 (T1: mean 3.27, SD 1.81; T2: mean 3.01, SD 1.96), and this change was significantly different from that in the control group, as indicated by the interaction term of group×T1 (β: −1.049, 95% CI −1.635 to −0.445; *P*=.001) and group×T2 (β: −.947, 95% CI −1.635 to −0.258; *P*=.007).

Figure 4. Mean scores of distress in the virtual reality and control groups at the 3 time points of data collection. VR: virtual reality.

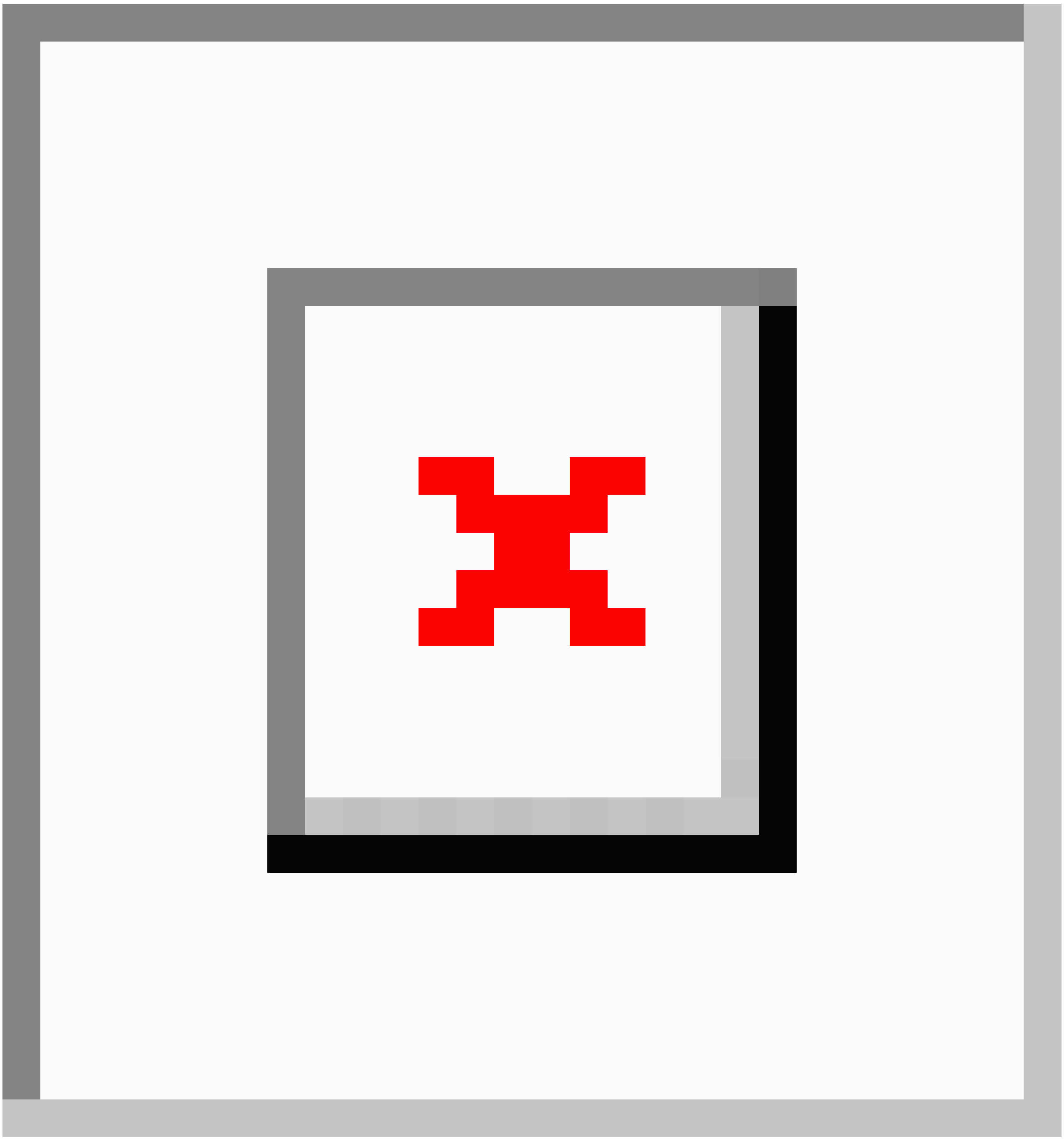


Table . Generalized estimating equation models for the comparison of outcomes across time between the control and intervention groups.

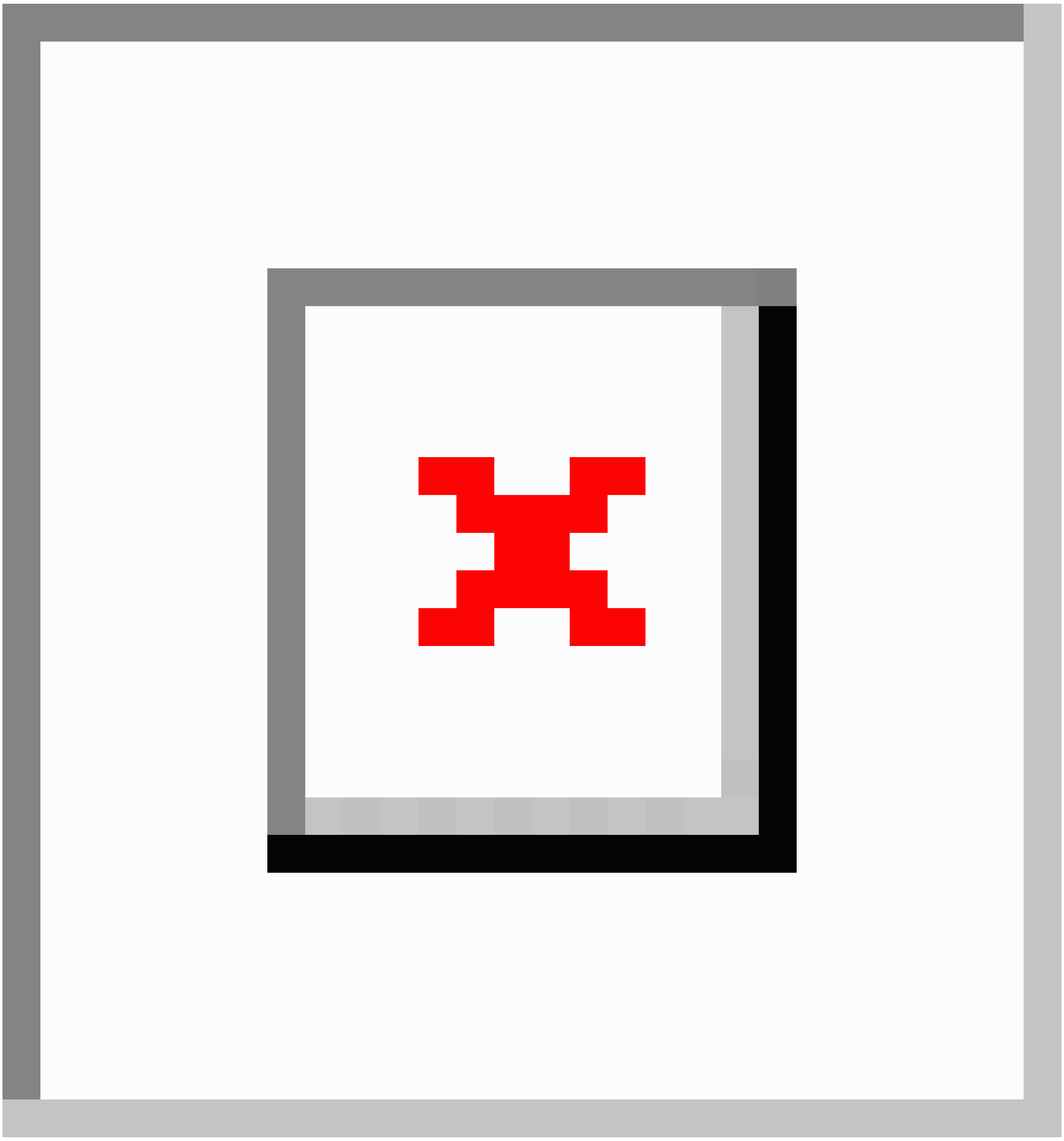
Outcome variables	Regression coefficients									
	Group		T1[a]		T2[b]		Group×T1		Group×T2	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
DT[c]	−.318 (−.794, 0.158)	.19	−.513 (−0.972, −0.053)	.03	−.872 (−1.417, −0.326)	.002	−1.049 (−1.635, −0.445)	.001	−.947 (−1.635, −0.258)	.007
QL[d]	2.052 (−2.704, 6.808)	.40	8.242 (3.346, 13.137)	.001	7.247 (2.452, 12.042)	.78	3.388 (−3.023, 9.799)	.30	6.698 (−0.147, 13.543)	.06
PF[e]	1.003 (−3.999, 6.006)	.69	2.008 (−3.731, 7.747)	.49	2.397 (−3.343, 8.138)	.41	6.087 (−1.396, 13.570)	.11	13.261 (5.434, 21.087)	.001
RF[f]	.897 (−5.575, 7.370)	.79	3.43 (−3.220, 10.081)	.31	5.409 (−1.149, 11.966)	.11	3.084 (−5.873, 12.041)	.50	10.491 (1.419, 19.563)	.02
EF[g]	2.437 (−1.861, 6.735)	.27	−5.259 (−10.603, 0.084)	.05	−1.011 (−6.891, 4.870)	.74	13.661 (6.991, 20.332)	<.001	14.271 (7.084, 21.459)	<.001
CF[h]	3.099 (−1.629, 7.827)	.20	−6.668 (−12.323, −1.014)	.02	−4.45 (−10.566, 1.666)	.15	12.83 (5.928, 19.733)	<.001	10.525 (2.953, 18.096)	.006
SF[i]	5.465 (−0.547, 11.476)	.08	−3.239 (−9.473, 2.887)	.30	1.922 (−3.434, 7.278)	.48	4.378 (−3.725, 12.482)	.03	8.132 (0.479, 15.785)	.04

[a]T1: the third months of intervention.
[b]T2: the sixth months of follow-up.
[c]DT: Distress Thermometer.
[d]QL: global health status or quality of life scale.
[e]PF: physical functioning.
[f]RF: role functioning.
[g]EF: emotional functioning.
[h]CF: cognitive functioning.
[i]SF: social functioning.

Effects on Health-Related QOL

As shown in Figure 5, the VR group showed a great improvement in the global health status scale at T1 and T2 but, when compared with the control group, statistical significance was not reached (group×T1: *P*=.30 and group×T2: *P*=.06). In terms of functional subscales, significant differences between the changes in the 2 groups were observed in the PF and RF scales at T2 (*P*=.001 and .02), and in the EF, CF, and SF scales at both T1 and T2 (*P*<.05; Table 3). At the same time, we also analyzed the other symptom subscales. Nausea and vomiting, appetite loss, dyspnea, and other symptoms of participants in the VR group were also significantly improved at different time points with respect to T0 (*P*<.05), as shown in Multimedia Appendices 3 and 4.

Figure 5. Mean scores on each dimension of quality of life in the virtual reality and control groups between the 3 time points of data collection. VR: virtual reality.



Results of the Mediating Analysis

According to the GEE model, the VR group had significantly greater improvement in distress compared with the control group across the preintervention and postintervention time points. The differential changes between the 2 groups supported the further exploration of the mediating roles. Table 4 shows that psychological distress had a significant indirect effect on the global health status in the participants of the VR group (indirect effect=6.245; 95% CI 2.965, 10.282). Similarly, the effects of VR on 5 health-related functional subscales were also mediated by improving psychological distress (indirect effect=4.572 - 6.672, *P*<.05).

Table . Mediation analysis results for the virtual reality effects on health-related quality of life via distress. The virtual reality intervention group at T0 was set as the independent variable (X), the global health status and 5 functioning subscales at follow-up T2 were set as dependent variables (Y), and the distress at T2 as a mediator (M). * *P*<.05.

Variables	a[a]	b[b]	c'[c]	ab[d]	95% CI
QL[e]	−1.265*	−4.938*	2.506	6.245*	2.965-10.282
PF[f]	−1.265*	−4.333*	8.784*	5.480*	2.645-8.725
RF[g]	−1.265*	−3.803*	6.578	4.810*	2.226-8.022
EF[h]	−1.265*	−5.275*	10.037*	6.672*	3.553-10.004
CF[i]	−1.265*	−3.615*	9.052*	4.572*	2.320-7.266
SF[j]	−1.265*	−3.724*	8.886*	4.711*	2.080-7.919

[a]a: the direct effect of X on M.
[b]b: the direct effect of M on Y when X is controlled.
[c]c : the direct effect of X on Y when M is controlled.
[d]ab: the indirect effect of X on Y through M.
[e]QL: global health status or quality of life scale.
[f]PF: physical functioning.
[g]RF: role functioning.
[h]EF: emotional functioning.
[i]CF: cognitive functioning.
[j]SF: social functioning.

Cybersickness Symptoms

Some cybersickness symptoms related to VR were analyzed through VRSQ. Data showed that in the first month of intervention, the frequency of patients with each symptom was less than 20% (32/163; Multimedia Appendix 5).

Discussion

Principal Findings

This study demonstrated that VR technology may help reduce distress and improve the health-related QOL of patients with breast cancer over time. By incorporating a mediation analysis, we showed that the QOL benefits of VR are manifested through its positive effects on psychological distress risk factors.

Psychological distress is a common side effect of cancer treatment. In 2018, a cross-sectional study completed by our team showed that 56.5% of cancer survivors had psychological distress and scored ≥4 on the DT [38]. However, the outbreak of COVID-19 aggravated the emotional distress of patients with cancer. In this study, the proportion rose to 74.9% (245/327) and it was necessary to take appropriate intervention measures to reduce the psychological pressure. VR allows patients to be completely engaged in an immersive environment, which may help distract them from noxious stimuli, thereby relieving psychological stress [19,28]. In our study, the distress level was significantly reduced in the VR group compared with the control group (undergoing usual care). This longitudinal effect did not disappear over time, suggesting that the effect of VR was not simply due to the novelty of the experience [22]. Yang et al [39] proposed a 4-layer theoretical framework of a potential VR intervention for mental health. Based on the hypothesis, VR delivers predictability in an unpredictable environment, and helps patients cope with their stress, evokes emotion, and diverts attention [40].

VR has shown promise in improving health status among patients with breast cancer [41]. Reynolds et al [30] reported that participants with metastatic breast cancer who were subjected to VR experiences showed clinically significant decreases in fatigue, stress, and pain during the trial and at follow-up. However, many studies focus on the efficacy of in-the-moment distraction from treatment, and the long-term effects of VR are still missing. Some of our preliminary findings are encouraging. The longitudinal effects of VR could be sustained at least for 6 months. At the same time, these positive effects were not only shown in the dimensions of PF and EF of QOL, but also in the dimensions of RF, CF, and SF, and even some chemotherapy side effects such as vomiting and insomnia were improved and relieved to varying degrees. This was also a comprehensive assessment of the effects of VR on QOL. In addition to the clear effect on pain management [42,43], this study also provided more possibilities for VR in improving patients' QOL and chemotherapy tolerance. In the future, VR may be considered for incorporation into the clinical setting to provide a convenient, attractive, and easily applied intervention.

The psychological and physiological fields are the most commonly studied areas for VR [26,44]. In our study, we focused on relatively stable hospitalized patients undergoing chemotherapy rather than surgical patients. We agree with Espinoza et al [45] that VR technology shifted patients' attention to the selected scene track, and reduced their attention to pain or anxiety. However, presumably this distraction effect did not occur in the days after the removal of the device, and so could not fully explain the maintenance of VR. Our other important finding was that psychological distress could play a mediating role in the VR-QOL interaction, which provides an explanation for the persistence of VR. Psychological factors play an important role in the development and regression of tumors [46]. Generally speaking, a good state of psychological health can stimulate a number of body functions and significantly boost the effectiveness of the chemotherapy [47]. Our results confirmed once again that improved mood can promote positive health outcomes and QOL. Several design characteristics of our VR system, such as guided relaxation and soothing music, support the mediation theory of positive mood. These observations further suggest that the intervention mechanism of VR is multifactorial, being mediated by attentional, cognitive, and emotional effects [48].

Owing to the specificity of the disease, the impact of the COVID-19, and the long study period, some participants dropped out of the study, and the overall attrition rate reached 35%, which was similar to previous studies [31] and within our estimated range. However, among the lost participants, only 18 patients were lost due to the substandard intervention frequency, accounting for less than 16%. At the same time, the VR equipment we used was comfortable to wear, and the designed scenes were natural. We also assessed the possible side effects of VR devices through the VRSQ scale, but all symptoms occurred at a frequency of less than 20%, which was considered negligible [49]. These findings are important and demonstrate that VR intervention can be a viable and acceptable treatment for patients with breast cancer [30].

Limitations

Although this study had several strengths and extended the literature on the use of VR, it was not without limitations. First, loss of muscle strength leading to a decrease in limb functionality is also a common complication seen in breast cancer survivors. This VR system did not incorporate exercise rehabilitation training such as improving patients' upper limb mobility, so the research on the effects of VR was incomplete. Second, although we designed 3 intervention scenarios, and required each patient to achieve multiple VR interventions, we did not further analyze the impact of the duration and frequency of interventions, and did not verify the effect of repeated VR exposure. Third, this study assessed the effects of VR through self-reported measures, but because the participants, investigators, and nurses were unlikely to be blind, this could have introduced bias into some of the results. In future, physiological indicators (such as stress-related molecules or brain waves) can be used to monitor changes and evaluate the effects of VR.

Conclusions

This study found that VR has the potential to sustainably relieve the level of distress and improve the QOL in women with breast cancer. Importantly, the positive results were sustained for at least 6 months. Furthermore, reducing emotional distress was identified as one of the possible mechanisms by which VR affects QOL. As physicians, it is our duty to participate in the development of these innovations to meet the clinical need for effective nonpharmacologic adjuncts, especially for patients with cancer or during special times such as the pandemic. We suggest that future research should be carried out in multiple centers and medical institutions to verify the effectiveness with other cancer treatments and to explore more possibilities for VR interventions.

Acknowledgments

This work was supported by the Key Research and Development Program of Liaoning Province (grant 2019JH2/10300013).

Data Availability

Research data are not publicly available due to privacy restrictions. For further access to the data, please contact the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist.

[PDF File, 2776 KB - games_v12i1e53825_app1.pdf]

Multimedia Appendix 2

Demographic and clinical characteristics of participants based on completion status at baseline.

[DOCX File, 19 KB - games_v12i1e53825_app2.docx]

Multimedia Appendix 3

Mean scores on the symptom subscales for intervention and control groups across study time points and the baseline comparisons.

[DOCX File, 18 KB - games_v12i1e53825_app3.docx]

Multimedia Appendix 4
Generalized estimating equation models for the comparison of outcomes across time between the control and intervention groups.
[DOCX File, 18 KB - games_v12i1e53825_app4.docx]

Multimedia Appendix 5
Percentage frequency of the occurrence of each symptom in the Virtual Reality Symptom Questionnaire.
[DOCX File, 16 KB - games_v12i1e53825_app5.docx]

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49. [doi: 10.3322/caac.21820] [Medline: 38230766]
2. Maass S, Boerman LM, Verhaak PFM, Du J, de Bock GH, Berendsen AJ. Long-term psychological distress in breast cancer survivors and their matched controls: a cross-sectional study. Maturitas 2019 Dec;130:6-12. [doi: 10.1016/j.maturitas.2019.09.003] [Medline: 31706438]
3. Badger T, Segrin C, Dorros SM, Meek P, Lopez AM. Depression and anxiety in women with breast cancer and their partners. Nurs Res 2007;56(1):44-53. [doi: 10.1097/00006199-200701000-00006] [Medline: 17179873]
4. Linden W, Vodermaier A, Mackenzie R, Greig D. Anxiety and depression after cancer diagnosis: prevalence rates by cancer type, gender, and age. J Affect Disord 2012 Dec 10;141(2-3):343-351. [doi: 10.1016/j.jad.2012.03.025] [Medline: 22727334]
5. Ramsenthaler C, Gao W, Siegert RJ, Edmonds PM, Schey SA, Higginson IJ. Symptoms and anxiety predict declining health-related quality of life in multiple myeloma: a prospective, multi-centre longitudinal study. Palliat Med 2019 May;33(5):541-551. [doi: 10.1177/0269216319833588] [Medline: 31060467]
6. Ng CG, Mohamed S, See MH, et al. Anxiety, depression, perceived social support and quality of life in Malaysian breast cancer patients: a 1-year prospective study. Health Qual Life Outcomes 2015 Dec 30;13:205. [doi: 10.1186/s12955-015-0401-7] [Medline: 26715073]
7. Huang X, Li X, Li J, et al. Chemotherapy-induced nausea and vomiting in breast cancer patients: a multicenter prospective observational study. Asia Pac J Oncol Nurs 2021;8(4):433-437. [doi: 10.4103/apjon.apjon-2120] [Medline: 34159237]
8. Cheng VCC, Wong SC, To KKW, Ho PL, Yuen KY. Preparedness and proactive infection control measures against the emerging novel coronavirus in China. J Hosp Infect 2020 Mar;104(3):254-255. [doi: 10.1016/j.jhin.2020.01.010] [Medline: 31962139]
9. Rubin GJ, Wessely S. The psychological effects of quarantining a city. BMJ 2020 Jan 28;368:m313. [doi: 10.1136/bmj.m313] [Medline: 31992552]
10. Lima CKT, Carvalho PMM, Lima I, et al. The emotional impact of coronavirus 2019-nCoV (new coronavirus disease). Psychiatry Res 2020 May;287:112915. [doi: 10.1016/j.psychres.2020.112915] [Medline: 32199182]
11. Duan L, Zhu G. Psychological interventions for people affected by the COVID-19 epidemic. Lancet Psychiatry 2020 Apr;7(4):300-302. [doi: 10.1016/S2215-0366(20)30073-0] [Medline: 32085840]
12. Gallo L, Minutolo A, De Pietro G. A user interface for VR-ready 3D medical imaging by off-the-shelf input devices. Comput Biol Med 2010 Mar;40(3):350-358. [doi: 10.1016/j.compbiomed.2010.01.006] [Medline: 20149912]
13. Mesa-Gresa P, Gil-Gómez H, Lozano-Quilis JA, Gil-Gómez JA. Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: an evidence-based systematic review. Sensors (Basel) 2018 Aug 1;18(8):2486. [doi: 10.3390/s18082486] [Medline: 30071588]
14. Turner WA, Casey LM. Outcomes associated with virtual reality in psychological interventions: where are we now? Clin Psychol Rev 2014 Dec;34(8):634-644. [doi: 10.1016/j.cpr.2014.10.003] [Medline: 25455627]
15. Chan JJI, Yeam CT, Kee HM, et al. The use of pre-operative virtual reality to reduce anxiety in women undergoing gynecological surgeries: a prospective cohort study. BMC Anesthesiol 2020 Oct 9;20(1):261. [doi: 10.1186/s12871-020-01177-6] [Medline: 33036555]
16. Shin H, Kim K. Virtual reality for cognitive rehabilitation after brain injury: a systematic review. J Phys Ther Sci 2015 Sep;27(9):2999-3002. [doi: 10.1589/jpts.27.2999] [Medline: 26504344]
17. Hoffman HG, Chambers GT, Meyer WJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med 2011 Apr;41(2):183-191. [doi: 10.1007/s12160-010-9248-7]
18. Ahmad M, Bani Mohammad E, Tayyem E, Al Gamal E, Atout M. Pain and anxiety in patients with breast cancer treated with morphine versus tramal with virtual reality. Health Care Women Int 2024;45(7):782-795. [doi: 10.1080/07399332.2023.2257627] [Medline: 37703384]
19. Zasadzka E, Pieczyńska A, Trzmiel T, Hojan K. Virtual reality as a promising tool supporting oncological treatment in breast cancer. Int J Environ Res Public Health 2021 Aug 19;18(16):8768. [doi: 10.3390/ijerph18168768] [Medline: 34444513]
20. Wong CL, Li H, Li CK, et al. Effects of immersive virtual reality for alleviating anxiety, nausea and vomiting among patients with paediatric cancer receiving their first chemotherapy: protocol for a randomised controlled trial. BMJ Open 2024 Feb 24;14(2):e079837. [doi: 10.1136/bmjopen-2023-079837] [Medline: 38401901]

21. Sharar SR, Alamdari A, Hoffer C, Hoffman HG, Jensen MP, Patterson DR. Circumplex model of affect: a measure of pleasure and arousal during virtual reality distraction analgesia. Games Health J 2016 Jun;5(3):197-202. [doi: 10.1089/g4h.2015.0046] [Medline: 27171578]
22. Indovina P, Barone D, Gallo L, Chirico A, De Pietro G, Giordano A. Virtual reality as a distraction intervention to relieve pain and distress during medical procedures: a comprehensive literature review. Clin J Pain 2018 Sep;34(9):858-877. [doi: 10.1097/AJP.0000000000000599] [Medline: 29485536]
23. Janssen A, Fletcher J, Keep M, et al. Experiences of patients undergoing chemotherapy with virtual reality: mixed methods feasibility study. JMIR Serious Games 2022 Feb 21;10(1):e29579. [doi: 10.2196/29579] [Medline: 35188474]
24. Moscato S, Sichi V, Giannelli A, et al. Virtual reality in home palliative care: brief report on the effect on cancer-related symptomatology. Front Psychol 2021;12:709154. [doi: 10.3389/fpsyg.2021.709154] [Medline: 34630217]
25. Chirico A, Maiorano P, Indovina P, et al. Virtual reality and music therapy as distraction interventions to alleviate anxiety and improve mood states in breast cancer patients during chemotherapy. J Cell Physiol 2020 Jun;235(6):5353-5362. [doi: 10.1002/jcp.29422] [Medline: 31957873]
26. Vass E, Simon V, Csukly G, Fekete Z, Kis B, Simon L. Virtual reality-based theory of mind intervention in schizophrenia: preliminary efficacy results. Compr Psychiatry 2022 Nov;119:152350. [doi: 10.1016/j.comppsych.2022.152350] [Medline: 36272240]
27. Atef D, Elkeblawy MM, El-Sebaie A, Abouelnaga WAI. A quasi-randomized clinical trial: virtual reality versus proprioceptive neuromuscular facilitation for postmastectomy lymphedema. J Egypt Natl Canc Inst 2020 Jun 15;32(1):29. [doi: 10.1186/s43046-020-00041-5] [Medline: 32537717]
28. Schneider SM, Kisby CK, Flint EP. Effect of virtual reality on time perception in patients receiving chemotherapy. Support Care Cancer 2011 Apr;19(4):555-564. [doi: 10.1007/s00520-010-0852-7] [Medline: 20336327]
29. Schneider SM, Hood LE. Virtual reality: a distraction intervention for chemotherapy. Oncol Nurs Forum 2007 Jan;34(1):39-46. [doi: 10.1188/07.ONF.39-46] [Medline: 17562631]
30. Reynolds LM, Cavadino A, Chin S, et al. The benefits and acceptability of virtual reality interventions for women with metastatic breast cancer in their homes; a pilot randomised trial. BMC Cancer 2022 Apr 2;22(1):360. [doi: 10.1186/s12885-021-09081-z] [Medline: 35366823]
31. Xiao J, Chow KM, Choi KC, et al. Effects of family-oriented dignity therapy on dignity, depression and spiritual well-being of patients with lung cancer undergoing chemotherapy: a randomised controlled trial. Int J Nurs Stud 2022 May;129:104217. [doi: 10.1016/j.ijnurstu.2022.104217] [Medline: 35339908]
32. Jacobsen PB, Donovan KA, Trask PC, et al. Screening for psychologic distress in ambulatory cancer patients. Cancer 2005 Apr 1;103(7):1494-1502. [doi: 10.1002/cncr.20940] [Medline: 15726544]
33. Aaronson NK, Ahmedzai S, Bergman B, et al. The European organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
34. Ames SL, Wolffsohn JS, McBrien NA. The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display. Optom Vis Sci 2005 Mar;82(3):168-176. [doi: 10.1097/01.opx.0000156307.95086.6] [Medline: 15767873]
35. McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. West J Emerg Med 2017 Oct;18(6):1075-1078. [doi: 10.5811/westjem.2017.8.35985] [Medline: 29085540]
36. Ma Y, Mazumdar M, Memtsoudis SG. Beyond repeated-measures analysis of variance: advanced statistical methods for the analysis of longitudinal data in anesthesia research. Reg Anesth Pain Med 2012;37(1):99-105. [doi: 10.1097/AAP.0b013e31823ebc74] [Medline: 22189576]
37. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [doi: 10.1037//0022-3514.51.6.1173] [Medline: 3806354]
38. Yu H, Li H, Zuo T, et al. Financial toxicity and psychological distress in adults with cancer: a treatment-based analysis. Asia Pac J Oncol Nurs 2022 Sep;9(9):100069. [doi: 10.1016/j.apjon.2022.04.008] [Medline: 35669288]
39. Yang T, Lai IKW, Fan ZB, Mo QM. The impact of a 360° virtual tour on the reduction of psychological stress caused by COVID-19. Technol Soc 2021 Feb;64:101514. [doi: 10.1016/j.techsoc.2020.101514] [Medline: 33424061]
40. Baños RM, Espinoza M, García-Palacios A, et al. A positive psychological intervention using virtual reality for patients with advanced cancer in a hospital setting: a pilot study to assess feasibility. Support Care Cancer 2013 Jan;21(1):263-270. [doi: 10.1007/s00520-012-1520-x] [Medline: 22688373]
41. Shahid Khan MU, Hussain HG, Farooq M, et al. The potential of virtual reality for rehabilitation intervention in females with breast cancer. Int J Surg 2022 Oct;106:106940. [doi: 10.1016/j.ijsu.2022.106940] [Medline: 36152921]
42. Sharar SR, Miller W, Teeley A, et al. Applications of virtual reality for pain management in burn-injured patients. Expert Rev Neurother 2008 Nov;8(11):1667-1674. [doi: 10.1586/14737175.8.11.1667]
43. Gershon J, Zimand E, Lemos R, Rothbaum BO, Hodges L. Use of virtual reality as a distractor for painful procedures in a patient with pediatric cancer: a case study. Cyberpsychol Behav 2003 Dec;6(6):657-661. [doi: 10.1089/109493103322725450] [Medline: 14756933]

44. Feyzioğlu Ö, Dinçer S, Akan A, Algun ZC. Is Xbox 360 Kinect-based virtual reality training as effective as standard physiotherapy in patients undergoing breast cancer surgery? Support Care Cancer 2020 Sep;28(9):4295-4303. [doi: 10.1007/s00520-019-05287-x] [Medline: 31907649]
45. Espinoza M, Baños RM, García-Palacios A, et al. Promotion of emotional wellbeing in oncology inpatients using VR. Stud Health Technol Inform 2012;181:53-57. [Medline: 22954828]
46. Harms CA, Cohen L, Pooley JA, Chambers SK, Galvão DA, Newton RU. Quality of life and psychological distress in cancer survivors: the role of psycho-social resources for resilience. Psychooncology 2019 Feb;28(2):271-277. [doi: 10.1002/pon.4934] [Medline: 30380589]
47. Wang B, Hao N, Zhang X. Factors influencing the psychology and quality of life in lung cancer patients. Saudi Med J 2017 Sep;38(9):948-951. [doi: 10.15537/smj.2017.9.18532] [Medline: 28889154]
48. Triberti S, Repetto C, Riva G. Psychological factors influencing the effectiveness of virtual reality-based analgesia: a systematic review. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):335-345. [doi: 10.1089/cyber.2014.0054] [Medline: 24892195]
49. Mayor S. Side-effects of cancer drugs are under-reported in trials. Lancet Oncol 2015 Mar;16(3):e107. [doi: 10.1016/S1470-2045(15)70023-9] [Medline: 25639368]

Abbreviations

CF: cognitive functioning
CONSORT: Consolidated Standards of Reporting Trials
DT: Distress Thermometer
EF: emotional functioning
GEE: generalized estimating equation
ITT: intention-to-treat
PF: physical functioning
QL: global health status or quality of life scale
QLQ-C30 : Quality of Life Questionnaire version 3.0
QOL: quality of life
RF: role functioning
SF: social functioning
VR: virtual reality
VRSQ: Virtual Reality Symptom Questionnaire

Edited by A Bustillo, T Leung; submitted 24.10.23; peer-reviewed by CL Wong, D Hu, M Coccia; revised version received 23.07.24; accepted 16.08.24; published 17.10.24.

Please cite as:
Li M, Yu Z, Li H, Cao L, Yu H, Deng N, Liu Y
Effects of Virtual Reality Therapy for Patients With Breast Cancer During Chemotherapy: Randomized Controlled Trial
JMIR Serious Games 2024;12:e53825
URL: https://games.jmir.org/2024/1/e53825
doi:10.2196/53825

Serious Game Development for Public Health: Participatory Design Approach to COVID-19 Quarantine Policy Education

Myunghwan Kwak[1,*]; Byeong-Je Kim[2,3,*], PhD; Ji-Bum Chung[1], PhD

1

2

3

*these authors contributed equally

Corresponding Author:
Ji-Bum Chung, PhD

Abstract

Background: Public health education plays a crucial role in effectively addressing infectious diseases such as COVID-19. However, existing educational materials often provide only foundational information, and traditional group education faces challenges due to social distancing policies.

Objective: Addressing these gaps, our study introduces a serious game called "Flattening the Curve." This interactive experience immerses learners in the role of quarantine policy managers, offering unique insights into the effects and challenges of social distancing policies.

Methods: The development of the game adhered to the SERES framework, ensuring a scientifically designed foundation. To achieve its learning objectives, the game incorporated learning and game mechanics including an agent-based infection model, a social distancing policy model, and an economic model, which were developed based on previous literature. After defining a broad concept of scientific and design foundations, we used a participatory design process. This study included 16 undergraduates and took place over one semester. Participants played the game, gave feedback, and answered surveys. The game was improved based on participants' feedback throughout the process. Participants' feedback was analyzed based on the Design, Play, and Experience framework. Surveys were conducted before and after the activity and analyzed to assess participants' evaluation of and satisfaction with the game.

Results: The game successfully achieved its learning objectives, encompassing a comprehensive understanding of infectious disease characteristics; the disease transmission process; the necessity and efficacy of quarantine policies and their delicate balance with economic factors; and the concept of flattening the curve. To achieve this, the game includes the following: (1) an agent-based infection model based on the modified Susceptible-Exposed-Infectious-Hospitalized-Recovered (SEIHR) model with five infectious disease scenarios; (2) a quarantine policy model with social distancing, travel control, and intensive care unit management; and (3) an economic model that allows users to consider the impact of quarantine policies on a community's economy. In response to participatory design feedback, the game underwent meticulous modifications, including refining game systems, parameters, design elements, the user interface, and interactions. Key feedback included requests for more scenarios and engaging yet simple game elements, as well as suggestions for improving the scoring system and design features. Notably, concerns about the fairness of the outcome evaluation system (star rating system), which could incentivize prioritizing economic activity over minimizing casualties, were raised and addressed by replacing the star rating system with a progress-based vaccine development system. Quantitative evaluation results reflect participants' positive assessments of the game through the learner-centric approach.

Conclusions: The serious game "Flattening the Curve," developed through a participatory design approach, emerges as a valuable tool for public health education, particularly concerning social distancing policies. The game and its source code are openly accessible online, enabling widespread use for research and educational purposes.

(JMIR Serious Games 2024;12:e54968) doi:10.2196/54968

KEYWORDS

online learning; serious game; simulation; quarantine policy; social distancing policy; game; public health education; infectious diseases; learner-centric

Introduction

Background

Since December 2019, the COVID-19 pandemic has profoundly impacted global society, affecting numerous domains such as daily life, health care, and education, among others. In an effort to mitigate the impacts of COVID-19, governments worldwide have implemented a range of social distancing measures, including compulsory mask wearing, curfews, stay-at-home orders, and lockdowns. However, the implementation of these policies has led individuals to experience economic, social, and psychological loss, resulting in diminished trust in their respective governments [1,2]. This sense of loss and eroded trust are fueling public rejection of social distancing policies and even leading to active defiance of governments. Indeed, several countries experienced increased public protests due to the social distancing policies [3-5].

People's noncompliance with social distancing policies can be attributed to various reasons. Previous research suggests that these reasons include a lack of knowledge about infectious disease and health, misunderstanding, and misinformation about mitigation policies [6-9]. To effectively prevent and control the transmission of infectious diseases such as COVID-19, adherence to public health policies, including quarantine measures, is crucial. Additionally, fostering public comprehension of the necessity and rationale underpinning these policies is of paramount importance [3].

To achieve these objectives, effective communication with the public and the provision of education aimed at enhancing the general public's understanding of policy validity and scientific knowledge are necessary. Currently, educational resources pertaining to infectious diseases predominantly focus on health-related information and individual behaviors, while the development of materials designed to foster comprehension of national- or community-level social distancing policies remains insufficient. Moreover, the majority of training materials employ a unilateral instructional approach, relying on simplistic documents, images, and videos. These conventional methods generally prove inadequate in fostering public empathy toward the policy-related challenges confronted by quarantine officers. On the other hand, game-based learning using serious games is known to increase education effectiveness and learning motivation through the interactive and participatory nature of games [10-12].

Due to the COVID-19 pandemic, the demand for health and safety education has increased, and many researchers have developed and used serious games for COVID-19–related education. The primary education topics include COVID-19 knowledge [13-15] or infection prevention behaviors and perceptions [10,15-18]. Since serious games are suitable for distance education and e-learning (posing less of a concern in terms of infection), they are an educational method that has been in the spotlight for many researchers in the pandemic era, coinciding with an increase in online education [19-22].

Earlier studies found that game-based learning could result in higher knowledge acquisition [12,23] and, more importantly, that games could make their players motivated and satisfied [11,23]. Digital game–based learning enabled by serious games can make a learning environment more active, engaging, and immersive, which is important for effective learning [24]. However, most game-based health and safety education focused on teaching safe behavior, knowledge, and adequate use of personal protective equipment to prevent COVID-19 infection. The transfer of mere knowledge has limitations in imparting an understanding of why compliance with quarantine policies is necessary and how the policies work. Therefore, by using existing education, learners were not properly motivated to go beyond personal safety behaviors and comply with quarantine policies. A serious game for quarantine policy education will be a fundamental driving force for the public to learn why the social distancing policy must be followed rather than just how to comply with it.

Objective

The objective of this study was to address a critical gap in existing educational resources for quarantine policies by developing an engaging and interactive learning tool. To achieve this objective, the game incorporates scientifically designed game-based learning mechanics including an agent-based infection model, a social distancing policy model, and an economic model. Furthermore, a participatory design (PD) process was conducted, with this learner-centric approach enhancing the educational effectiveness of the game. This study endeavors to enhance public understanding of quarantine policies by disseminating the developed serious game through web platforms, thereby promoting informed decision-making and adherence to quarantine measures.

Methods

General Study Design

To effectively deliver quarantine policy education addressing infectious diseases to the public, we developed a serious game named "Flattening the Curve" (FTC). The serious game used in this study was designed in a form that could motivate learners' interest while conveying scientific facts. The game was developed using the SERES framework [25]. The SERES framework is a 5-step framework for developing health care serious games. It has been applied to the development of various health care serious games and gamified e-learning materials.

In addition to using the SERES framework, the game was developed through a PD process to enhance the engagement and motivation of the learners [26,27]. PD is a research method in which users directly or indirectly participate in the design process and compose and interpret the design together with designers, researchers, and developers [28,29]. From an educational point of view, PD promotes learners' active engagement through motivation and internalization of concepts [30,31].

The design process was conducted as part of the undergraduate disaster management course at the Ulsan National Institute of Science and Technology, which was taught by the corresponding author during the second semester of 2021. Participants participated from the 2nd to the 15th week of the semester. To recruit participants, the authors explained to the participants the study's purpose, overall process, and tasks to be performed. The authors offered them the opportunity to participate in this activity as part of their coursework. All students voluntarily agreed to participate in the study.

This PD process reflected the iterative game design process, including design, prototype, and playtest [32]. Participants' comments and suggestions were analyzed and applied to the game using the 4 layers of the Design, Play, and Experience (DPE) framework [33]. During the PD process, participants answered two surveys, in the 3rd week and 14th week, respectively. Participants' key feedback and related implementation are presented in the Results section, and the details of the survey instrument are also presented in Multimedia Appendix 1.

Ethical Considerations

This study was reviewed by the Institutional Review Board at the Ulsan National Institute of Science and Technology and was determined to be exempt (UNISTIRB-21 - 58C). Due to the COVID-19 pandemic and the shift to remote learning, obtaining written informed consent was difficult. Therefore, verbal consent was obtained from all participants prior to their involvement in the study. All collected data were deidentified. There was no penalty for students who did not participate in part or all of the activity. Participants did not receive monetary compensation. However, to encourage active participation, participants who performed well in the activity were awarded extra credit for the course. This compensation was announced after all students had agreed to participate in the study, to ensure that the incentive did not coerce participation.

Scientific Foundations

The game was developed for people of all ages to facilitate an understanding of quarantine (social distancing) policies. Emphasis was placed on ensuring that the in-game explanations are easily comprehensible for players of all ages.

Understanding how and why a policy is important can lead to behavior changes in learners [34,35]. According to theories of human behavior related to safety and health, such as the Health Belief Model, the perceived benefits expected from engaging in certain behaviors can impact an individual's behaviors [36]. In other words, a thorough understanding of how quarantine policy benefits individuals and our society can influence an individual's compliance with policies. Thus, the primary learning objective of the game was for users to develop an understanding of the core concept of quarantine policies—"flattening the curve"—and recognizing the need for a balanced adjustment between preventive measures and economic consequences. Figure 1 illustrates the concept of flattening the curve, demonstrating how implementing social distancing measures can flatten the curve and consequently reduce the fatality rate [37]. In the absence of proper protective measures, the number of patients could easily overwhelm the health care system's capacity, resulting in inadequate treatment for some. Therefore, the point of flattening the curve is to ensure that more patients receive proper treatment, ultimately reducing the fatality rate [38]. Within the simulation, changes in the transmission pattern of infectious diseases can be observed according to policy implementation.

The game mechanics should be designed based on proper scientific foundations. Several models were suggested to predict the spread of the epidemic and the effect of social distancing policy [39-41] (Figure 2). Kermack and McKendrick [42] first proposed the Susceptible-Infectious-Recovered (SIR) model to explain the spread of infections transmitted from person to person through mathematical approaches. SIR is the most fundamental infectious disease model, consisting of Susceptible, Infectious, and Recovered or Removed compartments. Anderson and May [43] introduced the SEIR model, which includes an additional Exposed condition (latent period or incubation period). SEIHR is a model that extends the SEIR model by incorporating an isolated (Hospitalized) condition [40]. Further, Yang et al [41] suggested another modified version of SEIR: SEMCR, which divided the infectious stage into mild (M) and critical (C) cases. Patients in the critical stage have a higher probability of death.

Figure 1. An illustration of the concept of “flattening the curve.” Slowing the spread of infection is essential to ensure that people have easier access to treatment, while also allowing time to gradually increase the capacity of the health system.

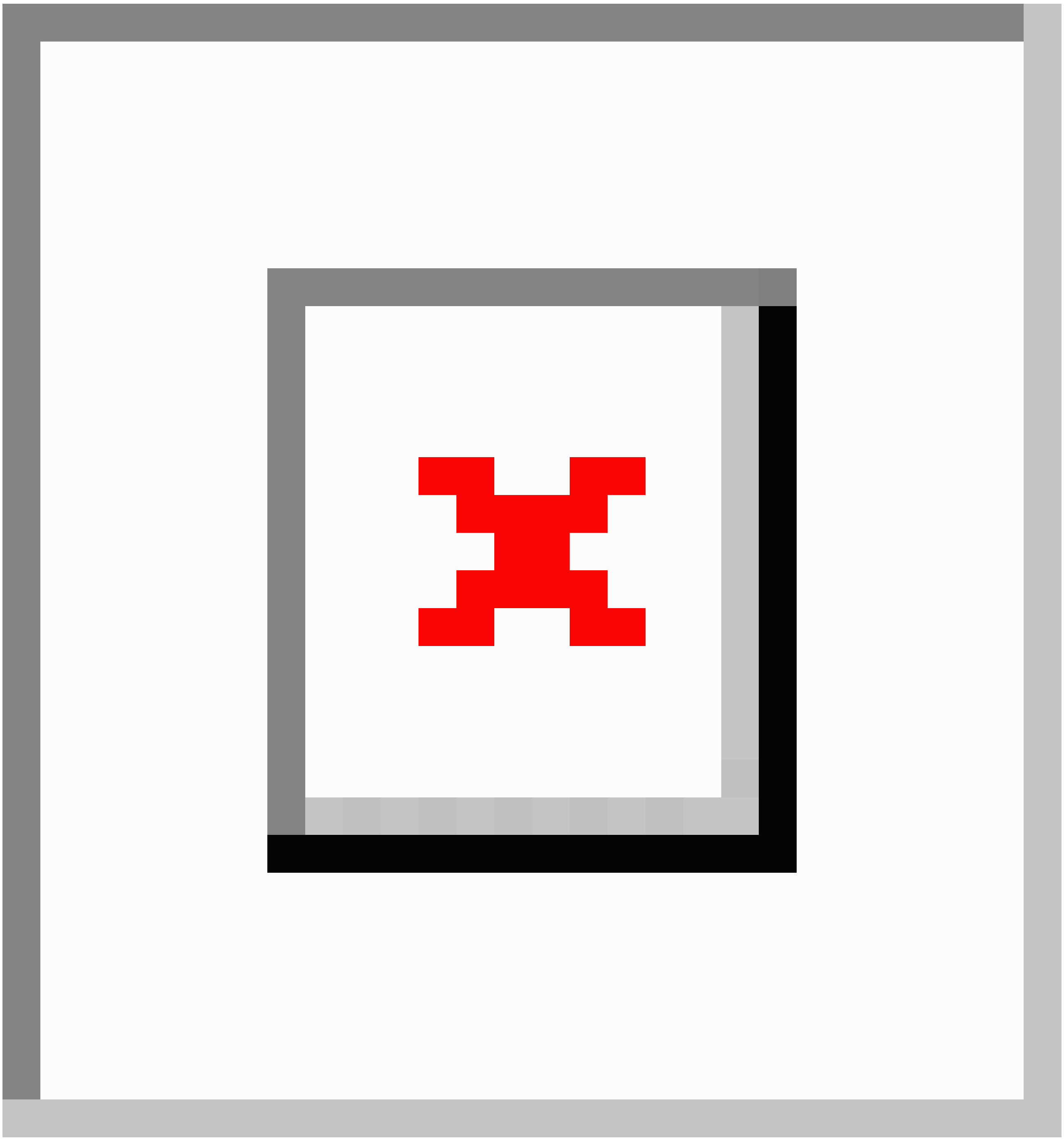
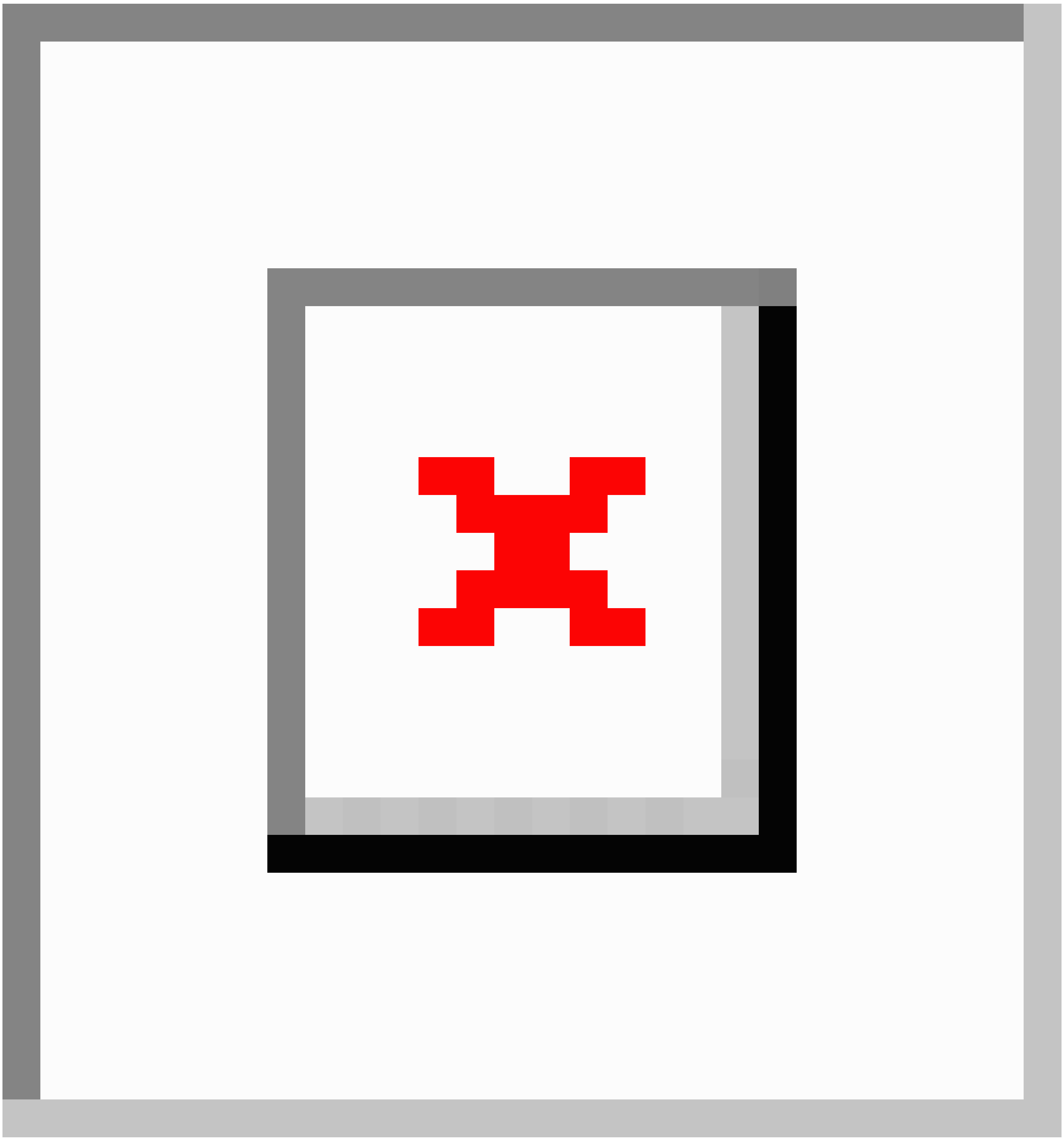


Figure 2. Flowcharts of epidemic models, edited from previous studies [39-41]. C: critical; D: dead; E: exposed or latent period; H: hospitalized; I: infectious; M: mild; R: recovered; S: susceptible.



In the FTC game, we applied a modified epidemic model by integrating SEIHR [40] and SEMCR models [41] to distinguish between mild and critical cases and also include the hospitalized condition.

Based on the modified epidemic model, the game was designed to model various well-known infectious diseases, such as COVID-19 (Delta and Omicron variants), measles, and Ebola virus. Major characteristics of diseases, such as reproduction number and mortality rate, were adjusted based on statistical data [44-50]. Characteristics of COVID-19 variants were based on statistics provided by the Korea Disease Control and Prevention Agency [51]. The implemented infectious disease transmission model has been consulted and verified by a medical expert.

Social distancing is a public health strategy that aims to reduce the spread of infectious diseases [52]. It involves maintaining physical distance from others, avoiding large gatherings, and minimizing contact with individuals who are sick.

Researchers have employed various methods to implement social distancing in epidemic models. Social distancing is typically achieved by increasing the physical distance between people or reducing the number of physical contacts between individuals [41,53,54]. Agarwal et al [55] utilized daily movement rates to predict the number of deaths per day to evaluate the effectiveness of COVID-19 social distancing policies. Miksch et al [53] used location factors to make social distancing decrease the number of visitors to specific places such as schools and workplaces.

However, many educational materials related to infectious diseases have not adequately addressed the economic damage caused by the infectious diseases and the corresponding quarantine policies. Various social, health, and sanitary policies to prevent the spread of infectious diseases have been found to bring about negative economic impacts [56]. In the case of the United States, government measures such as business closures and restrictions on economic activities to slow the spread of diseases led to a record high in unemployment claims [57]. Therefore, we built economic elements into the game to help players learn about the economic impacts of quarantine policies.

Design Foundations

The general design of the serious game was established according to Winn's DPE framework [33]. At the initial stage of game development, we decided what types of learning and game mechanics should be included in the game based on the previous literature [58-60]. The game genre was decided to be a simulation, as its learning mechanism focuses on constructivist learning principles [60], which posit that learning draws upon prior knowledge to alter existing schema [61]. Using simulations, students have more opportunities to learn by doing, discover interesting infectious disease problems, and gain indirect experience in dealing with public health issues in the real world [62].

Design requirements considered the platform and accessibility to potential learners. The game was developed using JavaScript and published through a web page to be accessible to anyone. Following the players' feedback, visual enhancements were implemented to enhance the game's appeal to a younger audience.

Results

Design of the Serious Game

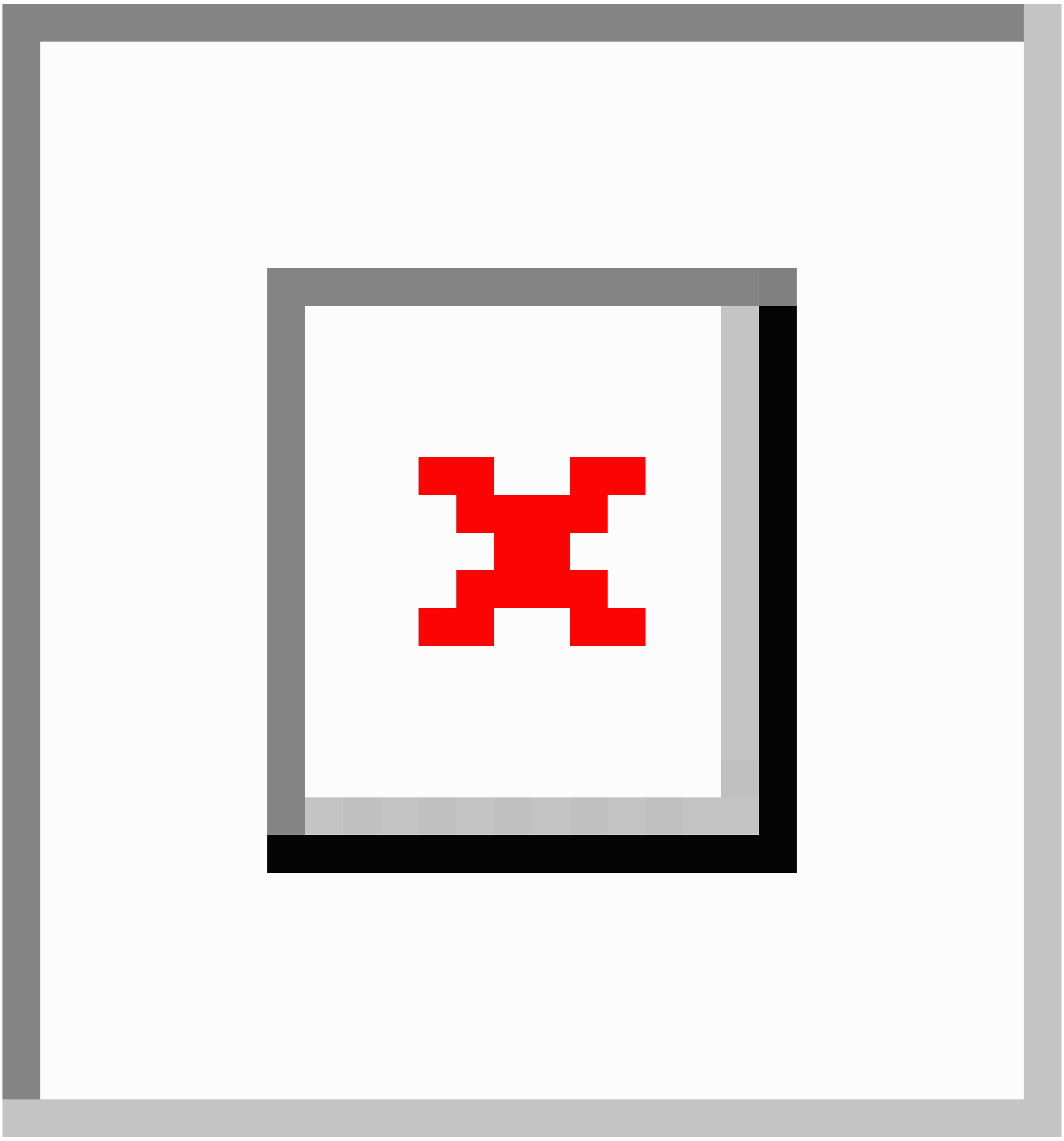
The game was developed as a turn-based strategic simulation game, with each turn corresponding to one week. In our interactive serious game, learners become a quarantine policy manager of a community. At the end of each week, users are allowed to choose appropriate quarantine measures. The spread of the epidemic is visualized on the game screen during the following week, and on weekends, users can check the status of the epidemic, apply quarantine measures, and continue on to the next turn. Players are provided with a budget necessary for implementing quarantine policies, which replenishes at the end of each month (every 4 weeks). The budget amount depends on the level of economic activity, as indicated by gross domestic product (GDP).

The objective of the gameplay is to minimize both casualties and economic damage until vaccine development is complete. Since the vaccine development speed depends on both the number of healthy people and GDP, faster vaccine development can be an indicator of a successful quarantine response.

To slow down the spread of disease, players should develop their own response strategies with various policies. These policies can be adjusted or changed on a weekly basis and each has its own advantages in reducing the spread of the disease. However, it is important for players to find a cost-effective combination of policies, as some have drawbacks (eg, social distancing policies may hinder economic activity, while policies such as travel control or increasing intensive care unit [ICU] capacity may require significant economic investments).

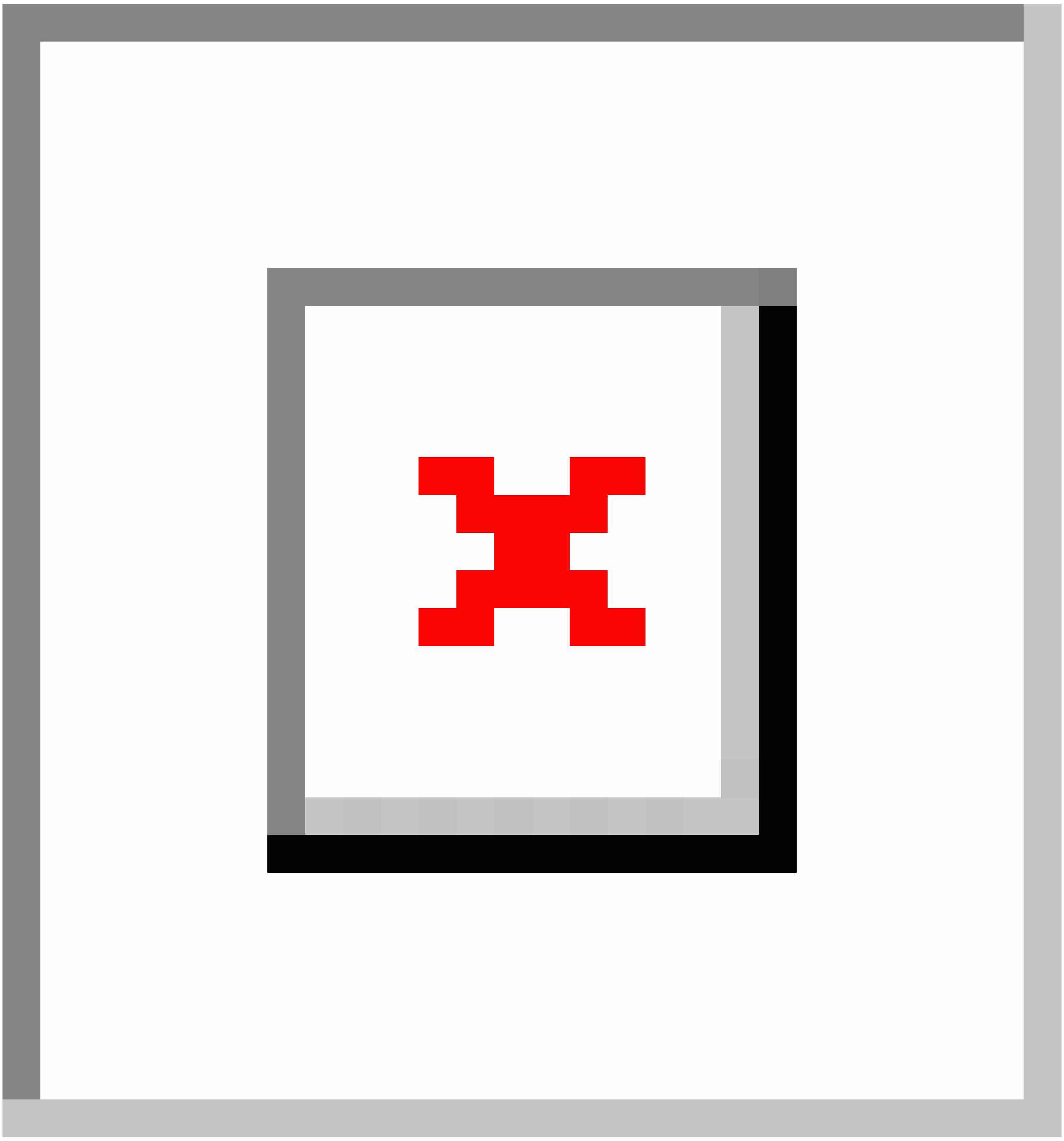
The game comprises 3 main panels: the simulation panel, information panel, and status panel (Figure 3). In the simulation panel, individuals are represented as round dots that move randomly. The color of each dot indicates its condition based on the modified SEIHR model. The information panel displays essential details about infectious diseases, such as the latent period, infectious period, infection rate, and severity. It also shows the number of people in each infectious stage, categorized according to the modified SEIHR model. The bottom of the panel features a graph illustrating overall disease trends.

Figure 3. Interface of the final version of the Flattening the Curve game. (A) Simulation panel: visualizes the agent's movements within the game environment. (B) Information panel: displays the current epidemic situation. At the end of each week, this panel also presents the weekly report. (C) Status panel: provides a summary of the key information. This panel also includes speed control and pause functions to allow players to adjust the pace of the game. GDP: gross domestic product; ICU: intensive care unit.



At the end of each week, the information panel transforms into the weekly report panel (Figure 4). Within the weekly report panel, there are two subtabs: Trends and Policy. The Trends tab includes tables and graphs that present the number of infections, critical cases, and deaths, as well as GDP, for the current week, providing a comparison with the previous week. This feature assists users in understanding more nuanced infection trends. In the Policy tab, players can implement age-based social distancing policies, implement travel control measures, and enhance ICU capacity.

Figure 4. Subtabs in the weekly report panel. (A) Trends tab. (B) Policy tab. GDP: gross domestic product; ICU: intensive care unit.



The status panel at the bottom consistently displays key information, including the current date, number of infections, ICU capacity, death toll, GDP, and budget status.

We made some important basic assumptions that make the game more explicit and learning oriented. These assumptions are carefully designed to ensure that the game retains its inherent learning capabilities and is sufficiently representative of the real world without becoming too complex. The assumptions are as follows.

1. Fixed population: there are 1000 people in the simulation at the start, with no population change due to death, birth, or migration outside of the epidemic.
2. Simplified interactions: a "contact" has occurred when dots overlap each other on the simulation panel.
3. Fixed border: the game world is assumed to have fixed rectangular boundaries and is divided into four in-game districts, whose geographic and social conditions are identical. This exists to determine the coverage of social distancing policies. There are no spatial features such as specific locations, workplaces, or homes.
4. Simplified economy: GDP is only influenced by how much people move around during the week. If social distancing policies are tightened or the number of deaths is high, people move around less, so the budget that the player receives decreases.

5. Simplified policy: quarantine policies, including the allocation of ICU beds, come into effect instantly, without any waiting or grace periods. In-game citizens do not raise objections to policies.

Application of the Scientific Foundation

Overview

The ultimate objective of the game is to comprehend the concept of flattening the curve, necessitating an understanding of infectious disease characteristics and control policies. There are 5 learning objectives in the game to achieve this. Figure 5 demonstrates a visualization of the learning objectives covered in the developed game.

The first learning objective focuses on understanding the characteristics of infectious diseases. The game incorporates characteristics of infectious diseases, such as the incubation period, transmission probability, and fatality rate. With the modified SEIHR model, the game simulates detailed infection and progression pathways. This interactive approach allows learners to gain insights into the distinct features and stages of various infectious diseases through gameplay of various scenarios.

The second learning objective involves understanding the process of disease transmission within society. The game takes the form of a city management simulation, effectively portraying the process of disease transmission as people move within the city.

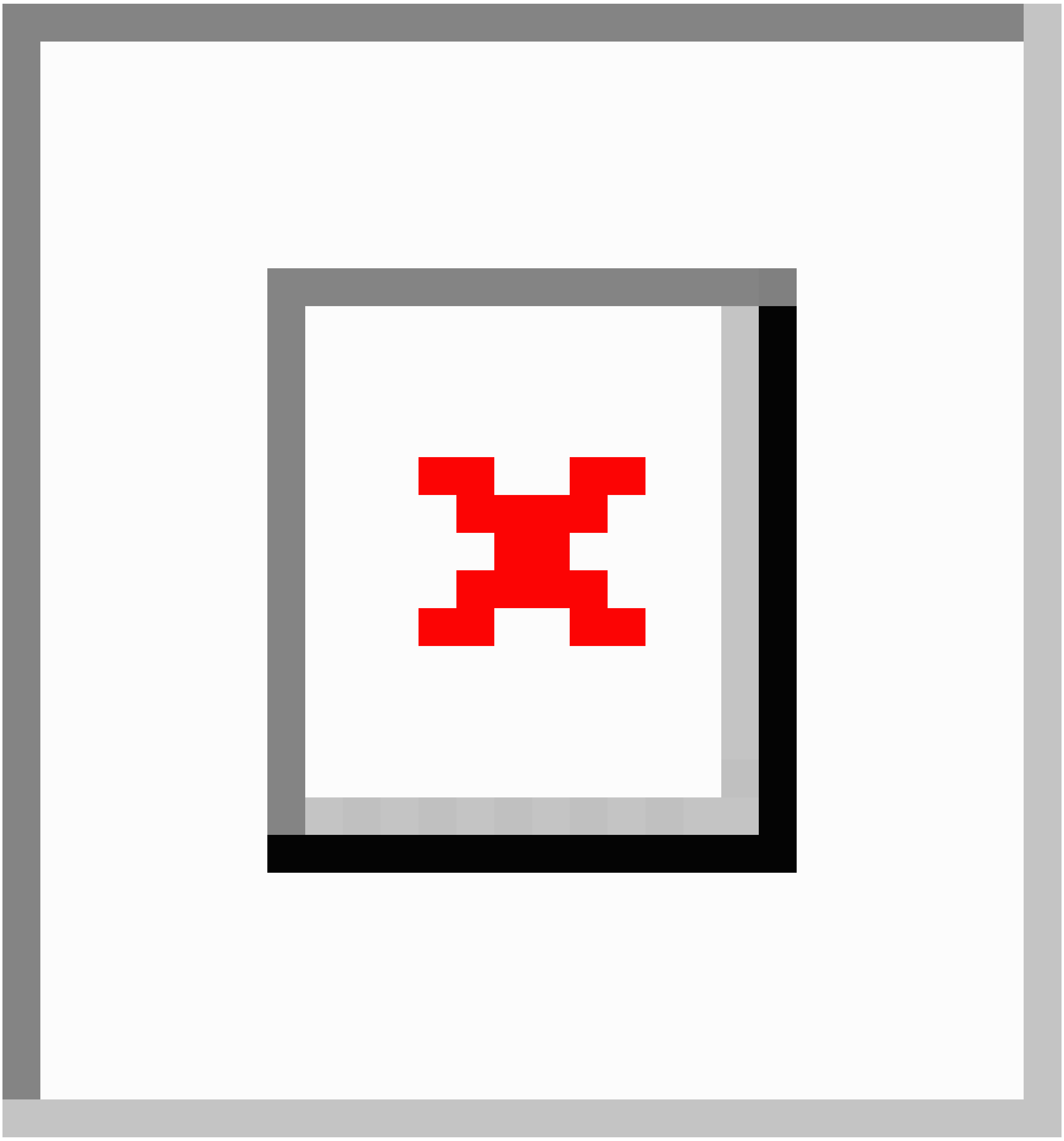
Building upon these objectives, the third learning objective emphasizes the necessity and efficacy of quarantine policies. During gameplay, learners are expected to actively participate in the implementation of quarantine policies in response to changing circumstances, thus developing a comprehensive understanding of the need for and effectiveness of policies.

The fourth learning objective addresses the balancing act between disease control policies and economic factors, illustrating the practical challenges of policy implementation. To convey the practical consequences of implementing quarantine policies, the game introduced economic concepts such as GDP and a budget, underscoring the imperative to strike a balance between epidemic control measures and economic considerations.

By building upon these 4 learning objectives, learners will eventually understand the concept of “flattening the curve” and gain the knowledge and skills to effectively manage infectious diseases.

All these learning objectives were delivered by scientifically designed game modules. The following sections describe how the modules were implemented in the game.

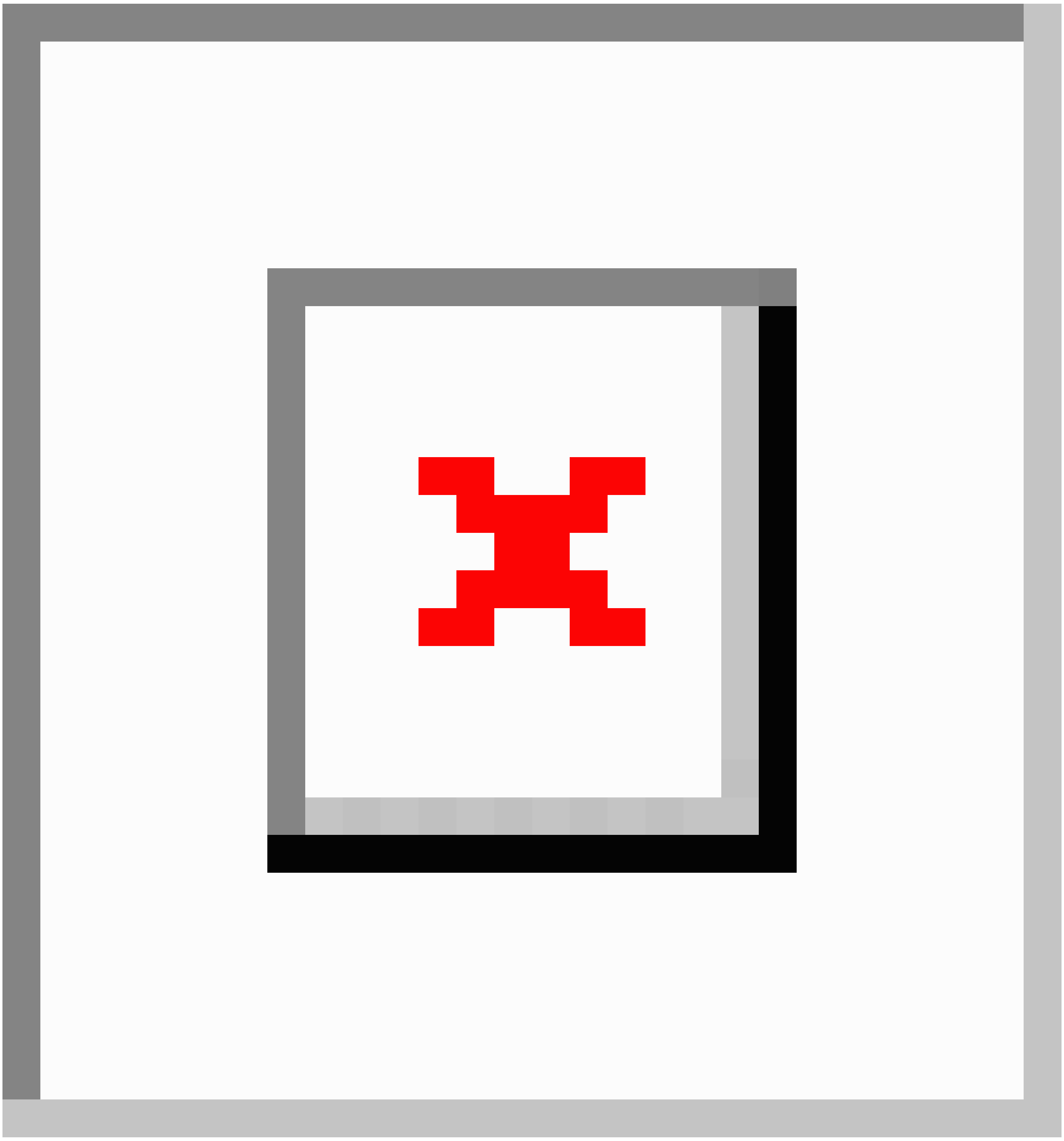
Figure 5. Learning objectives and corresponding game elements in the game. FTC: flattening the curve; GDP: gross domestic product; ICU: intensive care unit.



Infection Model

The modified SEIHR model implemented in the game consists of distinct stages: Susceptible (S), Exposed (E, latent period), Infected (I), Hospitalized (H), Recovered (R), and Dead (D) (Figure 6). When a noninfected agent (S) contacts a symptomatic agent (E2 or I1), they enter the latent period (E) based on infection probability. After a certain duration, the infection progresses to symptomatic mild stage (I1).

Figure 6. In-game description of the modified SEIHR model. ICU: intensive care unit; SEIHR: Susceptible-Exposed-Infectious-Hospitalized-Recovered.



In the earlier exposed phase (E1), agents do not have symptoms and are not infectious. Over time, they change to an exposed state with infectiousness (E2). Infected individuals are categorized into 2 groups based on the severity of symptoms. Mild patients (I1) can infect others and eventually recover (R) after a designated period. Some patients may advance to severe cases (I2) based on predefined probabilities linked to disease characteristics and age. If an ICU is available, severe patients can be isolated (H) and recover (R); if ICU isolation is not possible, these patients face mortality (D). The initial ICU bed capacity is limited and treatment in the ICU takes a considerable amount of time. Therefore, players are prompted to use their budget to expand ICU capacity. The ages and severity rates of agents in the game are distributed similarly to real-world statistics based on relevant data provided by the Korea Disease Control and Prevention Agency [51] (Table 1).

In addition, we have incorporated various infectious disease scenarios to facilitate a natural understanding of disease characteristics. Table 2 shows the latent and infection period, average infection rate, and morbidity of each disease implemented in the game. The included infectious diseases are COVID-19 and its variants (Delta and Omicron), measles, and the Ebola virus. This allows coverage of a spectrum of infectious diseases, ranging from those with low mortality rates and high infection rates to representative diseases with higher mortality rates and slower transmission.

Table . Population distribution and severe morbidity by age group (based on COVID-19, the default scenario).

	Age range (years)								
	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	≥80
Distribution, %	8	10	13	13	16	18	13	5	2
Morbidity, %	1.6	1.6	1.6	3.7	3.7	13.4	27	54	54

Table . Infectious disease scenarios and major characteristics.

Disease type	Latent period (days)	Infection period (days)		Infection rate, %	Morbidity, %
	S to E1[a]	I1 to R[b]	I1 to I2[c]		
COVID-19	3 - 4	19	4 - 5	5	12
COVID-19 (Delta variant)	3 - 4	19	4 - 5	13	14
COVID-19 (Omicron variant)	3 - 4	19	4 - 5	58	3
Ebola	1 - 2	18	3 - 4	2	55
Measles	3 - 4	17	2 - 3	17	0

[a]S to E1: Susceptible to Exposed (earlier, noninfectious stage).

[b]I1 to R: Infectious (mild case) to Recovered.

[c]I1 to I2: Infectious (mild case) to Infectious (critical case).

Quarantine Policy Model

Every 7 days, players can choose an appropriate quarantine policy based on the current status of disease spread. In the game, social distancing, travel control, and ICU expansion policies are implemented.

Social distancing policies are implemented based on the "containment and closure" category policies, according to the COVID-19 government response tracker from Oxford [63]. This category includes measures such as school closures, workplace shutdowns, cancellation of public events, restrictions on gatherings, public transport closures, stay-at-home orders, limitations on movement between cities or regions, and restrictions on international travel. As simplified by participants' feedback, the game only incorporates school closures (affecting those under 20 years), workplace shutdowns (affecting those between 20 and 50 years), and lockdowns (affecting all age groups). Although social distancing policies effectively control the spread of infectious diseases, they inevitably reduce overall public mobility, leading to a decrease in economic activity. This can constrain budgets for other control measures and potentially slow vaccine development. Consequently, players must carefully consider the long-term trade-offs when implementing social distancing strategies.

With the travel control policy, players can build walls at the boundaries of selected areas to prevent travel across those boundaries. This effectively restricts people within a zone from traveling to another zone, but it requires a weekly budget to maintain the walls. Furthermore, the policy has been designed to not entirely restrict travel between zones in order to accommodate for the restrictions of the real-world travel control policy. As a result, the effectiveness of the policy may vary depending on the pattern of infection spread. This requires players to carefully consider the spatial distribution of infections and strategically implement this policy to optimize timing and location to maximize impact.

ICU expansion permanently increases the number of available ICU beds through budget spending. The concept of flattening the curve is implemented through this. Players should slow the spread of infectious diseases through quarantine measures while continuously investing in ICUs to increase medical capacity and reduce deaths [64]. Expanding ICU capacity is crucial but expensive. As disease transmission stabilizes and ICU demand declines, some additional beds may become underutilized. This highlights the need for players to consider the long-term cost-effectiveness of such investments.

Economic Model

Although quarantine policies prove effective in curtailing the spread of infectious diseases, they carry the side effect of impeding economic activities [56,57]. To enhance learners' understanding of how quarantine policies impact a community's economy, we integrated a GDP-based economic module into the game. Economic activity within the game was designed to be correlated with the movement speed of agents, setting GDP as the total accumulation of economic activities over a week (ie, the sum of movement distances). The movement speed of each agent is contingent on factors such as age group, quarantine policies, and the prevailing infection situation. The economically active population aged 20 - 59 years moves at a faster pace than other age groups, with mild patients experiencing a 10% reduction in movement speed and severe patients being unable to move until recovery.

GDP functions as an indicator gauging the overall level of local economic activity, and players are allocated a budget each month proportional to the GDP of the preceding 4 weeks. This budget is essential for implementing travel control policies and expanding ICU capacity. Users face dilemmas, pondering whether to accept economic losses through stringent social distancing policies to slow the spread of infection or to mitigate economic damage through lenient social distancing policies while securing crucial budgets for augmenting ICU capacity.

Game Development With PD

To enhance participants' engagement and motivation when playing the serious game, we employed PD and analyzed learner feedback based on the DPE framework [33]. All 16 students in the class took part in the PD process. Throughout the activities, participants played the game, shared their experiences and opinions, and provided recommendations, which were mostly incorporated into the final version. Following the DPE framework, participants' feedback was categorized according to subcomponents of serious game design: learning, storytelling, gameplay, and user experience (Table 3).

Table . Participants' key feedback and related implementation examples.

User feedback by Design, Play, and Experience layers		In-game implementation
Learning		
	• Users can understand the dynamics of disease spread through visual simulations. • Learning about the differences in disease spread patterns based on virus characteristics. • Understanding the necessity and effectiveness of various public health policies. • Balancing adjustments are needed (eg, enhancing the effectiveness of social distancing policies).	Game parameters such as default movement speed, maximum impact of social distancing policies, and amount of budget were adjusted to achieve optimal game balance.
Storytelling		
	More scenarios are desired.	Implementation of various infectious diseases beyond COVID-19.
	Immersion in the narrative is hindered by the star rating system.	Addition of a vaccine development system to replace the scoring system.
Gameplay		
	New systems like locations and behavior patterns are needed.	Introduction of a 4-zone travel control system.
	There are too many elements requiring complex intervention.	Simplification of policy user interface.
User experience		
	User interface and design improvements are needed.	Design rework and implementation of a speed control feature.
	Lack of guidance on how to use the game (need for a tutorial) and need to add feedback elements like alerts and pop-ups.	Pop-ups for budget system and implementation of Key Facts.

In the learning component, participants generally agreed that the game enabled them to learn the intended learning objectives. However, in the initial version, feedback arose regarding the game's balance, causing issues like uncontrollable spread of disease or overly strong social distancing effects. To address these problems and enhance learning, efforts were made to balance the game, including lowering default agent movement speed and reducing the maximum impact of social distancing policies on movement speed.

Regarding the storytelling component, even though there was no direct narrative or story in the game, participants immersed themselves in the game by adopting the background narrative of being the "public health manager." Feedback suggested enhancing the narrative, leading to the introduction of various types of infectious disease scenarios that players could select at the beginning of the game.

In addition, the initial version of the game ended when the infectious disease was eradicated, and game results were rated on a 10-star scale based on cumulative confirmed cases (up to 1 star), deaths (up to 4), economic activity levels (up to 3), and remaining budget (up to 2). However, this star rating system impeded the players' empathy and made gameplay time too long. Further, a fairness issue [65] arose as players prioritized economic activity over managing deaths to gain more stars. To solve these problems, a new clear condition (vaccine development) was introduced. The progress of vaccine development increases each week based on the number of survivors without severe symptoms and the amount of GDP. Players can expedite vaccine development by focusing on minimizing casualties and economic damage. The final version of the game ends when the vaccine development progress reaches 100% or a year passes.

Regarding the gameplay components, participants mainly proposed new features such as events in new locations, gatherings, and activities in specific areas of the game. However, considering the potential excessive influence of these features on the game's outcomes, seen as too random, we have introduced a new social distancing policy, the travel control system (Figure 7). In the later stages of the PD process, after more game features had been implemented, opinions regarding the simplification and improvement of existing features were added. For example, players were originally able to manually set the reduction in movement speed for each age group at each policy stage and activate the policies. However, as players perceived this feature to be too complex and began to underutilize it, we opted for a user interface (UI) enhancement, replacing the user customization feature with a simpler on/off mechanism for policy activation (Figure 8).

Figure 7. (A) The image illustrates the implementation of travel control measures, where red lines indicate blocked movement between zones. (B) An in-game description of travel control measures.

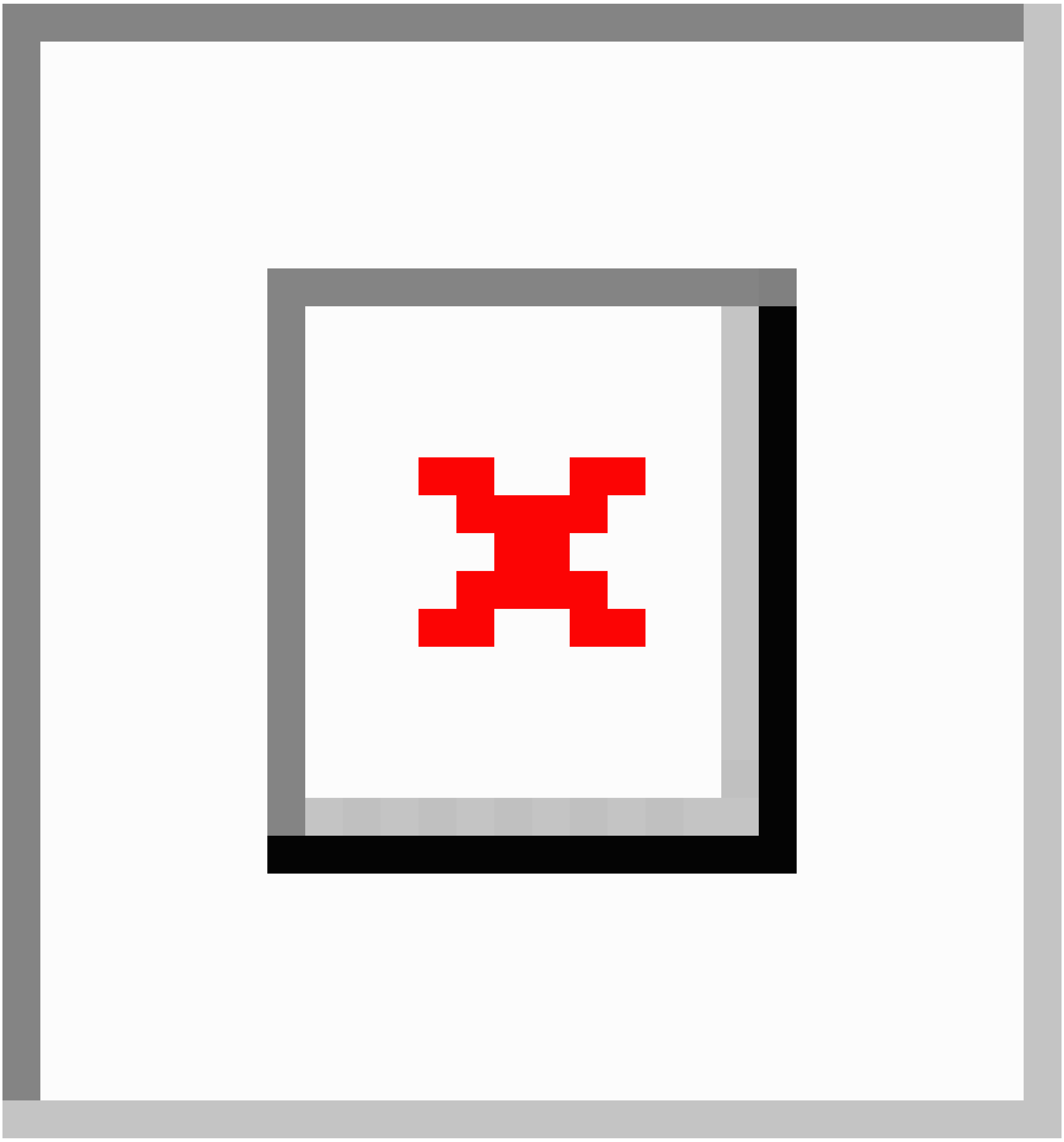
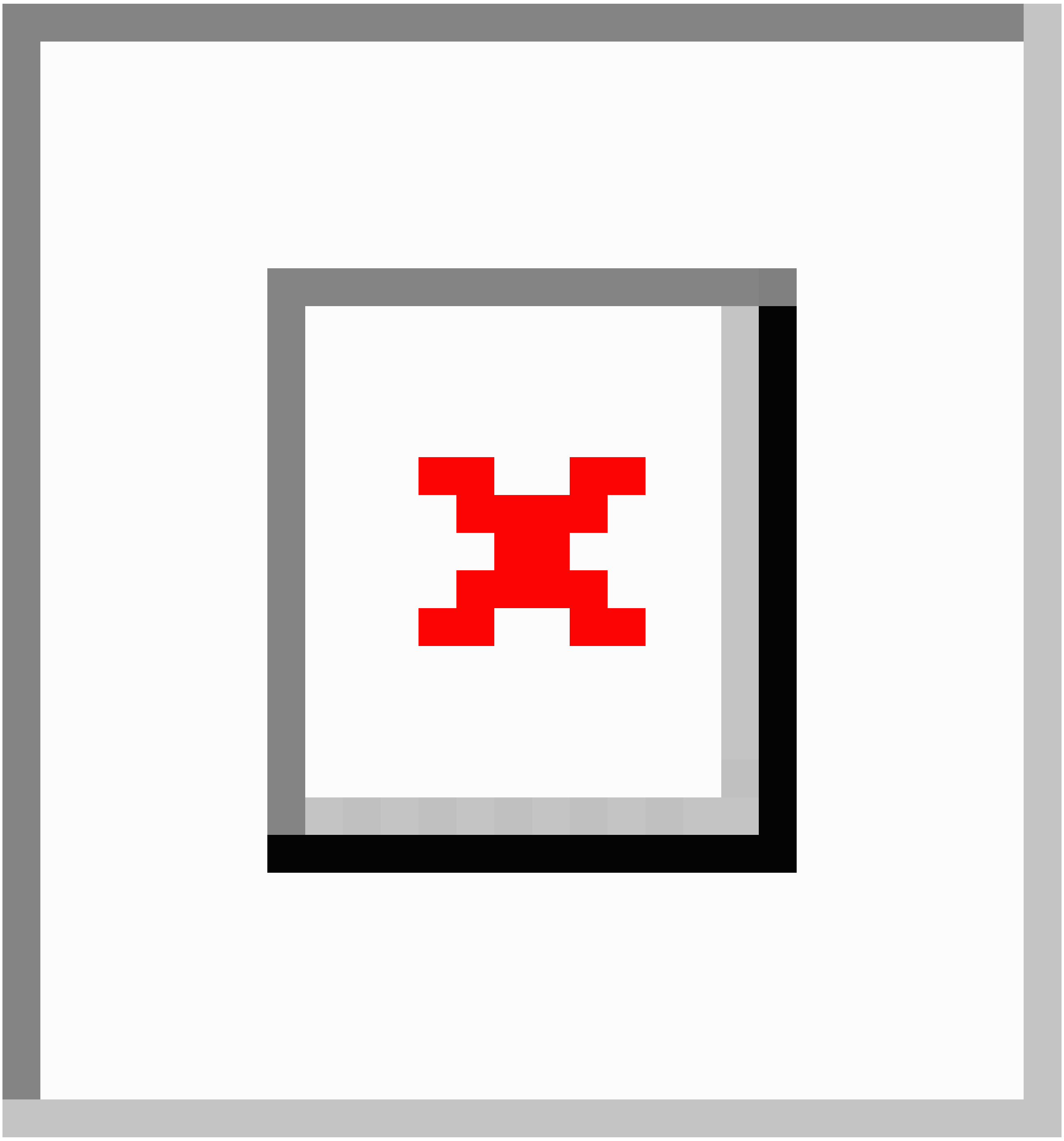


Figure 8. Example changes to improve UI for the social distancing policy. (A) The earlier version UI. (B) The final version UI. UI: user interface.



In terms of the user experience component, there were suggestions for improving the game's UI and design, as well as incorporating in-game feedback and tutorials. Responding to participants' opinions, we collaborated with design experts to completely rework the game design (Figure 9). To prevent monotony during repeated gameplay, a speed-up function was added, allowing users to accelerate simulation progress by up to 8 times. In response to feedback, budget notifications and pop-up alerts for budget shortages were implemented. Additionally, to assist new players in understanding and learning the game, we introduced a Key Facts tab providing essential information. Clicking on hint icons placed near elements requiring explanation adds the corresponding information to the Key Facts tab, allowing players to access explanations whenever needed (Figure 10).

Figure 9. Results of overall design improvements. (A) The initial version. (B) The final version of the game. GDP: gross domestic product; ICU: intensive care unit.

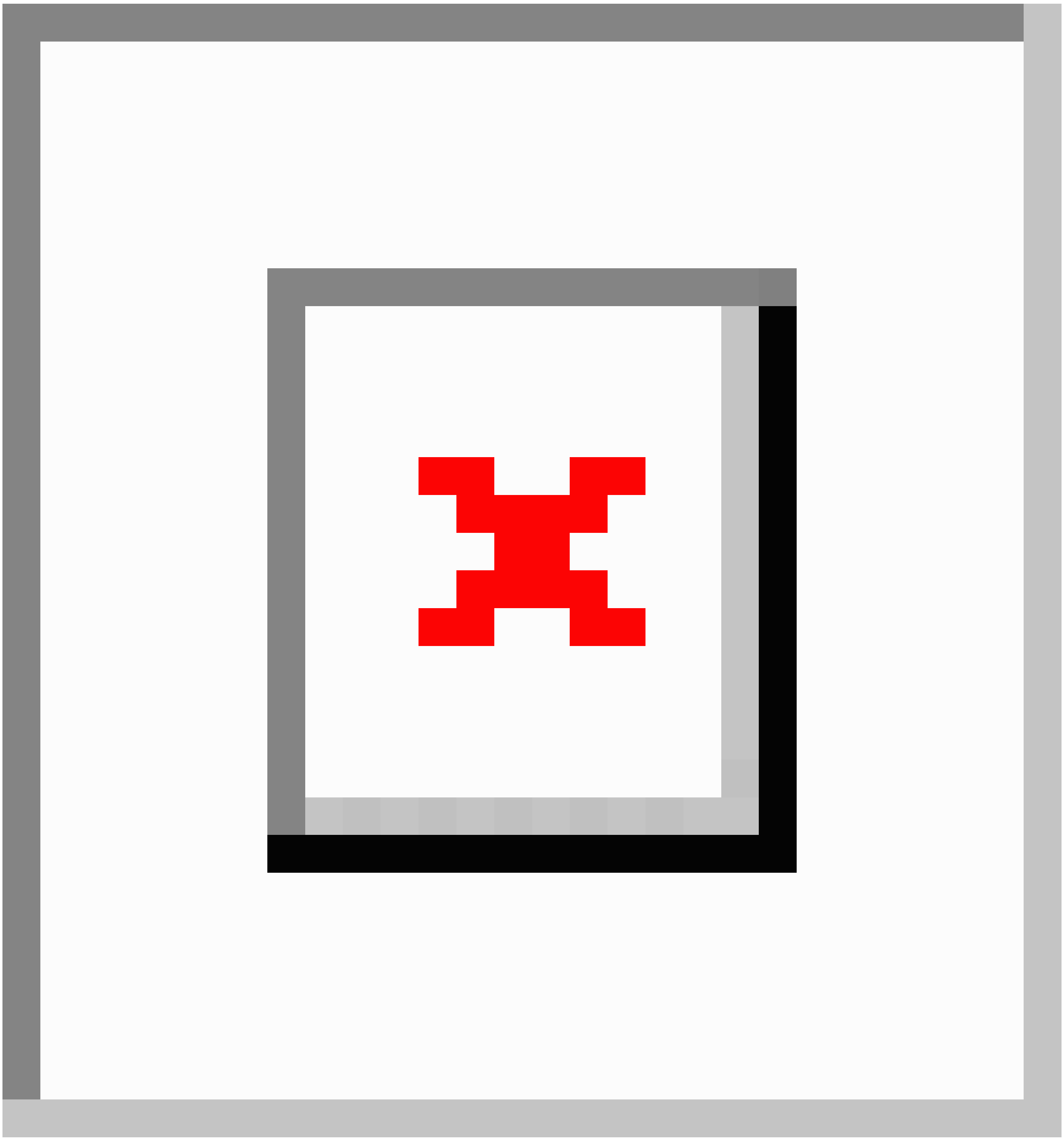
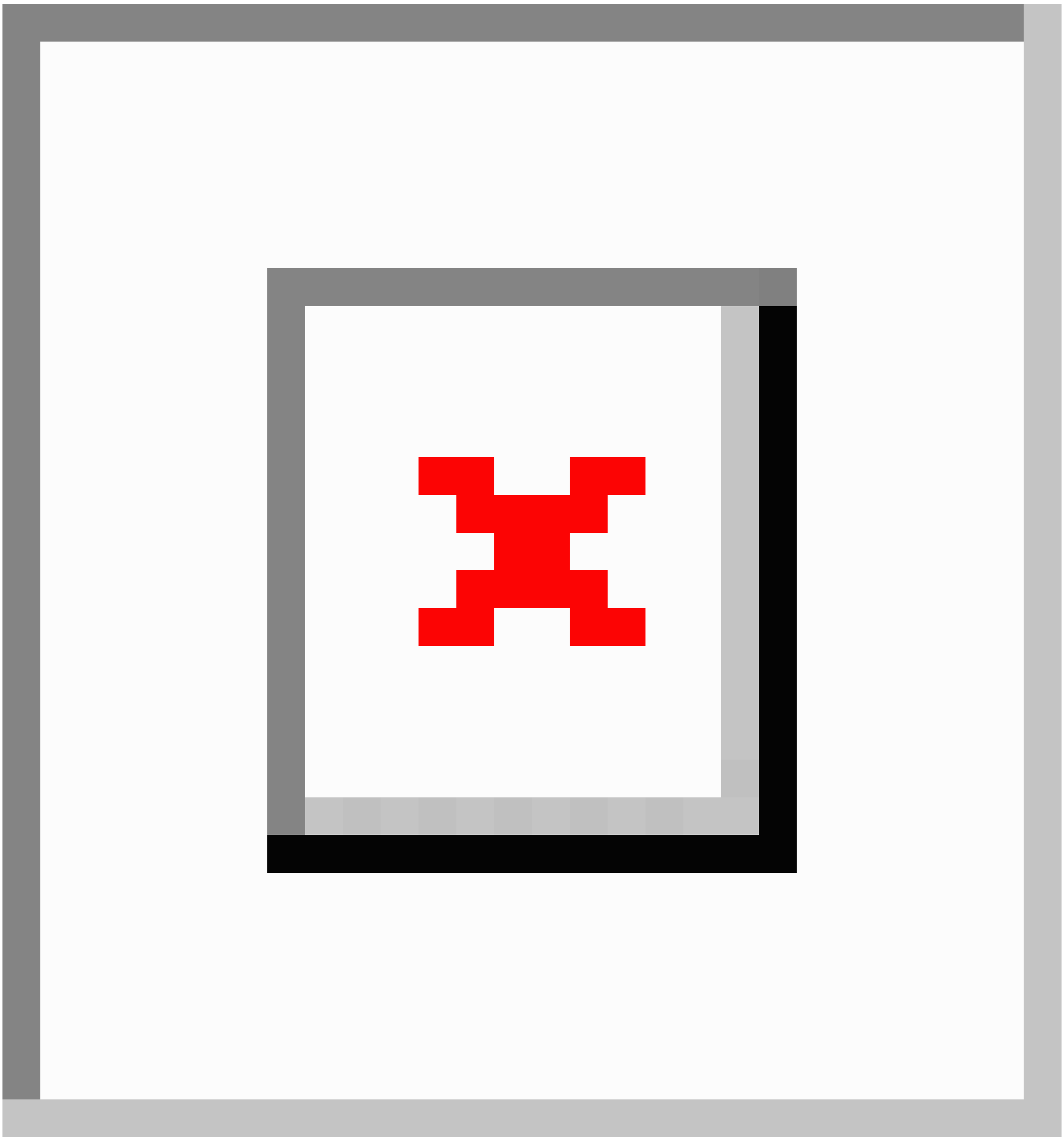


Figure 10. Overview of the Key Facts system. (A) When a player clicks a hint icon, (B) a message appears at the top, and (C) the Key Facts icon shows the count of newly discovered hints.



Game Evaluation

To assess if the game achieved its educational goals, we conducted surveys with the participants. The surveys measured the game's educational impact and its influence on participants' motivation to learn. In the earlier stage of PD, the first survey was conducted to measure participants' satisfaction with the initial version of the game. The second survey was conducted at a later stage and focused on examining participants' final evaluations of the game. Table 4 shows the descriptive statistics of the survey constructs.

Wilcoxon signed-rank tests were performed to examine the changes in participants' satisfaction between the 2 surveys. Although participants reported higher satisfaction in the second survey, the difference was not statistically significant (Z=−1.890, P=.06).

Participants also gave an overall positive evaluation of the game-based learning experience, with average scores of 6.19 and 6.06 for perceived learning and enjoyment, respectively. The perceived usefulness and ease of use, variables to examine the technical acceptance of FTC, were also evaluated positively, with an average of 5.75 and 5.63, respectively.

Table . Descriptive statistics of survey items (N=16).

Construct		Rating range	Values, mean (SD)
First survey			
	Satisfaction	1 - 4	2.94 (0.574)
Second survey			
	Satisfaction	1 - 4	3.25 (0.447)
	Perceived learning	1 - 7	6.19 (0.544)
	Enjoyment	1 - 7	6.06 (0.755)
	Usefulness	1 - 7	5.75 (0.848)
	Ease of use	1 - 7	5.63 (0.717)

Discussion

Principal Findings

Since the COVID-19 outbreak began in 2019, there has been a proliferation of educational materials on various preventive measures for infectious diseases. However, there remains a need for educational resources that facilitate a fundamental understanding of infectious diseases and quarantine policies. This study presents the development and design process of a quarantine policy education game ("Flattening the Curve"). The game serves as a comprehensive educational simulation, allowing users to learn about the characteristics of infectious diseases, the stages of transmission, the necessity and effectiveness of quarantine policies, and the crucial balance between these policies and economic considerations. Ultimately, the game aims to enhance learners' understanding of the concept of "flattening the curve."

The application of the SERES framework played a pivotal role in achieving a scientifically sound and educationally effective game [25]. Drawing insights from prior research, the scientific and design foundations of the SERES framework were found to be effective in developing health-related serious games [17,66]. Building on these foundations, this study effectively developed various learning components within the game, such as the infection model, social distancing model, and economic model.

This study shows that both the SERES framework and the PD process can be used simultaneously and interactively during the serious game development process. The effectiveness of education in an e-learning environment is highly dependent on the learner's motivation to learn [67], and an appropriate PD process during serious game development can lead to higher user satisfaction, which is essential for learner motivation [26,68]. The PD process was incorporated into the development of FTC to enhance the overall quality of the game and ensure learner engagement [69]. Participant feedback, categorized into the domains of learning, storytelling, gameplay, and user experience based on the DPE framework, guided the game's improvement and detailed development direction [33].

As the survey results showed, the final game received positive evaluations from participants. The PD activities made the serious game more enjoyable and satisfactory and contributed to the enhancement of the game's richness and user enjoyment.

Limitations and Further Research

Despite our efforts to increase the game's appeal to younger learners through design modifications guided by feedback from the PD process, this study has several limitations regarding its participants and evaluations. A comprehensive analysis of the game's learning effects was not conducted. Although validating the effectiveness of educational games is crucial, this study did not measure quantifiable learning outcomes, such as changes in knowledge or attitudes. Due to the nature of the PD process, participants engaged in continuous gameplay and provided feedback throughout a semester, posing challenges in evaluating the knowledge acquisition effects solely through gameplay. Consequently, the current evaluation primarily focused on qualitative assessments of learners' perceptions of the learning content.

Future research should target learners across various age groups, including children and adolescents. Measuring changes in knowledge related to infectious diseases and shifts in attitudes toward quarantine policies before and after gameplay can offer a more comprehensive evaluation of the learning effects of the FTC game. Additionally, we propose to evaluate the game with various game mechanics, such as difficulty and enjoyment. Investigating the relationship between game evaluations and learning effects will be instrumental in gaining a deeper understanding of the game's impact on learners.

Although not specifically addressed in this study, we have experimentally developed various multiplayer modes to respond to participant feedback and assessed their impact on learner engagement. Consistent with prior research [70], participants found the game more engaging in multiplayer mode than in single-player mode. Subsequent research could delve into how multiplayer mode can make serious games better, highlighting the effects of different gameplay modes on learning outcomes and immersion. This underscores the importance of sustained research and development efforts in the serious game domain, particularly those focused on enhancing learning experiences through innovative gameplay features.

Conclusions

"Flattening the Curve," a serious game designed to promote understanding of epidemic models and quarantine policies, was successfully developed by applying the SERES framework and the PD process.

The developed game is now publicly accessible online, allowing anyone to play [71]. The game is not only designed to educate about COVID-19–related quarantine policies but also serves broader educational purposes related to various types of infectious diseases. The game's source code is accessible to everyone, facilitating diverse educational applications and modifications to serve various research and educational objectives. The source code for the game is provided in a GitHub repository [72]. Resources are available under the license mentioned in the provided link. The authors welcome requests for additional information regarding the material presented in this paper.

Acknowledgments

This paper was financially supported by the Human Resource Development Project in Disaster Management and the Development of Life Safety Prevention Service Technology (2019-MOIS34-001 [RS-2019-ND618031]) of the Korea Ministry of the Interior and Safety.

Data Availability

As mentioned above, the source code of the game is provided in a GitHub repository [72]. The datasets generated during this study, including Korean-language data (transcripts and written assignments), are available from the corresponding author on reasonable request.

Authors' Contributions

JBC led the conceptualization, with MK and BJK supporting conceptualization. BJK led data curation, with support from MK. BJK did the formal analysis. JBC acquired funding and provided project administration, supervision, and resources. BJK led the investigation with support from MK and JBC. JBC led the development of the methodology with support from MK and BJK. MK was responsible for the software. BJK led validation with support from MK. MK was responsible for visualization, with support from BJK. BJK and MK wrote, reviewed, and edited the original draft with support from JBC.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey instrument details and reliability.
[DOCX File, 47 KB - games_v12i1e54968_app1.docx]

References

1. Williams SN, Armitage CJ, Tampe T, Dienes K. Public perceptions and experiences of social distancing and social isolation during the COVID-19 pandemic: a UK-based focus group study. BMJ Open 2020 Jul 20;10(7):e039334. [doi: 10.1136/bmjopen-2020-039334] [Medline: 32690752]
2. Baum NM, Jacobson PD, Goold SD. “Listen to the people”: public deliberation about social distancing measures in a pandemic. Am J Bioeth 2009 Nov;9(11):4-14. [doi: 10.1080/15265160903197531] [Medline: 19882444]
3. Taylor S, Asmundson GJG. Negative attitudes about facemasks during the COVID-19 pandemic: The dual importance of perceived ineffectiveness and psychological reactance. PLoS One 2021;16(2):e0246317. [doi: 10.1371/journal.pone.0246317] [Medline: 33596207]
4. Muhajarine N, Adeyinka DA, McCutcheon J, Green KL, Fahlman M, Kallio N. COVID-19 vaccine hesitancy and refusal and associated factors in an adult population in Saskatchewan, Canada: evidence from predictive modelling. PLoS One 2021;16(11):e0259513. [doi: 10.1371/journal.pone.0259513] [Medline: 34767603]
5. Bui LN, Berkowitz RL, Jilek W, et al. Public protests and the risk of novel coronavirus disease hospitalizations: a county-level analysis from California. Int J Environ Res Public Health 2021 Sep 8;18(18):9481. [doi: 10.3390/ijerph18189481] [Medline: 34574407]
6. Gisondi MA, Barber R, Faust JS, et al. A deadly infodemic: social media and the power of COVID-19 misinformation. J Med Internet Res 2022 Feb 1;24(2):e35552. [doi: 10.2196/35552] [Medline: 35007204]
7. Lee JJ, Kang KA, Wang MP, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. J Med Internet Res 2020 Nov 13;22(11):e22205. [doi: 10.2196/22205] [Medline: 33048825]
8. Freeman D, Waite F, Rosebrock L, et al. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. Psychol Med 2022 Jan;52(2):251-263. [doi: 10.1017/S0033291720001890] [Medline: 32436485]
9. Ahmed W, López Seguí F, Vidal-Alaball J, Katz MS. COVID-19 and the “Film Your Hospital” conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 Oct 5;22(10):e22374. [doi: 10.2196/22374] [Medline: 32936771]

10. Calik A, Cakmak B, Kapucu S, Inkaya B. The effectiveness of serious games designed for infection prevention and promotion of safe behaviors of senior nursing students during the COVID-19 pandemic. Am J Infect Control 2022 Dec;50(12):1360-1367. [doi: 10.1016/j.ajic.2022.02.025] [Medline: 35231565]
11. All A, Plovie B, Castellar EPN, Van Looy J. Pre-test influences on the effectiveness of digital-game based learning: a case study of a fire safety game. Comput Educ 2017 Nov;114:24-37. [doi: 10.1016/j.compedu.2017.05.018]
12. Zahed BT, White G, Quarles J. Play it safe: an educational cyber safety game for children in elementary school. Presented at: 2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games); Sep 4-6, 2019; Vienna, Austria. [doi: 10.1109/VS-Games.2019.8864594]
13. Chettoor Jayakrishnan G, Banahatti V, Lodha S. GOVID: repurposing serious game for enterprise COVID-19 awareness. Presented at: India HCI 2021; Nov 19-21, 2021; Virtual event India. [doi: 10.1145/3506469.3506471]
14. Hu H, Xiao Y, Li H. The effectiveness of a serious game versus online lectures for improving medical students' coronavirus disease 2019 knowledge. Games Health J 2021 Apr;10(2):139-144. [doi: 10.1089/g4h.2020.0140] [Medline: 33434099]
15. Kao D, Joshi A, Mousas C, et al. Fighting COVID-19 at Purdue University: design and evaluation of a game for teaching COVID-19 hygienic best practices. Presented at: FDG'21: The 16th International Conference on the Foundations of Digital Games 2021; Aug 3-6, 2021; Montreal, Canada. [doi: 10.1145/3472538.3472552]
16. Hill J, Corke E, Salawu M, Cotterell E, Russell M. Point of contact: investigating change in perception through a serious game for COVID-19 preventive measures. Proc ACM Hum Comput Interact 2021. [doi: 10.1145/3474701]
17. Suppan M, Catho G, Robalo Nunes T, et al. A serious game designed to promote safe behaviors among health care workers during the COVID-19 pandemic: development of "Escape COVID-19.". JMIR Serious Games 2020 Dec 3;8(4):e24986. [doi: 10.2196/24986] [Medline: 33242312]
18. Suppan M, Stuby L, Harbarth S, et al. Nationwide deployment of a serious game designed to improve COVID-19 infection prevention practices in Switzerland: prospective web-based study. JMIR Serious Games 2021 Nov 25;9(4):e33003. [doi: 10.2196/33003] [Medline: 34635472]
19. Hartikainen H, Iivari N, Kinnula M. Children's design recommendations for online safety education. Int J Child Comput Interact 2019 Dec;22:100146. [doi: 10.1016/j.ijcci.2019.100146]
20. Viitaharju P, Yliniemi K, Nieminen M, Karttunen AJ. Learning experiences from digital laboratory safety training. Educ Chem Eng 2021 Jan;34:87-93. [doi: 10.1016/j.ece.2020.11.009]
21. Bhute VJ, Inguva P, Shah U, Brechtelsbauer C. Transforming traditional teaching laboratories for effective remote delivery—a review. Educ Chem Eng 2021 Apr;35:96-104. [doi: 10.1016/j.ece.2021.01.008]
22. Shiradkar S, Rabelo L, Alasim F, Nagadi K. Virtual world as an interactive safety training platform. Inf 2021;12(6):219. [doi: 10.3390/info12060219]
23. Dzeng RJ, Hsueh HH, Chang RN. 3D game-based training system for hazard identification on construction site. Presented at: 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD); Aug 15-17, 2015; Zhangjiajie, China. [doi: 10.1109/FSKD.2015.7382339]
24. Wu WH, Hsiao HC, Wu PL, Lin CH, Huang SH. Investigating the learning - theory foundations of game - based learning: a meta - analysis. Comp Assisted Learn 2012 Jun;28(3):265-279. [doi: 10.1111/j.1365-2729.2011.00437.x]
25. Verschueren S, Buffel C, Vander Stichele G. Developing theory-driven, evidence-based serious games for health: framework based on research community insights. JMIR Serious Games 2019 May 2;7(2):e11565. [doi: 10.2196/11565] [Medline: 31045496]
26. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94. [doi: 10.2196/jmir.4444] [Medline: 27129447]
27. Ismail R, Ibrahim R, Yaacob S. Participatory design method to unfold educational game design issues: a systematic review of trends and outcome. Presented at: 2019 5th International Conference on Information Management (ICIM); Mar 24-27, 2019; Cambridge, United Kingdom. [doi: 10.1109/INFOMAN.2019.8714678]
28. Spinuzzi C. The methodology of participatory design. Tech Commun 2005;52(2):163-174 [FREE Full text]
29. Nunes EPS, Luz AR, Lemos EM, Nunes C. Approaches of Participatory Design in the Design Process of a Serious Game to Assist in the Learning of Hospitalized Children: Springer International Publishing; 2016:406-416.
30. Hansen EIK, Iversen OS. You are the real experts! - studying teenagers' motivation in participatory design. Presented at: IDC '13: Interaction Design and Children 2013; Jun 24-27, 2013; New York, USA. [doi: 10.1145/2485760.2485826]
31. Falcão TP, de Andrade e Peres FM, de Morais DCS, da Silva Oliveira G. Participatory methodologies to promote student engagement in the development of educational digital games. Comput Educ 2018 Jan;116:161-175. [doi: 10.1016/j.compedu.2017.09.006]
32. Tekinbas KS, Zimmerman E. Rules of Play: Game Design Fundamentals: MIT Press; 2003.
33. Winn BM. The design, play, and experience framework. In: Ferdig RE, editor. Handbook of Research on Effective Electronic Gaming in Education: Information Science Reference; 2009:1010-1024. [doi: 10.4018/978-1-59904-808-6.ch058]
34. Medeiros LC, Hillers VN, Chen G, Bergmann V, Kendall P, Schroeder M. Design and development of food safety knowledge and attitude scales for consumer food safety education. J Am Diet Assoc 2004 Nov;104(11):1671-1677. [doi: 10.1016/j.jada.2004.08.030] [Medline: 15499353]

35. Dragutinovic N, Twisk D. The effectiveness of road safety education: a literature review. : SWOV Institute for Road Safety Research; 2006 URL: https://swov.nl/en/publicatie/effectiveness-road-safety-education-literature-review [accessed 2024-09-24]
36. Schafer RB, Schafer E, Bultena GL, Hoiberg EO. Food safety: an application of the health belief model. J Nutr Educ 1993 Jan;25(1):17-24. [doi: 10.1016/S0022-3182(12)80183-X]
37. Covid-19 is now in 50 countries, and things will get worse. The Economist. 2020 Feb 29. URL: https://www.economist.com/briefing/2020/02/29/covid-19-is-now-in-50-countries-and-things-will-get-worse? [accessed 2024-09-22]
38. Gourinchas PO. Flattening the pandemic and recession curves. In: Baldwin R, Weder di B, editors. Mitigating the COVID Economic Crisis: Act Fast and Do Whatever: CEPR Press; 2020:31-40.
39. Dias S, Queiroz K, Araujo A. Introduction to group-structured-epidemic model. J Control Autom Electr Syst 2022 Feb;33(1):23-37. [doi: 10.1007/s40313-021-00841-7]
40. Ferrante M, Ferraris E, Rovira C. On a stochastic epidemic SEIHR model and its diffusion approximation. TEST (Madr) 2016 Sep;25(3):482-502. [doi: 10.1007/s11749-015-0465-z]
41. Yang P, Qi J, Zhang S, et al. Feasibility study of mitigation and suppression strategies for controlling COVID-19 outbreaks in London and Wuhan. PLoS One 2020;15(8):e0236857. [doi: 10.1371/journal.pone.0236857] [Medline: 32760081]
42. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc Lond A 1927 Aug;115(772):700-721. [doi: 10.1098/rspa.1927.0118]
43. Anderson RM, May RM. Directly transmitted infections diseases: control by vaccination. Science 1982 Feb 26;215(4536):1053-1060. [doi: 10.1126/science.7063839] [Medline: 7063839]
44. Ren SY, Wang WB, Gao RD, Zhou AM. Omicron variant (B.1.1.529) of SARS-CoV-2: mutation, infectivity, transmission, and vaccine resistance. World J Clin Cases 2022 Jan 7;10(1):1-11. [doi: 10.12998/wjcc.v10.i1.1] [Medline: 35071500]
45. Wolter N, Jassat W, Walaza S, et al. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. Lancet 2022 Jan 29;399(10323):437-446. [doi: 10.1016/S0140-6736(22)00017-4] [Medline: 35065011]
46. Chovatiya R, Silverberg JI. Inpatient morbidity and mortality of measles in the United States. PLoS One 2020;15(4):e0231329. [doi: 10.1371/journal.pone.0231329] [Medline: 32343688]
47. Francis MR. Just how contagious is COVID-19? This chart puts it in perspective. : Popular Science; 2020 URL: https://www.popsci.com/story/health/how-diseases-spread/ [accessed 2024-09-22]
48. Ebola virus disease. PAHO. URL: https://www.paho.org/en/topics/ebola-virus-disease [accessed 2022-03-30]
49. What you need to know about Ebola. CDC. URL: https://stacks.cdc.gov/view/cdc/25845 [accessed 2024-09-23]
50. Gastanaduy P, Haber P, Rota PA, Patel M. Measles. In: Epidemiology and Prevention of Vaccine-Preventable Diseases: CDC; 2021. URL: https://www.cdc.gov/pinkbook/hcp/table-of-contents/chapter-13-measles.html [accessed 2024-09-30]
51. Current status of COVID-19 [Website in Korean]. Korea Disease Control and Prevention Agency. 2024. URL: https://ncov.kdca.go.kr/pot/cv/trend/dmstc/selectMntrgSttus.do [accessed 2024-09-23]
52. Qualls N, Levitt A, Kanade N, et al. Community mitigation guidelines to prevent pandemic influenza - United States, 2017. MMWR Recomm Rep 2017 Apr 21;66(1):1-34. [doi: 10.15585/mmwr.rr6601a1] [Medline: 28426646]
53. Miksch F, Urach C, Einzinger P, Zauner G. A flexible agent-based framework for infectious disease modeling. In: Information and Communication Technology: Springer Berlin Heidelberg; 2014:36-45. [doi: 10.1007/978-3-642-55032-4_4]
54. Maharaj S, Kleczkowski A. Controlling epidemic spread by social distancing: do it well or not at all. BMC Public Health 2012 Aug 20;12:679. [doi: 10.1186/1471-2458-12-679] [Medline: 22905965]
55. Agarwal A, Alomar A, Sarker A, Shah D, Shen D, Yang C. Two burning questions on COVID-19: did shutting down the economy help? can we (partially) reopen the economy without risking the second wave? arXiv. Preprint posted online on May 10, 2020. [doi: 10.48550/arXiv.2005.00072]
56. Kristjanpoller W, Michell K, Minutolo MC. A causal framework to determine the effectiveness of dynamic quarantine policy to mitigate COVID-19. Appl Soft Comput 2021 Jun;104:107241. [doi: 10.1016/j.asoc.2021.107241] [Medline: 33679272]
57. Bauer L, Broady K, Edelberg W, O'Donnell J. Ten facts about COVID-19 and the US economy. : Brookings Institute; 2020 URL: https://www.brookings.edu/wp-content/uploads/2020/09/FutureShutdowns_Facts_LO_Final.pdf [accessed 2024-09-30]
58. Arnab S, Lim T, Carvalho MB, et al. Mapping learning and game mechanics for serious games analysis. Brit J Educ Tech 2015 Mar;46(2):391-411. [doi: 10.1111/bjet.12113]
59. Lameras P, Arnab S, Dunwell I, Stewart C, Clarke S, Petridis P. Essential features of serious games design in higher education: linking learning attributes to game mechanics. Brit J Educ Tech 2017 Jun;48(4):972-994. [doi: 10.1111/bjet.12467]
60. Silva FGM. Practical methodology for the design of educational serious games. Inf 2019;11(1):14. [doi: 10.3390/info11010014]
61. Almala AH. A constructivist conceptual framework for a quality e-learning environment. Dist Learn 2005;2(5):9-12 [FREE Full text]
62. Huang CY, Tsai YS, Wen TH. Simulations for epidemiology and public health education. In: Operational Research for Emergency Planning in Healthcare 2016, Vol. 2:176-202. [doi: 10.1007/978-1-137-57328-5_9]

63. Hale T, Angrist N, Goldszmidt R, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021 Apr;5(4):529-538. [doi: 10.1038/s41562-021-01079-8] [Medline: 33686204]
64. Barclay E, Scott D, Animashaun C. The US doesn't just need to flatten the curve. It needs to "raise the line.". Vox. 2020. URL: https://www.vox.com/2020/4/7/21201260/coronavirus-usa-chart-mask-shortage-ventilators-flatten-the-curve [accessed 2024-09-22]
65. Laine TH, Lindberg RSN. Designing engaging games for education: a systematic literature review on game motivators and design principles. IEEE Trans Learn Technol 2020;13(4):804-821. [doi: 10.1109/TLT.2020.3018503]
66. Sarasmita MA, Larasanty LPF, Kuo LN, Cheng KJ, Chen HY. A computer-based interactive narrative and a serious game for children with asthma: development and content validity analysis. J Med Internet Res 2021 Sep 13;23(9):e28796. [doi: 10.2196/28796] [Medline: 34515641]
67. Kim BJ, Chung JB. Is safety education in the e-learning environment effective? Factors affecting the learning outcomes of online laboratory safety education. Saf Sci 2023 Dec;168:106306. [doi: 10.1016/j.ssci.2023.106306]
68. Pollio A, Magee L, Salazar JF. The making of Antarctic futures: participatory game design at the interface between science and policy. Futures 2021 Jan;125:102662. [doi: 10.1016/j.futures.2020.102662]
69. Khaled R, Vasalou A. Bridging serious games and participatory design. Int J Child Comput Interact 2014 May;2(2):93-100. [doi: 10.1016/j.ijcci.2014.03.001]
70. Hansen A, Larsen KB, Nielsen HH, Sokolov MK, Kraus M. Asymmetrical multiplayer versus single player: effects on game experience in a virtual reality edutainment game. In: Augmented Reality, Virtual Reality, and Computer Graphics: Springer International Publishing; 2020:22-33. [doi: 10.1007/978-3-030-58465-8_2]
71. Kwak MH, Kim BJ, Chung JB. Flattening the Curve. GitHub. 2022. URL: https://brad9704.github.io/COVID_Simulation/en-us/covid_game.html [accessed 2024-09-22]
72. Kwak MH, Kim BJ, Chung JB. Flattening the Curve (source file). GitHub. 2022. URL: https://github.com/brad9704/COVID_Simulation/releases/tag/v1.0.0 [accessed 2024-09-22]

Abbreviations

C: Critical
D: Dead
DPE: Design, Play, and Experience
E1: Exposed (not contagious)
E2: Exposed (contagious)
FTC: Flattening the Curve
GDP: gross domestic product
H: Hospitalized
I1: Infected, mild
I2: Infected, severe
ICU: intensive care unit
M: Mild
PD: participatory design
R: Recovered
S: Susceptible
SEIHR: Susceptible-Exposed-Infectious-Hospitalized-Recovered
SEIR: Susceptible-Exposed-Infectious-Recovered
SEMCR: Susceptible-Exposed-Mild-Critical-Recovered
SIR: Susceptible-Infectious-Recovered
UI: user interface

Edited by A Mavragani; submitted 29.11.23; peer-reviewed by S Jonathan, V Wattanasoontorn; revised version received 29.07.24; accepted 06.08.24; published 15.10.24.

Please cite as:
Kwak M, Kim BJ, Chung JB
Serious Game Development for Public Health: Participatory Design Approach to COVID-19 Quarantine Policy Education
JMIR Serious Games 2024;12:e54968
URL: https://games.jmir.org/2024/1/e54968
doi:10.2196/54968

Using Games to Simulate Medication Adherence and Nonadherence: Laboratory Experiment in Gamified Behavioral Simulation

Umar Taj*, PhD; Aikaterini Grimani*, PhD; Daniel Read*, PhD; Ivo Vlaev*, PhD

Behavioural Science Group, Warwick Business School, University of Warwick, Scarman Road, Coventry, United Kingdom

*all authors contributed equally

Corresponding Author:
Aikaterini Grimani, PhD

Abstract

Background: Medical nonadherence is a significant problem associated with worse clinical outcomes, higher downstream rehospitalization rates, and a higher use of resources. To improve medication adherence, it is vital for researchers and practitioners to have a solid theoretical understanding of what interventions are likely to work. To achieve this understanding, we propose that researchers should focus on creating small-scale laboratory analogs to the larger real-world setting and determine what interventions, such as nudges or incentives, work to change behavior in the laboratory. To do this, we took inspiration from the literature on serious games and gamification and experimental economics. We call our approach "gamified behavioral simulation." In this paper, we modeled everyday life as the state of being engaged in a simple but addictive game, illness as being interruptions to the functionality of that game, treatment as being a series of actions that can be taken to prevent or mitigate those interruptions, and adherence as sticking to a prescribed rule for the application of those actions.

Objective: This study carries out a behavioral diagnosis of the medication adherence problem through a theoretically informed framework and then develops the gamified behavioral modeling approach to simulate medication nonadherence.

Methods: A laboratory experiment was conducted using a modified popular and addictive open-source video game called "2048," which created an abstract model for the medication adherence behavior observed in real life. In total, 509 participants were assigned to the control and 4 intervention groups ("incentive" group, "reminder" group, "commitment device" group, and "elongated duration for symptoms" group).

Results: The results of the modeling experiment showed that having theoretically informed interventions can increase the likelihood for them to be successful. In particular, there is evidence that the use of reminders improves the medication adherence rates for patients, and the same result was found in the modeling experiment, as they improved adherence significantly by 23% (95% CI −33.97% to −11.72%; *P*<.001). However, providing an incentive did not improve the adherence rate. We also tested the use of commitment devices, which, in line with real-world evidence, did not improve adherence rates. The fourth treatment tested elongated duration for symptoms, which attempted to show the power of modeling experiments where we test a what-if scenario that is extremely difficult to test in a real setting. The results indicated that if symptoms last longer, people did not adhere more to their medication regimen.

Conclusions: Gamified behavioral simulation is a useful tool to explain real health behaviors and help in identifying which interventions are most likely to work in a randomized trial.

(JMIR Serious Games 2024;12:e47141) doi:10.2196/47141

KEYWORDS

behavior change; experimental modeling; gamification; medication adherence; antibiotics; games; medication; testing behavior; clinical outcome; simulate; diagnosis; devices; symptoms; tool

Introduction

Overview

Medication adherence is the extent to which the patient's behavior matches the agreed recommendations from the prescriber. There is consistent evidence that regardless of what is being treated, nonadherence is associated with worse clinical outcomes and higher downstream rehospitalization rates [1]. Yet, several reviews have found that medication adherence among patients is low, especially in low-income countries [1-3]. In the United States, for example, 51% of patients with hypertension adhere to their medication, but the corresponding numbers in Gambia, Seychelles, and China are 27%, 26%, and 43%, respectively [4-8]. Even in the developed world, medication nonadherence is expensive. The annual cost to the

UK government has been estimated at £150 million (approximately US $270 million in 2005); for the United States, this figure is over US $1 billion [3].

The cause of medication nonadherence varies among patients and can be categorized as intentional or unintentional [9]. Intentional nonadherence involves patients who decide consciously not to take medication as instructed, based on perceptions, feelings, or beliefs [10]. Unintentional nonadherence occurs when the patient wants to take their medication as instructed but fails to do so. The most common factors associated with nonadherence include forgetfulness (50%), having other medications to take (20%), and being symptom-free (20%) [11]. To the extent that nonadherence is unintentional, patients are likely to welcome any nudge that helps them adhere to their doctor's recommendations.

New and innovative strategies are essential to successfully improve patient adherence to treatment. One potentially effective strategy for understanding patient medication adherence is the use of serious games and gamification [12] or the introduction of gaming elements into the medication experience. Games motivate users into engaging in an activity with a higher intensity and duration [13]. Game elements are activities, behaviors, and mechanisms designers incorporate into a specific context to create a gameful experience [13]. Introducing gaming elements into a nongaming context has the potential to transform routine tasks into more enjoyable and motivating experiences [14]. The aim of this study is to conduct a behavioral diagnosis of medication adherence issues using a theoretically informed framework and then to design, implement, and assess a game-based setting for simulating nonadherence behavior.

A range of studies have explored the use of games and technology (mainly via mobile health apps) to understand and improve treatment adherence. Tran et al [15] found that these features can lead to improved or sustained medication adherence but noted significant heterogeneity in patient populations and methodologies. de Vette et al [16] highlighted the potential of games to engage older people in telemedicine, while Brown et al [17] called for further research to understand the impact of game features on program adherence in web-based interventions for mental health. These studies collectively underscore the potential of games and incentives in improving medication adherence.

This paper takes inspiration from the literature on serious games and gamification and experimental economics. We call our approach "gamified behavioral simulation." We model everyday life as the state of being engaged in a simple but addictive game, illness as being interruptions to the functionality of that game, treatment as being a series of actions that can be taken to prevent or mitigate those interruptions, and adherence as sticking to a prescribed rule for the application of those actions. As far as we are aware, no previous research has addressed whether serious games can be used to model (simulate) medication adherence behavior and test the effectiveness of interventions to improve such adherence. This is a crucial gap in the literature that our study addresses.

Background

Existing literature indicates that interventions to improve medication adherence have had mixed results [3,8]. The Medical Research Council proposes that the development of a behavior change intervention should follow the same cycle as drug development: (1) a theory behind the design of behavioral intervention; (2) followed by modeling of the problem or behavior; and (3) finally, a randomized controlled trial (RCT) and implementation of the intervention [18].

Interventions developed to bring about behavior change often offer limited practical value, as they lack a theoretical basis for their selection and development [19]. The Medical Research Council proposes that an important early task for a researcher is to develop a theoretical understanding of the underlying process and constructs that might bring about behavior change. This helps in clearly understanding how successful interventions have had their effect, that is, which behavior change processes can be attributed to the observed change. A theoretical underpinning further allows the researcher to argue for the selection of a particular intervention [18].

The Medical Research Council proposed that this "theory" stage should be followed by a "modeling" stage, which can be considered as the equivalent to "testing with mice" in the drug development cycle. Modeling allows the researcher to investigate and identify the exact mechanisms that are bringing about the behavior change. It makes possible the study of isolated effects of different interventions. Modeling thus allows the researcher to make complex phenomena manageable and create knowledge about the underlying mechanisms of behavior change that might be quite difficult to uncover otherwise [18]. It is only once a clear understanding of what works has been achieved through the modeling stage that an RCT should be carried out to empirically test the intervention.

In this study, we focus on the "modeling" stage for developing behavior change interventions to increase patient's adherence to antibiotic medication. It is common in behavioral science and economics to model real-life behaviors in a laboratory setting in an attempt to control and make tractable the phenomenon that the researcher is interested in investigating. However, introducing this modeling approach in the development of complex interventions is not commonly found in the health behavior change literature [19]. We argue that using gamified behavioral simulation to first model treatment behavior can provide important information about the choice and design of the behavior change intervention to be tested in an RCT in a short time with a low cost.

Our experimental model for the nonadherent behavior in patients taking antibiotics was inspired by the conceptual approach developed by Kessler and Roth [20]. They used an abstract experimental interaction to model the effectiveness of the priority rule in increasing the registration of organ donors. An organ allocation policy that prioritizes registered donors on waiting lists was found to significantly boost donor registration. Their setting did not involve actual organ donation decisions and neither did they use any organ donation terminology during the experiment. However, they imposed real (monetary) costs to correspond to the analogous costs associated with decisions to donate or receive an organ. Their results showed that the priority rule condition had a significant positive impact on their laboratory-based "organ donation–like" decisions, and they used this to make a strong case in their paper to introduce the priority rule in the organ donation policy present in the United States. Subsequent work in the field indeed suggested that this laboratory method, designed without making any direct reference to organ donation, predicted organ donation attitudes and preferences [21] and even became the basis for policy decisions in at least 1 country [22].

We adopted a similar approach. The aim of this study was to develop an abstract experimental task that can capture the key elements of a real-world setting and thus simulate the nonadherence behavior of the patients taking antibiotic medication. At the heart of the task was a "game," which enabled us to make the experiment engaging and to incentivize behavior. This simulation served as a platform to test multiple interventions to positively increase the target behavior.

Our research question addresses a significant challenge in gamification technology: whether gamified behavioral modeling is a useful tool to explain real health behaviors and help in identifying which interventions are most likely to work in the real-world [23,24]. This challenge stems from the need to validate that behaviors observed within a gamified environment accurately reflect those in real-world settings. Additionally, it ensures that the motivational elements used in serious games are effective in driving the desired health behaviors and can be predictive of the outcomes of interventions in randomized trials. Addressing this challenge is crucial for the advancement of gamified applications in health-related fields, where the ultimate goal is to positively influence health outcomes and behaviors.

To address this question, we examined a range of interventions—those that have yielded both successful and unsuccessful outcomes in real-world trials—and assessed whether our gamified behavioral simulation can mirror these varied patterns of results. Therefore, our hypothesis was that our gamified behavioral simulation can accurately replicate these results, which implies that it would be a useful tool for predicting the success of interventions in the field.

Note that this approach differs from traditional methods to develop and test medication adherence interventions, which usually involve understanding barriers to action (adherence) and designing field trials to test interventions that address those behavior change challenges [25-27]. These interventions include patient education, medication regimen management, clinical pharmacist consultation, cognitive behavioral therapies, incentives, and various technology-based intervention and measurement components such as medication-taking reminders, support messages, and adherence measurement methods (eg, electronic drug monitors [pill bottles], sensor systems, and proximity sensing) [26]. Such field trials are costly and often cannot test every possible intervention strategy. Thus, evaluating our gaming methodology is essential, as it allows us to explore certain hypotheses about medication adherence within a controlled laboratory environment, where actual medication decisions are not involved. While not all aspects of medication adherence can be abstracted for study, there are critical elements of the adherence decision process that cannot be consistently manipulated in real-world settings, yet are amenable to manipulation in a laboratory setting (such as, different incentive schemes or elongated duration for symptoms).

Methods

Experiment Design

The laboratory experiments were conducted in December 2015 by Gallup Pakistan at 2 laboratories in Islamabad and Karachi. The laboratory experiments involved using a (then) popular and addictive open-source video game called "2048." The game's aim is to move numbered tiles in such a way that the total adds up to 2048 (Figure S1 in Multimedia Appendix 1).

The original 2048 game was modified to create an abstract model for the medication adherence behavior observed in real life. The game is the analog of "everyday life."

Participant Recruitment

The sample (N=509; n=305, 60% male and n=204, 40% female) was recruited through Gallup Pakistan. They were screened on the basis of whether they were able to browse on the internet, as the laboratory experiments requested to play a computer game (each of the control and treatment groups had the same representative proportions). Gallup Pakistan provided transport facilities to any participants who requested it. This was especially the case with female participants, as it was quite difficult to recruit them otherwise. If necessary, the timing of running the experiment was adjusted to accommodate participants after office hours. All participants provided written informed consent.

Procedure

The participants were asked to open the original 2048 game tab and practice for 10 minutes. At the end of the practice round, each participant was asked to close the original 2048 game tab and was directed to open an instruction video tab. The instructional video in Urdu (the national language of Pakistan) was made following a number of pilot tests (Multimedia Appendix 1) and extensive feedback from participants. It was recorded using CamStudio (Rendersoft Development). The instructional video explained in detail how to begin playing the experimental game and gave a demo of the experimental game. Three videos were made. The video for the control group was also used for the "commitment device" treatment condition and the "elongated duration for symptoms" treatment condition. The content of the "incentive" and "reminder" videos were the same except for some condition-specific information. Once the participants were finished watching the instructional video, they were asked to begin the experimental game.

The total duration of the game part of the experiment was 14 minutes and 30 seconds, and the participants were instructed to enter the code every minute or 14 times in total. The 14 code entries simulated a typical 7-day antibiotic medication course where patients are prescribed to take the pill twice a day with roughly 12-hour intervals. However, the inclusion of the initial practice rounds, the instruction video, and the payment process brought the whole task to 30 minutes on average.

Control Condition

When participants started the game, after the instructional video, the screen was blurred, making it very difficult for them to play the game. This blurriness simulated the onset of illness. To simulate the use of medication, participants were provided with a code, which they were required to enter every minute to clear the screen. The screen became clearer each time the code was entered. However, halfway through the game, the screen remained clear even though the participants were still expected to enter the code. If the code was not entered, however, there was a chance of "relapse," and the screen became blurry again.

As an analog to the pill pack, a pill counter was displayed on the left of the screen, showing how many times the code had been entered and how many times remained. Participants received a total of 14 opportunities to enter the code correctly, 1 every minute. If participants failed to adhere to this regimen, then, after the 14 opportunities, their codes become ineffective. This corresponded to the cost of forgetting to take one's medication. On the top left corner of the screen, a timer was displayed with the time elapsed since the start of the game. The timer was there to assist the participants in keeping track of time.

The show-up fee was Rs 80 (£0.50, approximately US $0.77 in 2015), and it was already included in the earnings; participants started the game with a score of 0 and earnings of Rs 80 (£0.50, approximately US $0.77 in 2015). Participants were rewarded on how well they scored in the 2048 game. The maximum money that participants could earn was Rs 500 (approximately £3 or US $4.59 in 2015). The final score and earning of the participant showed up on the screen along with their participant number, which served as an ID.

Participants were allowed to play the game as many times as possible during the duration of the experiment (which was 14 minutes 30 seconds). Once a player had no moves left on the board, a message box popped up on the screen giving the participant an option to restart the game (Multimedia Appendix 1).

To examine whether gamified behavioral modeling is a useful tool to explain real health behaviors and predict which interventions are most likely to work, we examined a range of interventions that have yielded both successful and unsuccessful outcomes in real-world trials, aiming to improve antibiotic medication adherence.

Treatment Condition 1: Incentive

For the "incentive" treatment condition, the design and mechanics of the experiment were exactly the same as the control experiment except that participants were given an incentive of Rs 5 (£0.02, approximately US $0.03 in 2015) every time they entered the correct code on time. This was an extra bonus on top of their usual earnings in the game, and the increment would show up in the earnings box on the top left corner of the screen. In addition, when participants entered the code correctly and on time, a message would flash on the screen informing them that they earned a bonus of Rs 5 (£0.02, approximately US $0.03 in 2015).

The instructional video before the "incentive" version of the game explicitly mentioned the Rs 5 (£0.02, approximately US $0.03 in 2015) bonus that they would earn upon entering the code correctly on time. The video also showed one such instance where the code was entered correctly on time and the earnings increased by Rs 5 (£0.02, approximately US $0.03 in 2015).

Treatment Condition 2: Reminder

For the "reminder" treatment condition, the design and mechanics of the experiment were exactly the same as the control experiment except that a message box would pop up on the screen when it was time to enter the code. The message box informed the participant that it was time to enter the code. This message box would appear in the middle of the board and stay for a few seconds before disappearing.

The instructional video for the "reminder" version of the game explicitly mentioned that a reminder message would show up when it would be time to enter the code. The video also showed one such instance where it was time to enter the code and the message box popped up on the screen.

Treatment Condition 3: Commitment Device

For the "commitment device" treatment condition, the design and mechanics of the experiment were identical to the control experiment except that at the start of the experiment, participants were asked to sign a sticker stating that they committed to entering the code as prescribed. The sticker was then pasted on the laptop they were using to play the experimental game. Participants watched the same instruction video as in the "control" condition.

Treatment Condition 4: Elongated Duration for Symptoms

An additional treatment condition was introduced in which symptoms took twice as long to disappear compared to the control condition. For the "elongated duration for symptoms" treatment condition, the rest of the design and mechanics of the experiment were exactly the same as the control experiment. Participants watched the same instruction video as in the "control" condition.

Ethical Considerations

The study was approved by the University of Warwick's ethics committee (ethical application reference 100/15‐16). All procedures were conducted in accordance with the relevant guidelines and regulations. Strict ethical and legal standards were upheld, ensuring that all personal information was securely stored, treated confidentially, and anonymized. Informed consent was obtained from all participants, permitting the use and publication of their data in this research. Participants received a show-up fee of Rs 80 (approximately US $0.77 in 2015) and had the opportunity to earn up to Rs 500 (approximately £3 or US $4.59 in 2015) based on their performance in the game. Participation in the study was entirely voluntary, and participants retained the right to withdraw at any time without any obligation for further contact from the study staff after withdrawal.

Results

We report results from 509 participants who participated in our laboratory experiment (Figure 1). There were 104 participants in the control group, 106 in the “incentive” group, 97 in the “reminder” group, 102 in the “commitment device” group, and 100 in the “elongated duration for symptoms” group.

Planned comparisons (ANOVA) were carried out to determine significant changes in adherence rates between the control group and treatment groups using the Bonferroni corrections post hoc test. The results showed that the adherence rate differed among the conditions ($F_{4,504}$=10.63; P<.001; n^2_p=0.078). The adherence rate in the control group was 44% (mean 44.16%, SD 27.45%), and similar adherence rates were seen in the “incentive” group (mean 52.02%, SD 29.41%), the “commitment device” group (mean 45.17%, SD 26.50%), and the “elongated duration for symptoms” group (mean 53.57%, SD 29.05%). In particular, the commitment device did not bring any change in the adherence rate compared to the control group (1%, 95% CI −11.99% to 9.98%). In the “incentive” as well as in the “elongated duration for symptoms” treatment group, the adherence rate improved by 8% (95% CI −18.74% to 3.02%) and 10% (95% CI −20.45% to 1.63%), respectively; however, the results were insignificant (*P*=.42 and *P*=.17, respectively). On the contrary, reminders had a different adherence rate (mean 67.01%, SD 27.17%) compared to the control group, as they improved adherence significantly by 23% (95% CI −33.97% to −11.72%; *P*<.001). The finding here suggests that simply reminding people to take their antibiotic medication can improve medication adherence significantly (Table 1).

Regarding the code entries, 63% (n=4489) were correct and on time across all conditions, while 21% (n=1497), although correct, were entered at the wrong time. This seems like a fairly reasonable result, as it was expected that some patients might take their pill at the wrong time. Only 7% (n=499) of the code entries consisted of wrong codes, which translated to a few patients taking the wrong pill. This result can be justified as patients who are on complex medication regimens (such as, for tuberculosis) do sometimes take the wrong pill.

The experiment included a consequence for nonadherence in the form of the screen relapsing to being blurry. To recall, once the participants’ screen was cleared, if they failed to enter the code, there was a 2% chance of “relapse” or the screen becoming blurry by 25%. The probability of “relapse” doubled each time the code was not entered. There were 73 participants in total who experienced relapse (each of these participants only experienced relapse once).

Our expectation was that once a participant had experienced relapse, they would be more adherent, but it seems that except for the incentive condition, adherence rates dropped after relapse. However, there may be an explanation for this result. Table 2 shows the round when relapse occurred for different participants. It can be seen that for more than half of the participants, relapse occurred when they only had 1, 2, or 3 code entries left. Since they knew that the game would end soon, they might have decided to forego entering the code and instead focus on the 2048 game. Since most participants experienced relapse very close to the end of the experiment, drawing meaningful insight into the relationship between relapse and adherence rate might not be appropriate.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram.

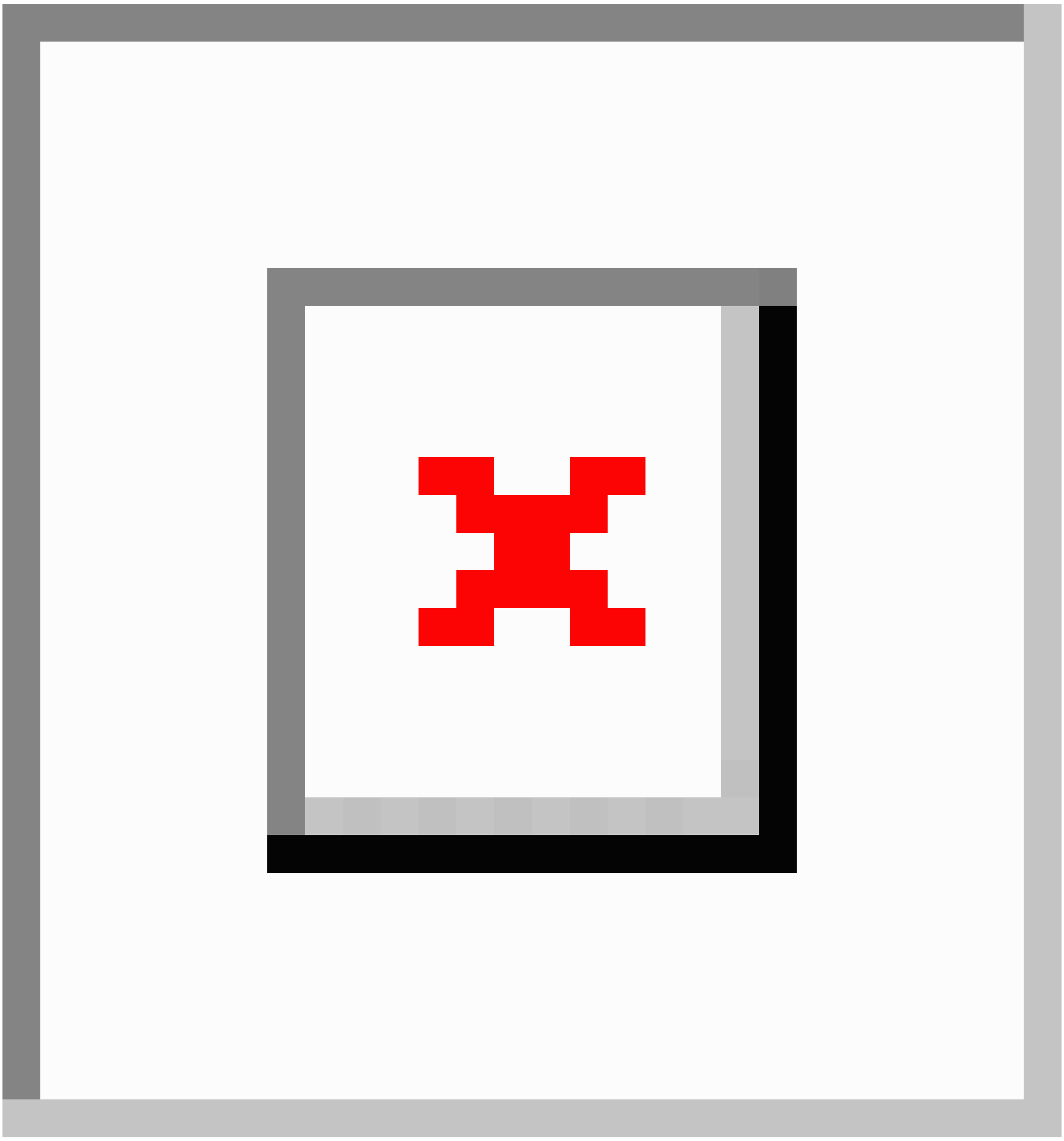


Table . Adherence rate of the control group and the intervention groups.

Group	Sample size, n (%)	Adherence rate (%)
Control group	104 (20)	44
Incentive group	106 (21)	52
Reminder group	97 (19)	67
Commitment device group	102 (20)	45
Elongated duration for symptoms group	100 (20)	54

Table . Adherence rates across the 5 conditions in the experiment.

Round when relapsed	Adherence rates (%)
7	4.1
8	5.5
9	4.1
10	19.2
11	8.2
12	16.4
13	21.9
14	20.5

Discussion

Overview

We intended to carry out a behavioral diagnosis of the medication adherence problem through a theoretically informed framework and then develop a model to simulate the nonadherence behavior. This constitutes the second stage of the theory-modeling-RCT cycle for developing behavior change interventions. The results of the modeling experiment showed that having theoretically informed interventions increased the likelihood for them to be successful. Furthermore, modeling can also help in identifying which interventions are most likely to work in the RCT.

It is worth discussing how good the 2048 experiment has been in modeling the nonadherence behavior. The modeling experiment seemed to have generated the same patterns that are expressed by patients such as forgetting to take the pill or adhering less to the regimen once the symptoms disappear. The results of the modeling experiment reflect this behavior of patients quite closely among the participants as well. The adherence rate found among the control group of the game experiment was 44%, which is quite similar to the adherence rate found among patients. While this statistic alone does not fully confirm the effectiveness of the gamification model, it is important to note that if the control group's adherence rate had been either very low or very high, it would have indicated that the 2048 experiment was not accurately modeling nonadherence behavior. Among the treatment conditions that were tested, it seems that reminders are only effective in increasing adherence rates.

Comparison to Prior Work

The results of the modeling experiment showed that providing incentives did not improve the adherence rate. Providing incentives was chosen as one of the treatment conditions because there is a lack of reinforcement for patients to continue adhering to their antibiotic medication once the symptoms disappeared [28]. However, Klein [29] maintains that rewarding behavior with money may have the undesirable effect of devaluing the intrinsic benefits of adherence, creating an even higher barrier to long-term adherence. Ideally, incentives should provide frequent, small (but tangible) rewards [30-32].

The results of the modeling experiment showed that providing reminders improves the adherence rate. There is evidence that the use of reminders improves the medication adherence rates for patients [33,34] because forgetting is the most prominent reason for nonadherence [28]. Using SMS text messaging for medication reminders appears to have a significant positive effect on medication adherence in other clinical areas such as, for example, mental health [35,36], with enduring effects [34]. The effectiveness of interventions using electronic reminders to improve adherence to antibiotics has not been tested in the field for real, so our experiment provides a promising indication that such an intervention is very likely to be effective.

While the use of reminders increased the adherence rate in the modeling experiments, we wanted to turn the tables and model an intervention that was yet to be carried out in a real-world setting. This motivated our selection of the use of commitment devices as one of the treatment groups in the modeling experiment. Commitment devices have been used in various contexts, such as smoking cessation, weight loss, exercise, and savings. Some studies have also explored the use of commitment devices for medication adherence among patients with chronic conditions, such as HIV, diabetes, or hypertension [28]. The results have been mixed, but some evidence suggests that commitment devices can improve medication adherence and health outcomes, especially when combined with other strategies, such as reminders, education, and incentives. Therefore, we aimed to test this strategy in our setting. We used the same protocol that was followed by the Department of Health and Boots UK in their RCT, and the results showed that the commitment sticker did not improve the adherence rates of the participants in the modeling experiment. Interestingly, a few months later, when the results of the RCT came out, it was found that the commitment sticker did not bring about any improvement in the adherence rates among patients as well. From the results obtained in the modeling experiment and their comparison with real-world analogs, it seems that the modeling experiment was able to simulate the nonadherence behavior of the patients.

The last treatment group of "elongated duration for symptoms" really shows the power of modeling experiments where we test a what-if scenario, which is extremely difficult to test in an RCT setting. The results from the experiment provide a proof of concept that if symptoms last longer people adhere more to their medication regimen. An interesting idea that comes to mind is that maybe the pills can be made in such a way that they keep the patients feeling sick until the last day of their treatment. Interestingly, a behavioral consultancy in India is testing a similar intervention. They are working on the issue of nonadherence to tuberculosis medication (which is an antibiotic medication but for a 6-month period) and have come across the same issue that patients stop taking their pills once the symptoms disappear. Rather than making people keep feeling sick (which is an extreme interpretation of the modeling experiment finding), they introduced an intervention to make patients realize that they are still sick even though the symptoms are gone.

In this study, we did not aim to identify which intervention or combination of interventions can be most effective in increasing medication adherence but rather to develop a model that can provide a platform to test various interventions and select the most effective ones to be included in an RCT. It is quite possible that an intervention is successful in an RCT but not acceptable by the target group. We would suggest that once we know from the modeling experiment which interventions show promise, we should carry out a reality check with the target group and understand how accepting they would be if that intervention was rolled out. In a recent study, researchers conducted an RCT to assess the value of SMS text message reminders as a means to improve medication adherence in patients receiving treatment for the prevention of cardiovascular disease [37]. They found that reminders improved adherence rate by 16%, but the interesting point to note is that for this study they contacted 7004 patients and only 303 agreed to receive text reminders. Although the study showed that reminders significantly improved adherence rates of the patients, the participant numbers also hint that many patients might not be interested in this service if it were to be rolled out. Hence, it is important to think about the acceptability of an intervention at the time of selecting interventions that are carried forward from the modeling stage to the RCT stage.

Limitations

The participants were screened on the basis of whether they were able to browse on the internet, as the laboratory experiments involved playing a computer game, which could be considered as a limitation. A significant proportion of the Pakistani population does not know how to use a computer, which restricted us to recruiting a representative sample of the computer-literate Pakistani population rather than a representative sample of the whole Pakistani population. Additionally, one argument that can be made against this modeling experiment is that participants might have adopted a strategy, whereby they enter the code to clear the screen, and once the screen is cleared, they cease to enter the code and maximize the time being spent to play the game, thereby maximize their earning in the experiment. However, this cannot be imagined to be a dominant strategy in the real-world setting, as patients do pay heed to the advice given by their doctors and have a high degree of trust and confidence in their doctors.

Conclusions

For this study, we set up the first behavioral science laboratory in Pakistan. We were keen on having the general public as our participants, and we knew that none of them would ever have participated in an experiment before, making them more receptive to the task. Using a game in the experiment proved to be a very important factor in attracting participants and greatly improved their engagement. As we said earlier, many participants requested to play the game again because they enjoyed it. We believe that more attention should be paid by researchers on how to keep the participants engaged when designing an experiment.

Our investigation answered our research question: whether gamified behavioral modeling is a useful tool to explain real health behaviors and help in identifying which interventions are most likely to work in a randomized trial. We examined a range of interventions that have been tested in real-world settings, and we discovered that our gamified behavioral simulation can model these varied patterns of results. We confirmed our hypothesis that the gamified behavioral simulation can replicate such results, which implies it would be a useful tool for predicting the success of interventions in the field.

Developing behavior change interventions is a complex process and thus requires a systematic way to approach the problem. The process should start from developing a theoretical understanding of the behavior at hand, followed by modeling the behavior to identify the exact mechanisms that might bring about the desired behavior change. Once the most suitable interventions are identified in the modeling stage, then an RCT should be carried out to test the intervention in a real-world setting. In this study, we have showed how this process can be followed in relation to the problem of nonadherence to antibiotic medication. Care must always be taken in extrapolating results from the laboratory to the real world, and caution is particularly called for when the laboratory model abstracts away from some important features such as the feelings of the patients or the environment in which patients are making their decisions. However, generating a hypothesis through simple modeling experiments can help us in developing the most effective interventions that can then be tested in the real-world setting (through RCTs).

Acknowledgments

This research is supported by the National Institute for Health Research Applied Research Centre West Midlands. The views expressed are those of the authors and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

UT, DR, and IV were involved in conceptualizing the study, while UT handled data curation. The investigation was conducted by UT, AG, DR, and IV, with all 4 contributing to the development of the study's methodology. UT and AG performed the statistical analyses. Supervision was provided by DR and IV. UT and AG wrote the original draft, with AG leading the review and editing process. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Information detailing the design and methodology.
[DOCX File, 1828 KB - games_v12i1e47141_app1.docx]

References

1. Chaudri NA. Adherence to long-term therapies: evidence for action. Ann Saudi Med 2004 Jun 3;24(3):221-222. [doi: 10.5144/0256-4947.2004.221]
2. Chesney M. Adherence to HAART regimens. AIDS Patient Care STDS 2003 Apr;17(4):169-177. [doi: 10.1089/108729103321619773] [Medline: 12737640]
3. Horne R, Weinman J, Barber N, et al. Concordance, adherence and compliance in medicine taking. Association of Health Professions in Ophthalmology. 2005 Dec. URL: https://www.ahpo.net/assets/NCCSDO%20Compliance%202005.pdf [accessed 2022-07-21]
4. Bovet P, Burnier M, Madeleine G, Waeber B, Paccaud F. Monitoring one-year compliance to antihypertension medication in the Seychelles. Bull World Health Organ 2002;80(1):33-39. [Medline: 11884971]
5. Graves JW. Management of difficult-to-control hypertension. Mayo Clin Proc 2000 Mar;75(3):278-284. [doi: 10.4065/75.3.278] [Medline: 10725955]
6. van der Sande MA, Milligan PJ, Nyan OA, et al. Blood pressure patterns and cardiovascular risk factors in rural and urban Gambian communities. J Hum Hypertens 2000 Aug;14(8):489-496. [doi: 10.1038/sj.jhh.1001050] [Medline: 10962516]
7. Guo H, He H, Jiang J. Study on the compliance of antihypertensive drugs in patients with hypertension. Chin J Epidemiol 2001 Dec;22(6):418-420. [Medline: 11851054]
8. Checchi KD, Huybrechts KF, Avorn J, Kesselheim AS. Electronic medication packaging devices and medication adherence: a systematic review. JAMA 2014 Sep 24;312(12):1237-1247. [doi: 10.1001/jama.2014.10059] [Medline: 25247520]
9. Hsu C, Lemon JM, Wong ES, et al. Factors affecting medication adherence: patient perspectives from five veterans affairs facilities. BMC Health Serv Res 2014 Nov 13;14:533. [doi: 10.1186/s12913-014-0533-1] [Medline: 25391694]
10. Wroe AL. Intentional and unintentional nonadherence: a study of decision making. J Behav Med 2002 Aug;25(4):355-372. [doi: 10.1023/a:1015866415552] [Medline: 12136497]
11. Aggarwal B, Pender A, Mosca L, Mochari-Greenberger H. Factors associated with medication adherence among heart failure patients and their caregivers. J Nurs Educ Pract 2015;5(3):22-27. [doi: 10.5430/jnep.v5n3p22] [Medline: 25635204]
12. De Croon R, Geuens J, Verbert K, Vanden Abeele V. A systematic review of the effect of gamification on adherence across disciplines. In: Fang X, editor. HCI in Games: Experience Design and Game MechanicsThird International Conference, HCI-Games 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I: Springer; 2021:168-184. [doi: 10.1007/978-3-030-77277-2_14]
13. Altuwayrib SA, Win KT, Freeman M. Gamified medication adherence applications for chronic health conditions: scoping review. In: Meschtscherjakov A, Midden C, Ham J, editors. Persuasive Technology: 18th International Conference, Persuasive 2023, Eindhoven, The Netherlands, April 19–21, 2023, Proceedings: Springer; 2023:307-321. [doi: 10.1007/978-3-031-30933-5_20]
14. Abdul Rahim MI, Thomas RH. Gamification of medication adherence in epilepsy. Seizure 2017 Nov;52:11-14. [doi: 10.1016/j.seizure.2017.09.008] [Medline: 28934624]
15. Tran S, Smith L, El-Den S, Carter S. The use of gamification and incentives in mobile health apps to improve medication adherence: scoping review. JMIR Mhealth Uhealth 2022 Feb 21;10(2):e30671. [doi: 10.2196/30671] [Medline: 35188475]
16. de Vette F, Tabak M, Dekker-van Weering M, Vollenbroek-Hutten M. Engaging elderly people in telemedicine through gamification. JMIR Serious Games 2015 Dec 18;3(2):e9. [doi: 10.2196/games.4561] [Medline: 26685287]
17. Brown M, O'Neill N, van Woerden H, Eslambolchilar P, Jones M, John A. Gamification and adherence to web-based mental health interventions: a systematic review. JMIR Ment Health 2016 Aug 24;3(3):e39. [doi: 10.2196/mental.5710] [Medline: 27558893]

18. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
19. Michie S, Johnston M, Abraham C, et al. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33. [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
20. Kessler JB, Roth AE. Organ allocation policy and the decision to donate. Am Econ Rev 2012 Aug;102(5):2018-2047. [doi: 10.1257/aer.102.5.2018] [Medline: 29115804]
21. Robitaille N, Mazar N, Tsai CI, Haviv AM, Hardy E. Increasing organ donor registrations with behavioral interventions: a field experiment. J Mark 2021 May;85(3):168-183. [doi: 10.1177/0022242921990070]
22. Stoler A, Kessler JB, Ashkenazi T, Roth AE, Lavee J. Incentivizing authorization for deceased organ donation with organ allocation priority: the first 5 years. Am J Transplant 2016 Sep;16(9):2639-2645. [doi: 10.1111/ajt.13802] [Medline: 27013023]
23. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106. [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
24. Schmidt-Kraepelin M, Warsinsky S, Thiebes S, Sunyaev A. The role of gamification in health behavior change: a review of theory-driven studies. Presented at: International Conference on System Sciences; Jan 7-10, 2020; Hawaii. [doi: 10.24251/HICSS.2020.155]
25. Kuntz JL, Safford MM, Singh JA, et al. Patient-centered interventions to improve medication management and adherence: a qualitative review of research findings. Patient Educ Couns 2014 Dec;97(3):310-326. [doi: 10.1016/j.pec.2014.08.021] [Medline: 25264309]
26. Kini V, Ho PM. Interventions to improve medication adherence: a review. JAMA 2018 Dec 18;320(23):2461-2473. [doi: 10.1001/jama.2018.19271] [Medline: 30561486]
27. Zullig LL, Peterson ED, Bosworth HB. Ingredients of successful interventions to improve medication adherence. JAMA 2013 Dec 25;310(24):2611-2612. [doi: 10.1001/jama.2013.282818] [Medline: 24264605]
28. Alsan M, Beshears J, Armstrong WS, et al. A commitment contract to achieve virologic suppression in poorly adherent patients with HIV/AIDS. AIDS 2017 Jul 31;31(12):1765-1769. [doi: 10.1097/QAD.0000000000001543] [Medline: 28514277]
29. Klein DE. The forest and the trees: an integrated approach to designing adherence interventions. AMJ 2008 Nov 1;2009(13):181-184. [doi: 10.4066/AMJ.2009.83] [Medline: 26213585]
30. Volpp KG, Loewenstein G, Troxel AB, et al. A test of financial incentives to improve warfarin adherence. BMC Health Serv Res 2008 Dec 23;8:272. [doi: 10.1186/1472-6963-8-272] [Medline: 19102784]
31. Volpp KG, Pauly MV, Loewenstein G, Bangsberg D. P4P4P: an agenda for research on pay-for-performance for patients. Health Aff (Millwood) 2009;28(1):206-214. [doi: 10.1377/hlthaff.28.1.206] [Medline: 19124872]
32. John LK, Loewenstein G, Troxel AB, Norton L, Fassbender JE, Volpp KG. Financial incentives for extended weight loss: a randomized, controlled trial. J Gen Intern Med 2011 Jun;26(6):621-626. [doi: 10.1007/s11606-010-1628-y] [Medline: 21249462]
33. Vervloet M, Linn AJ, van Weert JCM, de Bakker DH, Bouvy ML, van Dijk L. The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: a systematic review of the literature. J Am Med Inform Assoc 2012;19(5):696-704. [doi: 10.1136/amiajnl-2011-000748] [Medline: 22534082]
34. Menon V, Selvakumar N, Kattimani S, Andrade C. Therapeutic effects of mobile-based text message reminders for medication adherence in bipolar I disorder: are they maintained after intervention cessation? J Psychiatr Res 2018 Sep;104:163-168. [doi: 10.1016/j.jpsychires.2018.07.013] [Medline: 30081390]
35. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. Psychiatry Res 2012 Dec 30;200(2-3):89-95. [doi: 10.1016/j.psychres.2012.07.034] [Medline: 22901437]
36. Santo K, Singleton A, Rogers K, et al. Medication reminder applications to improve adherence in coronary heart disease: a randomised clinical trial. Heart 2019 Feb;105(4):323-329. [doi: 10.1136/heartjnl-2018-313479] [Medline: 30150326]
37. Wald DS, Bestwick JP, Raiman L, Brendell R, Wald NJ. Randomised trial of text messaging on adherence to cardiovascular preventive treatment (INTERACT trial). PLoS One 2014 Dec 5;9(12):e114268. [doi: 10.1371/journal.pone.0114268] [Medline: 25479285]

Abbreviations

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 09.03.23; peer-reviewed by A Biro, R Gorantla, T Baranowski; revised version received 15.02.24; accepted 17.03.24; published 24.09.24.

Please cite as:
Taj U, Grimani A, Read D, Vlaev I
Using Games to Simulate Medication Adherence and Nonadherence: Laboratory Experiment in Gamified Behavioral Simulation
JMIR Serious Games 2024;12:e47141
URL: https://games.jmir.org/2024/1/e47141
doi:10.2196/47141

Acceptability of a Serious Game About Proton Radiotherapy Designed for Children Aged 5 to 14 Years and Its Potential Impact on Perceived Anxiety: Feasibility and Randomized Controlled Pilot Trial

Catarina Cederved[1,2], PhD; Gustaf Ljungman[1], Prof Dr Med; Jon Back[3], PhD; Charlotte Ångström-Brännström[1,4], PhD; Gunn Engvall[1], PhD

1
2
3
4

Corresponding Author:
Catarina Cederved, PhD

Abstract

Background: Children who are going to undergo radiotherapy have displayed fear and anxiety. Therefore, a web-based serious game was developed as a psychological preparation to investigate if it could affect anxiety levels. In an earlier stage, children with experience of radiotherapy had been part of the developmental process.

Objective: The study aimed to investigate the feasibility in terms of reach, usability, and acceptability of a serious game about proton radiotherapy and to pilot that it did not increase anxiety levels in children aged 5 to 14 years undergoing radiotherapy.

Methods: The design was a randomized controlled pilot trial with predefined feasibility criteria. In total, 28 children were assessed for eligibility, and 23 met the inclusion criteria. They were consecutively randomized into 1 of 2 study arms. One child was excluded after randomization. If randomized into arm 1, the children received the intervention before treatment started. Children in arm 2 were treated as controls. Questionnaires with fixed answers were used to assess anxiety levels (an adapted version of the State-Trait Anxiety Inventory for Children) and experiences of gameplay (an adapted version of Player Experience of Need Satisfaction [PENS]). The children were asked to answer questionnaires at 5 different measurement occasions during their radiotherapy treatment.

Results: In arm 1, age ranged from 5 to 13 (mean 8.4, SD 2.4) years. In arm 2, age ranged from 5 to 11 (mean 7.6, SD 2.3) years. The sample consisted of 15 girls and 7 boys. The feasibility criterion that the children should play the game for 20 minutes or more was not met. Mean playtime for children in arm 1 was 32.1 (SD 23.8) minutes, where 18 children had played for at least 15 minutes. The criterion that 70% (n=16) or more of the participants should return all of the questionnaires was not met; however, more than 73% (n=16) returned the PENS questionnaires. The State-Trait Anxiety Inventory for Children was returned by 73% (n=16) on day 0, 77% (n=17) on day 1, 82% (n=18) on day 3, 82% (n=18) on day 6, and 86% (n=19) on day 15.

Conclusions: All feasibility criteria set for the study were not met, suggesting that adaptions need to be made if a future study is to be undertaken. Further, the analysis revealed that there was no indication that playing increased the children's self-reported anxiety. The PENS questionnaire adapted for children showed promising results regarding player satisfaction when using the serious game. When studying children with severe conditions and young age, 5 measurement occasions seemed to be too many. Measuring both player satisfaction or experience and knowledge transfer would be preferable in future studies.

Trial Registration: ClinicalTrials.gov NCT04728555; https://clinicaltrials.gov/study/NCT04728555

(JMIR Serious Games 2024;12:e54082) doi:10.2196/54082

KEYWORDS

anxiety; feasibility; acceptability; pediatric oncology; psychological preparation; proton radiotherapy; serious game; games; cancer

Introduction

Background

Children with cancer are subjected to medical procedures that are both tiresome and painful [1]. Some children need to undergo radiotherapy. The treatment does not hurt but can be perceived as frightening [2]. Radiotherapy is usually given over a longer period of up to 6 weeks [3]. When radiotherapy is performed, the children are left alone in the enormous treatment room [4]. Since the children need to be completely still during the procedure, they are also fixated to stay immobilized [4,5]. When radiotherapy is targeted at the head, a mask is made and firmly fixated to the table [6]. The experience has been described as stressful and invoking anxiety and fear in both the children undergoing radiotherapy and their parents [7,8]. Several of the children are sedated, especially preschool children [2] because of the anxiety the procedure induces in the children [9]. Informing and preparing children who are going to undergo radiotherapy in order to decrease sedation or anesthesia (hereafter mentioned as sedation) or anxiety levels have been studied, and the effects have been minor [10,11]. However, with extensive psychological preparation, a decrease in the number of sedations was reported by Clerici et al [12]. An increased proportion of proton beam therapy is given as the preferred irradiation type, as it lessens short- and long-term toxicities and improves the quality of life outcomes for the children [13]. To receive proton therapy, the family often needs to go to a clinic far away from their home since only a few clinics provide such treatment [14,15]. In Sweden, where the study was executed, there is one national clinic providing proton beam therapy. Leaving their familiar surroundings can lead to further stress [16] on top of the pain and distress already caused by other cancer-related procedures the child has to endure [17]. It can therefore be argued that it is necessary to prepare the children by informing them about what radiotherapy is, what they can expect from the treatment that they are going to receive, and introducing them to the environment of the clinic so that they can feel that they are as much in control of the situation as possible [18].

In school, educational games have been proven to affect children's learning [19]. Games can be enjoyable, create engagement, and induce behavior changes [20]. Using digital tools and games to prepare children before procedures and having them take part in the development process are becoming more common [21-23]. Children with cancer have suggested that information should be given to elevate understanding about a specific treatment, which also encourages their ability to cope with the procedure [24]. Knowledge about the procedure increased, and the level of anxiety decreased with a digital tool when tested on children who were to undergo planned hospital procedures [25]. Several games have been developed with the intent to increase knowledge about the disease and treatment for children with cancer [26]. To decrease anxiety, a serious game about proton radiotherapy was developed together with children to function as interactive information before the procedure [23]. However, how well the game functions needs to be evaluated. Therefore, as psychological preparation before radiotherapy, a serious game as a supplement to preparation already in place at a proton clinic was evaluated for acceptability and impact.

Aim

The aim of the study is to investigate the feasibility in terms of reach, usability, and acceptability of a serious game about radiotherapy and to pilot that it does not increase self-reported anxiety in children aged 5 to 14 years undergoing radiotherapy.

Methods

Study Design

The study was designed as a feasibility and randomized controlled pilot study. The study provided an intervention using a prospective, waiting list control design.

Serious Game

For the intervention, children who would be undergoing radiotherapy played a web-based serious game on their tablet or computer to familiarize themselves with the radiotherapy treatment. The parents received a link to the game via email so that the children could access the game whenever they wanted. They used their own devices to play the game; however, they could also use the computers at the hotel they stayed at in conjunction with the treatment. No game controller was used, instead, the game was played with the computer's mouse or by touching the tablet's touchscreen. The game was designed as a doll house experience [27] where the children played an avatar that was in a proton radiotherapy clinic. In the game, a map was used that the child could click on to access different rooms in the clinic to explore them independently. The rooms contained information about the procedure of radiotherapy for educational purposes as displayed in Figures 1 and 2. It also contained mini-games and game elements that were placed there to make exploration of the play setting interesting [23]. The game had been developed through an iterative process together with 9 children with experiences of radiotherapy treatment, which influenced the game's design and led to numerous changes in it [23]. The changes pertained to the design of how the treatment was displayed and explained within the game and the language [23]. To be able to explore the clinic's facilities and get informed about the procedure of radiotherapy is a form of psychological preparation [28]. The purpose of psychological preparation is to mitigate the fear of the unknown [29] and is intended to reduce anxiety before procedures [30].

Figure 1. Screenshot from the game the participants were subjected to within the randomized controlled trial showing the first encounter of where radiotherapy is performed and where the avatar will later undergo radiotherapy.

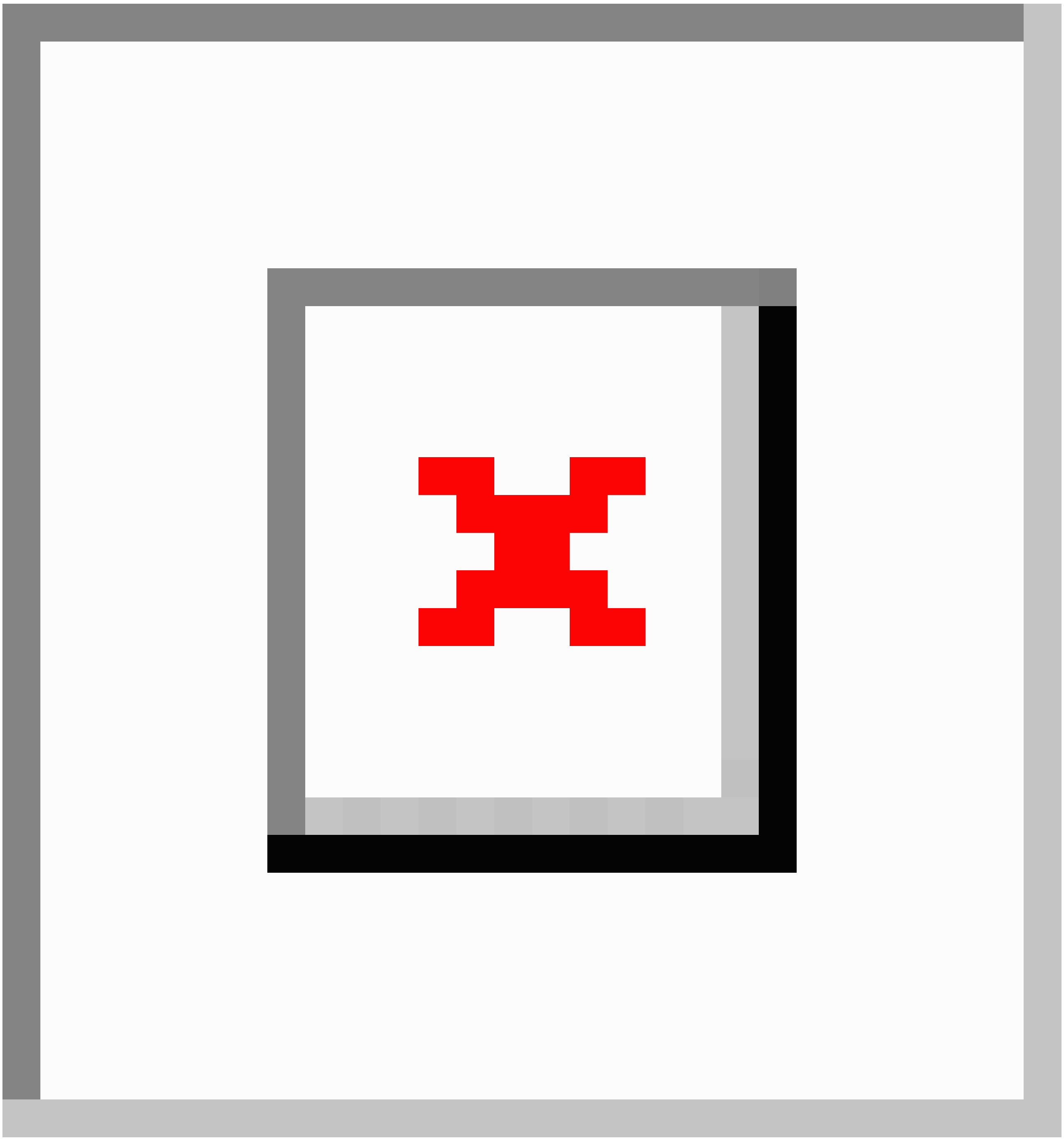
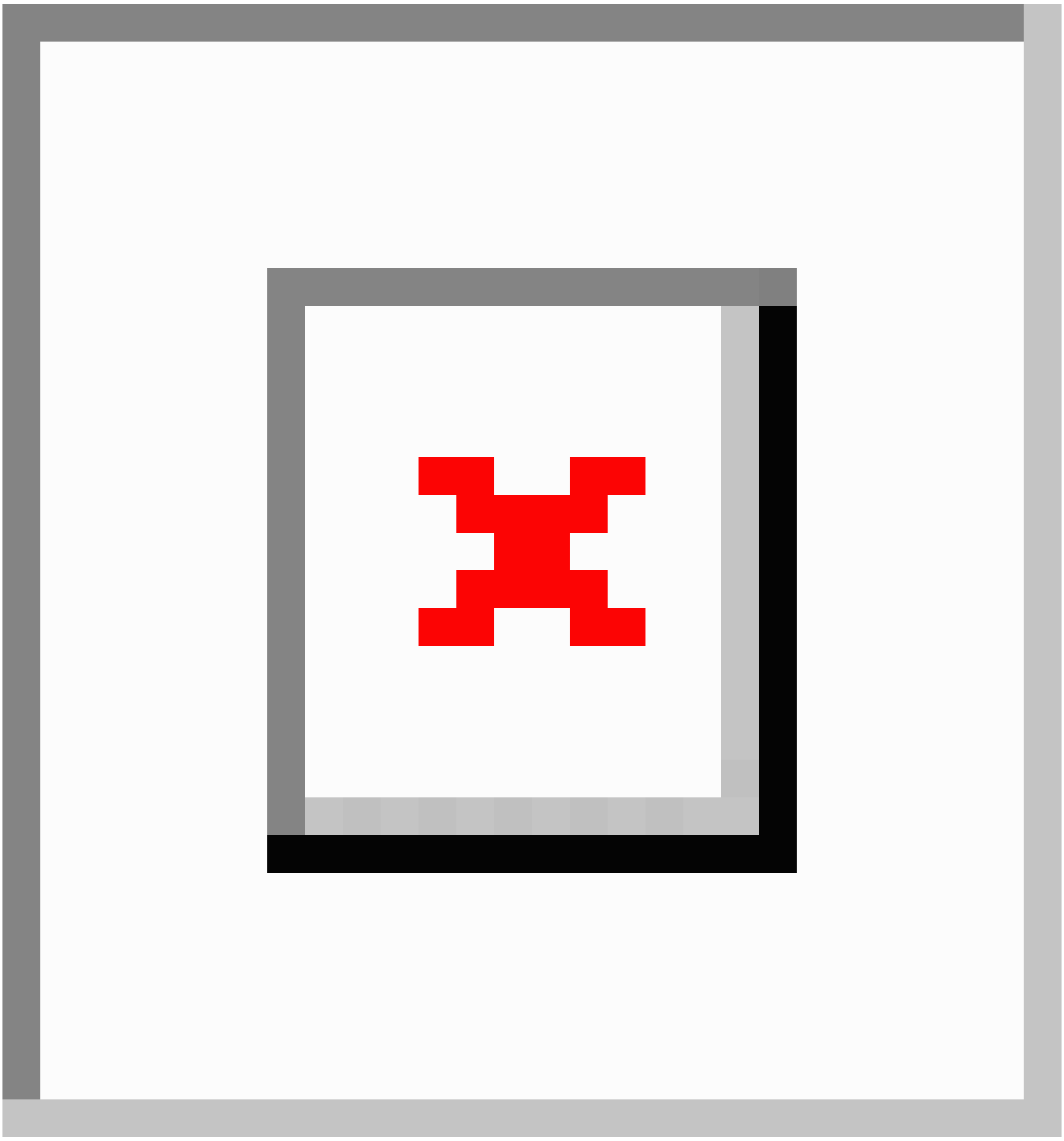


Figure 2. Screenshot from the game the participants were subjected to within the randomized controlled trial showing the monitor room where the staff provide therapy and observe the avatar undergoing radiotherapy.



Recruitment

Recruitment started in February 2021 and completed in August 2022. During that period, 26 children were considered eligible for inclusion. When children between the ages of 5 and 15 years, living in Sweden, had been scheduled to have radiotherapy at the clinic, a written invitation about the study was sent to them by post. A few days later, a pediatric or oncology nurse or a coordinator contacted the family by telephone with information about the study and asked for permission for the researcher to contact them. Upon agreement, the researchers contacted them and gave further information orally, and if oral consent was granted, they were then emailed all further information about the study. Exclusion criteria were children living in countries other than Sweden, inability to understand Swedish, and severe mental disability. Two children were excluded due to language barriers during recruitment.

The children were divided into 3 age groups: ages 5-7, 8-10, and 11-14 years. They were then randomly assigned into 1 of 2 study arms using stratified randomization in blocks to achieve balanced groups. The parents received information about which arm their child had been randomized to. Study arm 1 received the intervention 1 or more days before starting their therapy. Children assigned to arm 2 received the game 3 days after they started their therapy since it was not considered ethical to invite the children to a game study and then not let them play the game [31].

Ethical Considerations

Ethics approval was received by the Swedish Ethical Review Authority (2020‐05578) before the study was started. The study was registered at ClinicalTrials.gov (NCT04728555). Upon agreement to participate in the study, the children received a 6-digit participation ID, which was used to sign into the game to be able to track time played, and the number was also used as identification for the questionnaires; hence, the data became anonymous. There was one separate list where participants' names and contact information were kept together with their participation ID. The children's legal guardian signed a written consent form on their child's behalf and their own behalf. Children also gave their assent to participate. Neither the children nor their guardians received any compensation for participating in the study.

Feasibility Criteria and Measurements

The feasibility criterion for success was that 80% (n=18) of the participants had played the game for 20 minutes or more. Further, 70% (n=17) or more of the participants should have returned all questionnaires partially or fully answered State-Trait Anxiety Inventory for Children-State Anxiety (STAIC(S)) and Player Experience of Need Satisfaction (PENS). The premises of the threshold were set to find out which days most children filled out the STAIC(S) to establish for future studies how many measurement occasions are plausible. A hypothesis was tested, which was that children in arm 1 should not communicate more anxiety than children allocated to arm 2. In addition, reach was measured by assessing whether prospective participants received information about the study.

There were 4 different questionnaires distributed to participating children: the STAIC(S) and State-Trait Anxiety Inventory for Children-Trait Anxiety [32], the PENS [33], and a customized questionnaire about radiotherapy developed by the research team (which will be reported elsewhere). Parents answered an adult version of the State-Trait Anxiety Inventory (STAI), which included 10 questions on a 4-point Likert scale [34]. They also answered a questionnaire providing background variables. Children who were not able to read were helped by their parents to answer the questionnaires and interpret the scale alternatives.

An adapted version of the short STAIC with a 3-point Likert scale was used to measure the children's anxiety [32]. The short form of STAIC consists of 6 questions; however, 2 questions were added from the long version of STAIC to provide further insight into anxiety, making a range from 8 to 24 points. Hereafter, the adapted version used in the intervention will be referred to as STAIC(S). According to the guidelines of interpretation, a calculation based on 8 questions indicated 12 points as the cutoff for anxiety. To not overestimate the children's anxiety, the cutoff was set at 13 points for feeling anxious. The State-Trait Anxiety Inventory for Children-Trait Anxiety is a questionnaire consisting of 10 questions on a 3-point Likert scale [32].

The PENS questionnaire consists of 16 questions on a 7-point Likert scale [33]. Approval was given for the questionnaire to be translated into Swedish and adapted for children by the owner company, Immersyve. It was first translated from English into Swedish by a group consisting of 2 pediatric nurses, 1 researcher in informatics and media, and 1 pediatric psychologist and then backward by an interpreter with English as their mother tongue. Two questions were related to multiplayer games, which were not relevant to the study since the game is a single-player game, so they were excluded. The questionnaire was tested for face validity on 4 healthy children between the ages of 5 and 11 years through interviews. The questionnaire was made in 2 versions, 1 for younger children (5‐7 years) and 1 for older children (8‐14 years). The younger population had difficulty understanding the 7-point Likert scale. Therefore, it was changed to a 3-point Likert scale. Furthermore, the language was adapted to the young population and consisted of 8 questions (range 8‐24). For the older children (8‐14 years), the questionnaire consisted of 14 questions (range 14‐98). The cutoff level for satisfaction with the game experiences was set at 16 for PENS (5‐7 years) and at 65 for PENS (8‐14 years), reaching two-thirds of the total score (66%) for both scales. The fourth questionnaire (customized questionnaire about radiotherapy) included 6 questions about radiotherapy on a 4-point Likert scale.

Data Collection

At 1‐3 days prior to the start of treatment, the children randomized into arm 1 (intervention) received a web link to the game, and children randomized into arm 2 (control) received the web link after their third day of treatment. The children and parents received instructions stating that they were free to play the game as much or little as they liked. Each participant was given a unique participation number sent to their guardian's email, which was used to access the game. A software engine was used to store the playtime for each number, and the data were deleted after a month. A member of the team who only had the participation number collected the data from the engine. The playtime was measured cumulatively, and no record was kept on how many times each participant had chosen to access the game. For 2 players, the information was lost. Therefore, the parents were emailed to make an estimation of their children's playtime. Children in both arms answered the same questionnaires 1 day before treatment started (day 0), the first day of treatment, the third day of treatment, the sixth day of treatment, and the last questionnaire approximately 15 (±4) days into treatment. The children received a diary notes form on their first day at the clinic where they were asked open questions about how they perceived the game, which will be reported elsewhere. The questionnaires were administered to the children and parents by the receptionist staff at the clinic before they were due to have the treatment. The answered questionnaires were left in a mailbox at the clinic. The children and parents were asked to arrive before their appointment on the days of the study so that they would have time to answer the questionnaires before that day's treatment. The total time children spent playing the game was collected digitally for each participant. STAIC(S) was collected on all 5 measuring occasions, PENS at 1 time, and radiotherapy questionnaire on 4 occasions. Parents answered demographic questions and STAI at the first measuring point. During participation in the study, the children and their parents received information according to standard care.

Data Analysis

Descriptive statistics were used to gain insights into the collected quantitative data. To analyze differences between the arms, Fisher exact test, Mann-Whitney *U* test, Pearson chi-square, and Wilcoxon matched-pair signed rank were used. Spearman rank order correlation was used to analyze statistically significant associations between variables within the sample. In addition, Cronbach α was used to calculate the internal consistency between concepts. SPSS (version 28.0; IBM Corp) was used for the analyses, and the findings were considered to be statistically significant if a *P* value <.05 was reached. The CONSORT (Consolidated Standards of Reporting Trials) guideline was followed for statistical analysis and to report the study [35].

Results

Feasibility in Terms of Reach and Background Variables

The study aimed to reach all patients within the age group who were to have radiotherapy. After the study had been finished, the nurses at the radiotherapy clinic went through the records of all children who had passed the clinic within the study's timeframe. There had been 28 children who had been assessed for inclusion, of which, 26 of the children were eligible to take part in the study (Figure 3). A total of 3 children declined participation, and 1 child was lost to participation due to administrative issues, resulting in 22 (85%) children's completion of at least one questionnaire of the study. There were no differences between the intervention arm and controls regarding sex, age, diagnosis, sedation, and parents' educational level.

In total, 11 children were randomized into the intervention arm, and another 11 into the control arm. In arm 1, age ranged from 5 to 13 (mean 8.4, SD 2.4) years. In arm 2, age ranged from 5 to 11 (mean 7.6, SD 2.3) years. For the variances of ages, see Table 1. The sample consisted of 15 girls. A majority of the children (n=16) were diagnosed with brain tumors and 6 with extracranial solid tumors. The number of fractions received varied from 14 to 33 with a mean of 27 (SD 5.7) fractions. In total, 11 children received sedation, and 6 of these children received the intervention. All 5- and 6-year-old children were sedated, and 3 of 4 children among the 7-year-olds were also sedated. The sedated children's answers are included in the analysis. Most of the children (n=21) lived with 2 parents, while 1 child lived with a single parent. A total of 16 children lived with a parent who had attended higher education and 4 lived with a parent who had finished high school, and there were 2 children for whom no data were attained.

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flowchart depicting enrollment and the days the questionnaires were administered and how many participants returned them for each arm. PENS: Player Experience of Need Satisfaction; RT.Q.: customized questionnaire about radiotherapy; STAI: State-Trait Anxiety Inventory; STAIC(S): State-Trait Anxiety Inventory for Children-State Anxiety; STAIC-T: State-Trait Anxiety Inventory for Children-Trait Anxiety.

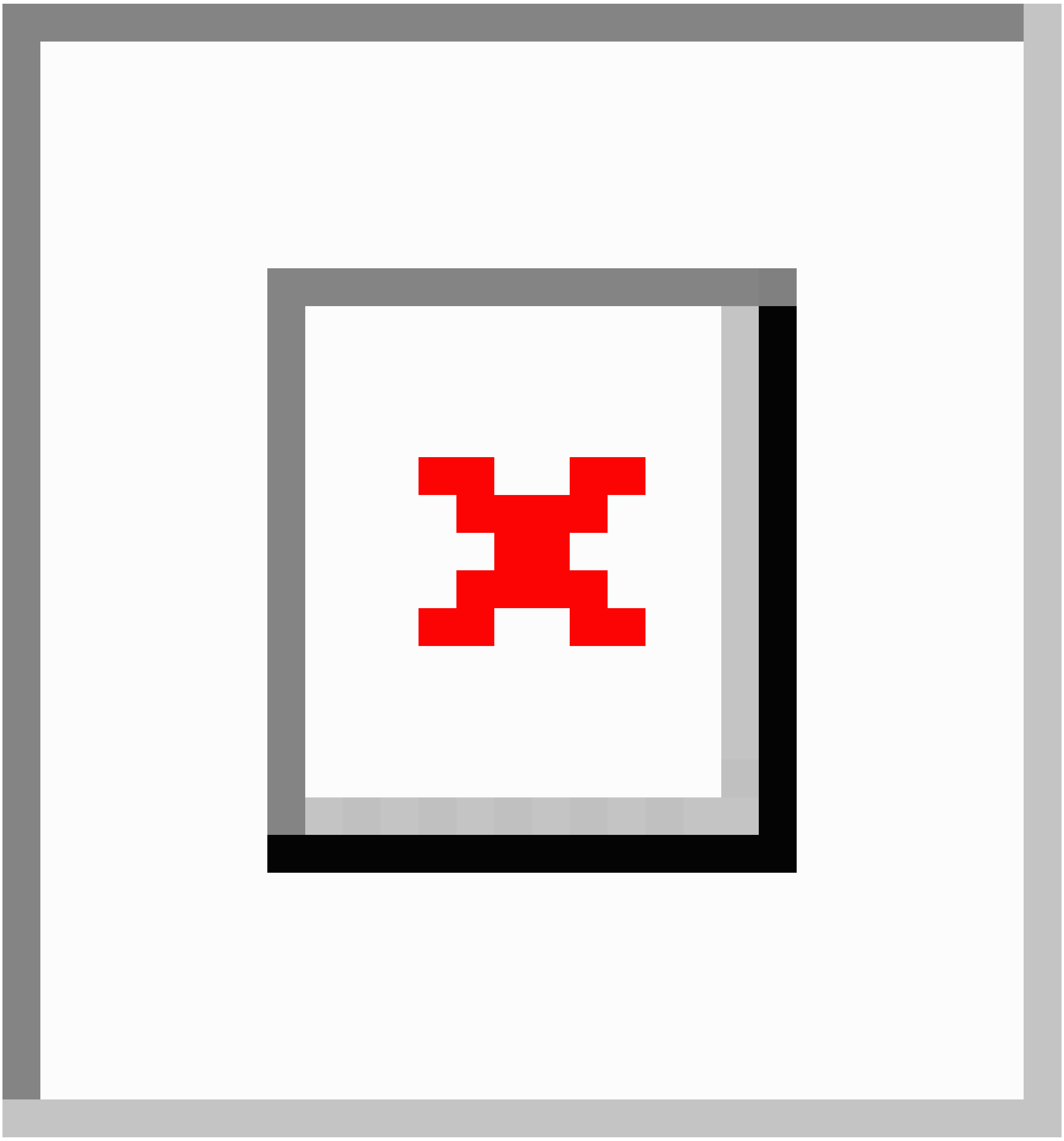


Table . Children's (n=22) background variables divided by allocated arm. Displaying number of children's sex, age groups, sedated or awake status during radiotherapy, diagnosis, and parental education level with the total amount of each arm for each variable and total when merging the 2 arms.

		Arm 1, n (%)	Arm 2, n (%)	Total, n
Sex				
	Boy	2 (18)	5 (45)	7
	Girl	9 (82)	6 (55)	15
	Total	11	11	22
Age groups (years)				
	5 - 7	5 (50)	5 (50)	10
	8 - 10	4 (44)	5 (56)	9
	11 - 14	2 (67)	1 (33)	3
	Total	11	11	22
Sedated				
	Sedated	6 (55)	5 (45)	11
	Awake	5 (45)	6 (65)	11
	Total	11	11	22
Diagnosis				
	Brain tumor	8 (50)	8 (50)	16
	Extracranial solid tumor	3 (50)	3 (50)	6
	Total	11	11	22
Level of parents' education				
	Higher education	9 (56)	7 (44)	16
	High school	1 (25)	3 (75)	4
	Total	10	10	20

Feasibility in Terms of Usability and Acceptability of the Intervention

The predefined feasibility criterion was that 80% (n=18) of the children should play for at least 20 minutes to determine whether they found the game acceptable. There were 6 (27%) participants who played the game for less than 20 minutes, and 18 (82%) of the participants played for more than 15 minutes; hence, the criterion was not fulfilled. However, the mean time playing was 31 minutes for the entire sample (median of 27 minutes) spanning from not playing the game at all to playing for 85 minutes (Table 2). The intervention arm had a mean time playing the game of 32.1 (SD 23.8) minutes, while the mean time was 29.9 (SD 19.0) minutes for the control arm. There were no differences in playtime between the two.

To measure the participants' self-rated anxiety, the STAIC(S) was administered at 5 measurement points. The estimation was that 70% (n=17) of the participants should return the forms partially or fully completed. In total, 9 (41%) children returned all 5 forms of STAIC(S), and 13 (59%) children failed to return 1 or more forms. To measure the usability of the game, their experiences of the game were measured through the 2 versions of the PENS questionnaire, which 16 (73%) participants returned. A total of 8 participants answered the 5 - 7 years version, and 8 participants answered the PENS questionnaire (8 - 14 years).

Table . Description of children's playtime of the serious game, Player Experience of Need Satisfaction (PENS) score, and if the child was sedated or awake during radiotherapy (n=22).

Participant	Age (years)	Time played (minutes)	PENS score	Sedated
Intervention[a]				
E1	5	29	11[b]	Sedated
E2	7	18	19[b]	Sedated
E3	7	20	18[b]	Awake
E4	7	58	19[b]	Sedated
E5	7	85	18[b]	Sedated
E6	8	10	—[c]	Sedated
E7	8	23	45[d]	Awake
E8	8	27	70[d]	Sedated
E9	10	45	—	Awake
E10	12	0	—	Awake
E11	13	38	67[d]	Awake
Control [e]				
L1	5	0	—	Sedated
L2	5	10	—	Sedated
L3	5	20	16	Sedated
L4	5	60	22	Sedated
L5	6	23	20	Sedated
L6	9	17	34	Awake
L7	9	35	66	Awake
L8	9	38	—	Awake
L9	9	60	88	Awake
L10	10	39	47	Awake
L11	11	27	26	Awake

[a]Age: mean 8.4 (SD 2.4) years and time played: mean 32.1 (SD 23.8) minutes.

[b]PENS (children 5 - 7 years) score range 8 - 24, mean 17.9 (SD 3.3) based on both arms (n=8).

[c]Missing value.

[d]PENS (children 8 - 14 years) score range 14 - 98, mean 55.4 (SD 20.8) based on both arms (n=8).

[e]Age: mean 7.6 (SD 2.3) years and time played: mean 29.9 (SD 19.0) minutes.

Anxiety

To measure whether children who received the intervention communicated an increase in anxiety compared to the controls on their first day of radiotherapy, the Mann-Whitney *U* test was used. No perceived differences between the arms could be found (*P*=.81). There were no differences measured in anxiety between the arms across all 5 assessment occasions.

The number of children who indicated anxiety (13 - 24 points) through the STAIC(S) is presented in Table 3, where the children receiving the intervention and the control arm are displayed separately. There were 6 children sedated in the intervention arm and 5 in the control. The following response rates at each measurement occasion were collected: 73% (n=16) on day 0, 77% (n=17) on day 1, 82% (n=18) on days 3 and 6, and 86% (n=19) on day 15 (Table 3). Over time, the children's anxiety decreased, but no significant difference could be found between the arms. There was no significant correlation found between STAIT (trait anxiety) and STAIC(S) (state anxiety) in the current sample (*P*=.34). The items' internal consistency of STAIC(S) was calculated by Cronbach α based on 89 questionnaires that had been obtained during the study, and the α value was 0.85.

On day 0, the parental STAI questionnaire was administered to 43 parents, and 33 answers were obtained. The hypothesis was that if the parents presented higher levels of anxiety, so would their children. Since there were considerably more mothers who returned the parental STAI, the hypothesis was only tested between mothers and their children. No correlation could be found between the mother's reported trait anxiety levels and the children's before treatment started. No correlation could be found between the mother's state anxiety levels and those of their children.

Table . The number of children reporting anxiety (13 - 24 points) through the State-Trait Anxiety Inventory for Children-State Anxiety (STAIC(S)) questionnaire at 5 measurement occasions (n=22).

Measurement occasions	Intervention			Total, n	Control			Total, n
	Awake, n (%)	Sedated, n (%)	Total, n		Awake, n (%)	Sedated, n (%)	Total, n	
Day 0 (n=16)	2 (29)	4 (57)	6	7	3 (33)	2 (22)	5	9
Day 1 (n=17)	1 (11)	4 (44)	5	9	4 (50)	2 (25)	6	8
Day 3 (n=18)	1 (11)	3 (33)	4	9	2 (22)	1 (11)	3	9
Day 6 (n=18)	2 (22)	3 (33)	5	9	1 (11)	1 (11)	2	9
Day 15 (n=19)	2 (20)	2 (20)	4	10	1 (11)	1 (11)	2	9

Serious Game About Radiotherapy

No correlation between time playing the game and anxiety levels measured on day 1 was found in the intervention arm nor was it found when combining the 2 arms on day 6. To check the translated and modified PENS questionnaire's items for internal consistency and accuracy, we used the concepts that were established in PENS (version 1.6). PENS (5 - 7 years) consisted of 8 questions: 2 items related to competence, 2 items to autonomy, and 4 items to presence, with a sample unit of 8. Six respondents were girls. The Cronbach α score for competence was 0.89, autonomy 0.57, and presence 0.28. Descriptive statistics of the concepts for PENS (5 - 7 years) are reported in Table 4. For PENS (5 - 7 years), the score could range between 8 and 24, and the mean score was 17.9 (SD 3.3; Table 2). No correlation between the time playing the game and the scoring of the questionnaire for the PENS (5 - 7 years) was established. In total, 7 (88%) of the participants had a score that was 16 or higher, indicating that they found the game experience satisfying (Table 2).

PENS (8 - 14 years) consisted of 14 questions, and there were 4 items related to competence, 3 to autonomy, and 7 to presence, with a sample unit of 8. Descriptive statistics of the concepts for PENS (8 - 14 years) are reported in Table 5. The Cronbach α score for competence was 0.85, autonomy was 0.87, and presence was 0.91. For PENS (8 - 14 years), the score could range between 14 and 98, and the mean score was 55.4 (SD 20.8; Table 2). A correlation between the time playing the game and the scoring of the questionnaire for the PENS (8 - 14 years) could not be found with a *P* value of .05. However, it is noteworthy that it resulted in a *P* value of .06. In total, 4 (50%) of the participants had a score that was 65 or higher, indicating that they found the game experience satisfying.

Table . The three concepts by subscales present in Player Experience of Need Satisfaction (5-7 years; calculated from replies by 8 participants).

Concepts	Scores from participants aged 5 - 7 years on a scale ranging from 1 to 3		
	Min	Max	Mean
Competence	2.38	2.63	2.50
Autonomy	1.63	2.50	2.06
Presence	1.88	2.50	2.13

Table . The three concepts by subscales in Player Experience of Need Satisfaction (8-14 years; calculated from replies by 8 participants).

Concepts	Scores from participants aged 8 - 14 years on a scale ranging from 1 to 7		
	Min	Max	Mean
Competence	3.89	6.11	5.22
Autonomy	3.78	4.78	4.11
Presence	2.63	4	3.20

Discussion

Principal Findings

Web-based games are easily accessible and can work as psychological preparation for children with cancer [26]. Preparing children who are going to undergo radiotherapy with the help of a serious game could be a means to make the experience somewhat less frightening. All 3 predefined feasibility criteria of the study were not met. Less than 80% of the children played the game for 20 minutes or more. More than 70% returned the PENS questionnaire. Less than 70% returned all the questionnaires of STAIC(S). Although, on the last 3 occasions of assessment, more than 80% of the questionnaires were returned. The results did not show that the serious game increased the children's anxiety, and over the trajectory of the study's time, there was a decrease in anxiety levels in the 2 arms. The study achieved its desired reach by recruiting through nurses who contacted children who were to visit the clinic. Randomization was also done according to the pilot study protocol.

Anxiety

Since the sample size was small, no difference could be found in anxiety levels between the ones receiving the intervention and the control arm, which was similar to findings in earlier studies [10,11]. The study relies on the participating children being able to identify their own feelings and express them accordingly to the scale in the STAIC(S). The scale, in its full version, is one of those most used in research and has been found to be both valid and reliable among children with cancer [36]. Children need to have had experiences of feelings to have learned their meaning if they are to be able to label them with words on a questionnaire. [37]. Nevertheless, it is preferable that children themselves report their symptoms [38]. The STAIC(S) is not validated for younger children (5 - 7 years), which means another instrument would be preferable; for example, the STAIC(S) redesigned and including pictures aimed at younger children could be an option [39].

The sample included children who were sedated. Children who are sedated are not at risk of not being still during radiotherapy [5]; nevertheless, they still can feel fear and anxiety concerning the procedure [40]. Therefore, to include their answers about how they perceived the game and rated their anxiety is valuable to report since few earlier studies have measured children's self-rated anxiety when undergoing radiotherapy.

Since 2 items were added to the short form of STAIC (including 8 instead of 6 questions), Cronbach α was used to assess the internal consistency by reliability tests, and the α value was 0.85, which is considered to prove high internal consistency between the items [41].

Playing Time and PENS Scores

The feasibility criteria of the study were not met since fewer children than predicted played the game for 20 minutes or more. The children were free to play the game as much or little as they liked, resulting in a few children not playing the game at all. However, the time spent playing the game was almost the same in the arms, indicating an interest in the game among the children participating in the study. Other studies have used various specified game dosages and a set time playing the intervention, which in some cases resulted in a correlation to statistically significant findings [42]. To not have a predetermined play time but instead monitor how much time children spend playing the game is a way to assess whether children find serious games acceptable [43]. Cronbach α was used to assess the internal consistency through reliability tests of the 2 adapted PENS questionnaires developed for the study of the concepts [41]. For PENS (8 - 14 years), the reliability was high. However, the number of participants that the calculation was established for was low. Therefore, the questionnaire should be validated by factor analysis, and reliability tested on a larger sample of children in a Swedish context [44]. Further, the validation of the English version was conducted using university students [33]. For PENS (5 - 7 years), the reliability was low for 2 of the concepts and high for 1. This could be an effect of the change from the 7-point Likert scale to a 3-point Likert scale and that it consists of fewer questions. Therefore, the PENS (5 - 7 years) also needs to be validated for children in a Swedish context.

Methodological Considerations

Following the intention-to-treat analysis, the 2 children who never opened or played the game are part of the analyzed material as they answered questions. For future studies, predetermined play time of the intervention could be a means to establish whether it has an effect on what the study measures [42] and exclude those not achieving predetermined play time in the analysis. Fewer children than predicted returned all of the STAIC(S) questionnaires. For future studies, fewer assessment point days including fewer questions would be preferred when researching the population. In addition, another option to evaluate the effect of the intervention could be to measure the number of sedations before and after the intervention. For the study to reach power, a power calculation was made showing that 60 children needed to be enrolled. A strength of the study was that the clinic involved has a nationwide uptake and therefore reached all children within the age groups. However, after 1.5 years, only 28 children had been assessed for eligibility. The study was run during the COVID-19 pandemic, which might have affected referrals to the clinic. Despite prospective inclusion, there were more girls than boys included, which could be due to chance on account of the small sample size. Future studies need to have an even sex distribution to ensure generalizability. A weakness in the study was that it was unable to reach the estimated goal for power; and therefore, the criterion of a difference in the anxiety levels between the arms could not be found. With a larger study sample size, the criterion might have been met. Therefore, collaboration with other proton clinics is warranted to obtain a larger sample. Hence, the evaluation of the feasibility of the study provided information on how to improve the study protocol for future studies.

Furthermore, it is not possible to know if the children chose to answer the questionnaires before the treatment or after. On the questionnaires, the child or parent wrote the date of treatment they were in; however, the date sometimes differed from the date in their treatment plan. Staff had on occasion missed handing out the questionnaires on the correct date. In the future, a more precise collection is warranted, for example, having the questionnaires on the web would be favorable [45], and reminders could increase response rates [46].

Implications for Future Studies

Some lessons learned through the feasibility study are the following: make sure that children in the study population under investigation have the possibility to take part as planned. Time needs to be set off for answering questionnaires within the hospital procedure schedule. When studying children with severe health conditions, it is important to be mindful of how many questions is possible for them to answer before they run out of energy. To test if games do what they are intended for within medicine and indicate if they would be used, it would be preferable to measure both player satisfaction and knowledge transfer.

Conclusions

All feasibility criteria set for the study were not met, suggesting that adaptions need to be made if a future study is going to be undertaken. It is preferable to do a feasibility study since it is a way to detect the advantages and disadvantages of the study protocol and to optimize future studies. There was no indication found that playing the serious game increased the children's self-reported anxiety toward undergoing radiotherapy. The PENS questionnaire showed promising results regarding player satisfaction in the use of the serious game within health care. When serious games are used as interventions, it is necessary to evaluate player satisfaction in future studies, and for that reason, the PENS questionnaire for children needs to be validated. To avoid dropout from the study, it would be commendable to not have as many as 5 measurement occasions for a group of the current age.

Acknowledgments

The authors would like to thank the children and parents who took the time to participate in the study and the personnel at the clinic for helping with the administration of the study. The study was funded by the Swedish Childhood Cancer Fund (grant KP2020-0010).

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author or Karin Enskär (karin.enskar@uu.se) upon reasonable request.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH (version 1.6.1) checklist.
[PDF File, 1225 KB - games_v12i1e54082_app1.pdf]

References

1. Kazak AE, Noll RB. The integration of psychology in pediatric oncology research and practice: collaboration to improve care and outcomes for children and families. Am Psychol 2015 Feb;70(2):146-158. [doi: 10.1037/a0035695] [Medline: 25730721]
2. Klosky JL, Tyc VL, Tong X, et al. Predicting pediatric distress during radiation therapy procedures: the role of medical, psychosocial, and demographic factors. Pediatrics 2007 May;119(5):e1159-e1166. [doi: 10.1542/peds.2005-1514] [Medline: 17452492]
3. Verma V, Beethe AB, LeRiger M, Kulkarni RR, Zhang M, Lin C. Anesthesia complications of pediatric radiation therapy. Pract Radiat Oncol 2016 May;6(3):143-154. [doi: 10.1016/j.prro.2015.10.018] [Medline: 26725960]
4. Boterberg T, Dieckmann K, Gaze M, editors. Radiotherapy and the Cancers of Children, Teenagers, and Young Adults: Oxford University Press; 2021.
5. Boterberg T, Dunlea C, Harrabi S, et al. Contemporary paediatric radiation oncology. Arch Dis Child 2023 May;108(5):332-337. [doi: 10.1136/archdischild-2021-323059]
6. Bucholtz JD. Comforting children during radiotherapy. Oncol Nurs Forum 1994 Jul;21(6):987-994. [Medline: 7971421]
7. Woodman H. Put yourself in their shoes… the vulnerability of children and their families when attending for radiotherapy treatment: the role of the specialist paediatric radiotherapy radiographer. J Radiography 2013 Nov;19(4):311-314. [doi: 10.1016/j.radi.2013.08.006]
8. Engvall G, Ångström-Brännström C, Mullaney T, et al. It is tough and tiring but it works—children's experiences of undergoing radiotherapy. PLoS One 2016 Apr;11(4):e0153029. [doi: 10.1371/journal.pone.0153029] [Medline: 27055258]
9. Goldman N, Gallitto M, Wu CC. Managing anxiety and minimizing sedation requirements in the pediatric radiation oncology population. Appl Rad Oncol 2021 Mar;10(1):14-17. [doi: 10.37549/ARO1265]

10. Gårdling J, Törnqvist E, Månsson ME, Hallström IK. Age-appropriate preparations for children with cancer undergoing radiotherapy: a feasibility study. J Child Health Care 2017 Dec;21(4):370-380. [doi: 10.1177/1367493517727070] [Medline: 29110518]
11. Engvall G, Lindh V, Mullaney T, Nyholm T, Lindh J, Ångström-Brännström C. Children's experiences and responses towards an intervention for psychological preparation for radiotherapy. Radiat Oncol 2018 Jan 22;13(1):9. [doi: 10.1186/s13014-017-0942-5] [Medline: 29357940]
12. Clerici CA, Ferrari A, Pecori E, et al. Age-appropriate multidisciplinary approach to young children with cancer undergoing radiotherapy: the SIESTA procedure. Pediatr Blood Cancer 2021 Jan;68(1):e28650. [doi: 10.1002/pbc.28650] [Medline: 32779871]
13. Baliga S, Yock TI. Proton beam therapy in pediatric oncology. Curr Opin Pediatr 2019 Feb;31(1):28-34. [doi: 10.1097/MOP.0000000000000724] [Medline: 30531222]
14. H-Yazdi N, Meadows R. Child proton beam therapy: a qualitative study of parental views on treatment and information sources. Radiography (Lond) 2021 Feb;27(1):101-107. [doi: 10.1016/j.radi.2020.06.016]
15. Gaito S, Aznar MC, Burnet NG, et al. Assessing equity of access to proton beam therapy: a literature review. Clin Oncol (R Coll Radiol) 2023 Sep;35(9):e528-e536. [doi: 10.1016/j.clon.2023.05.014] [Medline: 37296036]
16. Cockle SG, Ogden J. The radiation vacation: parents experiences of travelling to have their childrens brain tumours treated with proton beam therapy. Health Psychol Open 2016 Jan;3(1):2055102916649767. [doi: 10.1177/2055102916649767] [Medline: 28070403]
17. Linder LA, Hooke MC. Symptoms in children receiving treatment for cancer-part II: pain, sadness, and symptom clusters. J Pediatr Oncol Nurs 2019 Jul;36(4):262-279. [doi: 10.1177/1043454219849578] [Medline: 31307323]
18. Bray L, Appleton V, Sharpe A. The information needs of children having clinical procedures in hospital: will it hurt? Will I feel scared? What can I do to stay calm? Child Care Health Dev 2019 Sep;45(5):737-743. [doi: 10.1111/cch.12692] [Medline: 31163093]
19. Whitton N. Digital games and learning. In: Research and Theory: Routledge; 2014. URL: https://www.taylorfrancis.com/books/9781136216442 [doi: 10.4324/9780203095935]
20. Baranowski T, Blumberg F, Buday R, et al. Games for health for children—current status and needed research. Games Health J 2016 Feb;5(1):1-12. [doi: 10.1089/g4h.2015.0026]
21. Ingadottir B, Laitonen E, Stefansdottir A, et al. Developing a health game to prepare preschool children for anesthesia: formative study using a child-centered approach. JMIR Serious Games 2022 Jan 20;10(1):e31471. [doi: 10.2196/31471] [Medline: 35049507]
22. Aufegger L, Bùi KH, Bicknell C, Darzi A. Designing a paediatric hospital information tool with children, parents, and healthcare staff: a UX study. BMC Pediatr 2020 Oct 8;20(1):469. [doi: 10.1186/s12887-020-02361-w] [Medline: 33032549]
23. Cederved C, Back J, Ångström-Brännström C, Ljungman G, Engvall G. Co-creation of a serious game about radiotherapy: participatory action research study with children treated for cancer. JMIR Hum Factors 2022 May;9(2):e34476. [doi: 10.2196/34476]
24. Jibb LA, Croal L, Wang J, et al. Children's experiences of cancer care: a systematic review and thematic synthesis of qualitative studies. Oncol Nurs Forum 2018 Jul 2;45(4):527-544. [doi: 10.1188/18.ONF.527-544] [Medline: 29947357]
25. Bray L, Sharpe A, Gichuru P, Fortune PM, Blake L, Appleton V. The acceptability and impact of the Xploro digital therapeutic platform to inform and prepare children for planned procedures in a hospital: before and after evaluation study. J Med Internet Res 2020 Aug;22(8):e17367. [doi: 10.2196/17367]
26. Majid ESA, Garcia JA, Nordin AI, Raffe WL. Staying motivated during difficult times: a snapshot of serious games for paediatric cancer patients. IEEE Trans Games 2020 Nov;12(4):367-375. [doi: 10.1109/TG.2020.3039974]
27. Hunicke R, LeBlanc M, Zubek R. MDA: a formal approach to game design and game research. Presented at: AAAI Workshop on Challenges in Game AI; Jul 25-29, 2004; San Jose, CA URL: https://cdn.aaai.org/Workshops/2004/WS-04-04/WS04-04-001.pdf [accessed 2024-09-13]
28. Dai Y, Livesley J. A mixed-method systematic review of the effectiveness and acceptability of preoperative psychological preparation programmes to reduce paediatric preoperative anxiety in elective surgery. J Adv Nurs 2018 May 13;74(9):2022-2037. [doi: 10.1111/jan.13713] [Medline: 29754399]
29. Carleton RN. Fear of the unknown: one fear to rule them all? J Anxiety Disord 2016 Jun;41:5-21. [doi: 10.1016/j.janxdis.2016.03.011] [Medline: 27067453]
30. Lazarus RS. Emotions and interpersonal relationships: toward a person-centered conceptualization of emotions and coping. J Pers 2006 Feb;74(1):9-46. [doi: 10.1111/j.1467-6494.2005.00368.x] [Medline: 16451225]
31. Damsma Bakker A, van Leeuwen R, Roodbol P. Ethical considerations regarding the inclusion of children in nursing research. Nurs Ethics 2021 Feb;28(1):106-117. [doi: 10.1177/0969733020948120] [Medline: 32996358]
32. Spielberger CD, Edwards CD, Montouri J, Lushene R. State-Trait Anxiety Inventory for Children: Consulting Psychologists Press; 1973.
33. Johnson D, Gardner MJ, Perry R. Validation of two game experience scales: the Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). Int J Hum Comput Stud 2018 Oct;118:38-46. [doi: 10.1016/j.ijhcs.2018.05.003]

34. Spielberger CD, editor. Manual for the State-Trait Anxiety, Inventory: Consulting Psychologists Press; 1970.
35. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010 Mar 23;340:c332. [doi: 10.1136/bmj.c332] [Medline: 20332509]
36. Mahakwe G, Johnson E, Karlsson K, Nilsson S. A systematic review of self-report instruments for the measurement of anxiety in hospitalized children with cancer. Int J Environ Res Public Health 2021 Feb 16;18(4):1911. [doi: 10.3390/ijerph18041911] [Medline: 33669455]
37. Hoemann K, Feldman Barrett L. Concepts dissolve artificial boundaries in the study of emotion and cognition, uniting body, brain, and mind. Cogn Emot 2019 Jan 2;33(1):67-76. [doi: 10.1080/02699931.2018.1535428]
38. Cheng L, Wang L, He M, Feng S, Zhu Y, Rodgers C. Perspectives of children, family caregivers, and health professionals about pediatric oncology symptoms: a systematic review. Support Care Cancer 2018 Sep;26(9):2957-2971. [doi: 10.1007/s00520-018-4257-3] [Medline: 29774474]
39. Nilsson S, Holstensson J, Johansson C, Thunberg G. Children's perceptions of pictures intended to measure anxiety during hospitalization. J Pediatr Nurs 2019 Jan;44:63-73. [doi: 10.1016/j.pedn.2018.10.015] [Medline: 30683283]
40. McMullen KP, Hanson T, Bratton J, Johnstone PAS. Parameters of anesthesia/sedation in children receiving radiotherapy. Radiat Oncol 2015 Mar 11;10:65. [doi: 10.1186/s13014-015-0363-2] [Medline: 25889312]
41. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
42. Holtz BE, Murray K, Park T. Serious games for children with chronic diseases: a systematic review. Games Health J 2018 Oct;7(5):291-301. [doi: 10.1089/g4h.2018.0024] [Medline: 29957082]
43. Maddison R, Baghaei N, Calder A, et al. Feasibility of using games to improve healthy lifestyle knowledge in youth aged 9-16 years at risk for type 2 diabetes: pilot randomized controlled trial. JMIR Form Res 2022 Jun 17;6(6):e33089. [doi: 10.2196/33089] [Medline: 35713955]
44. Bonett DG. Sample size requirements for comparing two alpha coefficients. Appl Psychol Meas 2003 Jan;27(1):72-74. [doi: 10.1177/0146621602239477]
45. Evans JR, Mathur A. The value of online surveys: a look back and a look ahead. Internet Res 2018 Aug 6;28(4):854-887. [doi: 10.1108/IntR-03-2018-0089]
46. Van Mol C. Improving web survey efficiency: the impact of an extra reminder and reminder content on web survey response. Int J Soc Res Methodol 2017 Jul 4;20(4):317-327. [doi: 10.1080/13645579.2016.1185255]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
PENS: Player Experience of Need Satisfaction
STAI: State-Trait Anxiety Inventory
STAIC: State-Trait Anxiety Inventory for Children
STAIC(S): State-Trait Anxiety Inventory for Children-State Anxiety

Edited by A Mavragani; submitted 29.10.23; peer-reviewed by A Ghosh, S Liszio; revised version received 28.06.24; accepted 02.07.24; published 23.09.24.

Please cite as:
Cederved C, Ljungman G, Back J, Ångström-Brännström C, Engvall G
Acceptability of a Serious Game About Proton Radiotherapy Designed for Children Aged 5 to 14 Years and Its Potential Impact on Perceived Anxiety: Feasibility and Randomized Controlled Pilot Trial
JMIR Serious Games 2024;12:e54082
URL: https://games.jmir.org/2024/1/e54082
doi:10.2196/54082

Using Virtual Reality in a Rehabilitation Program for Patients With Breast Cancer: Phenomenological Study

Shih-Chung Wu[1], MD; Chia-Wen Chuang[2,3], MSN; Wen-Chun Liao[4,5], PhD; Chung-Fang Li[2], MSN; Hsin-Hsin Shih[4,5], PhD

1

2

3

4

5

Corresponding Author:
Wen-Chun Liao, PhD

Abstract

Background: Surgery is an essential treatment for early-stage breast cancer. However, various side effects of breast cancer surgery, such as arm dysfunction and lymphedema, remain causes for concern. Rehabilitation exercises to prevent such side effects should be initiated within 24 hours after surgery. Virtual reality (VR) can assist the process of rehabilitation; however, the feasibility of applying VR for rehabilitation must be explored, in addition to experiences of this application.

Objective: This study explored patients' attitudes toward and experiences of using VR for their rehabilitation to determine the feasibility of such VR use and to identify potential barriers.

Methods: A phenomenological qualitative study was conducted from September to December 2021. A total of 18 patients with breast cancer who had undergone surgical treatment were interviewed using open-ended questions. The Colaizzi 7-step procedure for phenomenological analysis was used for data analysis. To ensure high study reliability, this study followed previously reported quality criteria for trustworthiness.

Results: Three themes were identified: (1) VR was powerful in facilitating rehabilitation, (2) early and repetitive upper limb movements were an advantage of VR rehabilitation, and (3) extensive VR use had challenges to be overcome. Most of the interviewed patients reported positive experiences of using VR for rehabilitation. Specifically, VR helped these patients identify appropriate motion and angle limits while exercising; in other words, knowledge gained through VR can play a key role in the rehabilitation process. In addition, the patients reported that the use of VR provided them company, similar to when a physiotherapist is present. Finally, the gamified nature of the VR system seemed to make VR-based rehabilitation more engaging than traditional rehabilitation, particularly with respect to early rehabilitation; however, the high cost of VR equipment made VR-based rehabilitation difficult to implement at home.

Conclusions: The interviewed patients with breast cancer had positive experiences in using VR for rehabilitation. The high cost of both VR equipment and software development presents a challenge for applying VR-based rehabilitation.

(JMIR Serious Games 2024;12:e44025) doi:10.2196/44025

KEYWORDS

breast cancer; rehabilitation; virtual reality; VR; virtual reality design process; VR design process; feasibility; accessibility

Introduction

Breast cancer is a major global health problem. In Taiwan, more than 10,000 women are diagnosed with breast cancer every year; approximately 80% of these diagnoses are early-stage breast cancer and most require surgical treatment [1]. More than 1 in 5 women who are breast cancer survivors might eventually develop upper limb lymphedema [2]. In addition, the life expectancy of patients with stage 0 to stage 3 breast cancer is 20 to 32 years [3]. Breast cancer treatments include surgery, radiation therapy, chemotherapy, endocrine therapy, and targeted therapy. Because of advancements in diagnosis and treatment in recent years, the 5-year relative survival rate of breast cancer is now higher than 80% [4]. However, both ancillary dissection and axillary radiation are known to increase the incidence of lymphedema and axillary web syndrome [5]; associated symptoms include pain, upper extremity weakness, paresthesia, and limited range of motion (ROM), each of which can not only delay the commencement of radiation therapy but also lead to patients being unable to perform basic self-care tasks. This is a problem because less-mobile patients are more likely to experience side effects after surgery; for example, a frozen shoulder is a side effect that commonly occurs in the short term as a result of immobility after surgery [6]. These associated symptoms may persist for up to 1.5 years after surgery and therefore predispose patients to depression and other mental health conditions that can become long-term health problems [7,8].

Studies have suggested that the early commencement of postoperative rehabilitation exercises, including upper extremity abduction and flexion exercises, can reduce the risk of upper extremity pain and dysfunction [9,10]. Although the optimal time to start ROM exercises remains unclear, one study suggested that activity recovery tends to be more successful following earlier rehabilitation [11]. Typically, rehabilitation is recommended to be commenced within the first 7 days after surgery, with shoulder ROM initially being limited to 90° [12]. Hence, a patient's willingness to undergo rehabilitation after surgery to prevent their arms from becoming motionless is essential. However, after surgery, a patient's motivation to undergo physical rehabilitation tends to be low because of wound pain and fear of movement [13,14].

Virtual reality (VR) can reduce fear of movement and improve motivation to engage in rehabilitation. In addition, VR can promote compliance and rehabilitation success among patients [15-17]. Furthermore, VR appears to be effective in increasing shoulder ROM compared with standard physiotherapy for postoperative rehabilitation in patients with breast cancer [18]. The use of VR for rehabilitation is not new; however, feasibility studies that use VR to provide individualized, progressive practice for arm movements after surgery for breast cancer remain scarce. Feasibility research is often conducted prior to a randomized controlled trial [19]. In recent years, most studies investigating VR for rehabilitation in patients with breast cancer have used a randomized controlled design and have been prospective, adequately powered, and methodologically rigorous. Nevertheless, according to According to the recommendations of the VR-CORE (Clinical Outcomes Research Experts) [20], the development of new VR content should include 3 phases: VR1 to VR3. VR1 provides guidance for the development of new VR content; VR2 constitutes early testing related to the feasibility, acceptability, tolerability, and initial clinical efficacy of VR; and VR3 evaluates the efficacy of VR in comparison with controls. Bypassing feasibility and accessibility tests to jump immediately to clinical trials may lead to a gap between program development and bedside application and cause barriers to implementing VR in practice. Even though there is sound evidence that VR is effective in improving rehabilitation effects in patients with breast cancer, VR is not currently used in many clinical settings. According to Brennan et al [21], barriers to the implementation of VR include usability problems, cognitive limitations, cost, a lack of patient safety, low patient motivation, and negative patient responses. In addition, VR's unfamiliarity to many patients may dampen their willingness to use VR [22]; thus, patients' attitudes, experiences, and opinions related to VR must be understood. Further explorations of feasibility and accessibility are recommended for the use of VR in health care [23,24].

To the best of our knowledge, most studies regarding postoperative rehabilitation in patients with breast cancer have been clinical trials (ie, VR3) that evaluated only treatment effects, such as pain, grip power, muscle power, or ROM; VR2 studies are skipped. This study used a phenomenological approach to investigate the feasibility and experiences of a VR rehabilitation program (ie, this was a VR2 study) in patients with breast cancer. Therefore, this study was conducted to answer the following questions: (1) What are the lived experiences of patients using VR for arm rehabilitation? (2) How do they perceive the process? In answering these questions, this study determined the clinical effectiveness of and barriers to VR rehabilitation as a prelude to a definitive randomized controlled trial.

Methods

Design and Development of the VR System

The VR system in this study was designed and developed in accordance with the VR–CORE–VR1 guidelines [20]. Our VR design team was a multidisciplinary team comprising users and technicians. The rehabilitation program was developed in 3 stages. The first stage was based on inspiration gained through empathy. Specifically, interviews were held with 2 patients who were invited to share their rehabilitation experiences and express their perspectives and opinions to facilitate the design of an effective VR-based treatment program. The interview questions in this stage considered (1) smart device use behavior and (2) willingness and acceptability to use VR rehabilitation systems. The second stage involved ideation through team collaboration. Specifically, our team comprised experts in surgery, rehabilitative physicians, and software engineering, as well as a case manager. Innovative ideas were generated through methods such as brainstorming and collective ideation, and the ideas that were most suitable for prototyping were then compiled. The third stage involved prototyping based on user feedback. Specifically, further discussions were held with the aforementioned 2 patients to obtain further information. Our team then built prototypes to test 3 VR rehabilitation exercises prior to implementation (Table 1). Ultimately, we extracted 3 key sets of limb motions that are described as follows. The first motion was a Whac-A-Mole–like game that involved the abduction of the shoulder (no more than 90°). The second motion was wiping a table and involved external forearm rotation. The third motion was combing one's hair and involved flexion of the shoulder (no more than 90°) (Figures 1-3). This study used the PC-based Oculus Quest 2 head-mounted display (HMD) to provide the most immersive VR experience possible.

Table . Summary of our design principles and strategies based on recommendations for best practices in VR1 (virtual reality phase 1) studies [20].

Design principles and strategies	Our practices
Inspiration through empathy, recruitment, observation, patient interviews, and expert interviews	Two patients with breast cancer who had experienced axillary web syndrome were invited for individual interviews regarding their relevant needs, experiences, fears, and expectations; a group discussion was then conducted to determine the patients' needs.
Ideation through team collaboration, sharing stories and notes, and generating ideas	We analyzed, aggregated, and discussed the stories and data obtained in the previous phase. Our team then formulated 10 ideas for rehabilitation actions. After considering the restrictions on drainage tube indwelling the day after surgery (shoulder joint movement must not exceed 90°), we extracted 3 motions: (1) Whac-A-Mole, which involved abduction of the shoulder; (2) wiping a table, which involved external forearm rotation; and (3) combing one's hair, which involved flexion of the shoulder.
Prototyping through continuous feedback, building a prototype, and repeatedly testing the prototype	We drew the height of a mallet on a wall and placed a picture of a gopher on a table to simulate the Whac-A-Mole game. For the wiping the table and combing hair scenarios, we used a table and a comb, respectively, and we then collected feedback from the patients after their first use. Our team found that the Whac-A-Mole game was difficult to perform owing to the different heights of the patients and seats (eg, a patient may need to raise their shoulder angle excessively to pick up the mallet, leading to pain and an inability to perform the exercise, forcing it to be abandoned). Thus, the software engineer adjusted the height of the controllable mallet and the adjustable gopher table to meet the need of each individual patient.

Figure 1. Whac-A-Mole–like scenario.

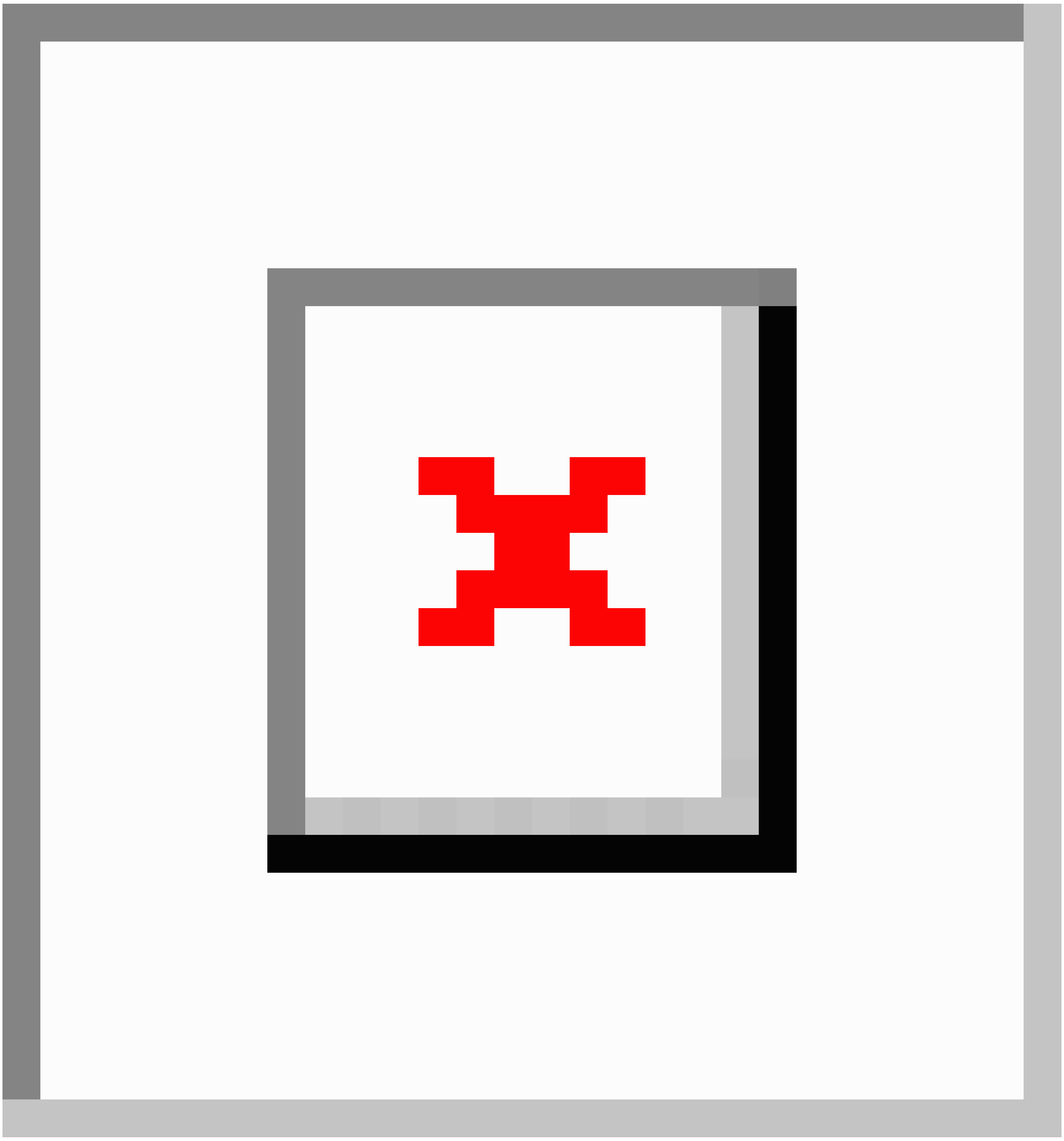


Figure 2. Wiping a table scenario.

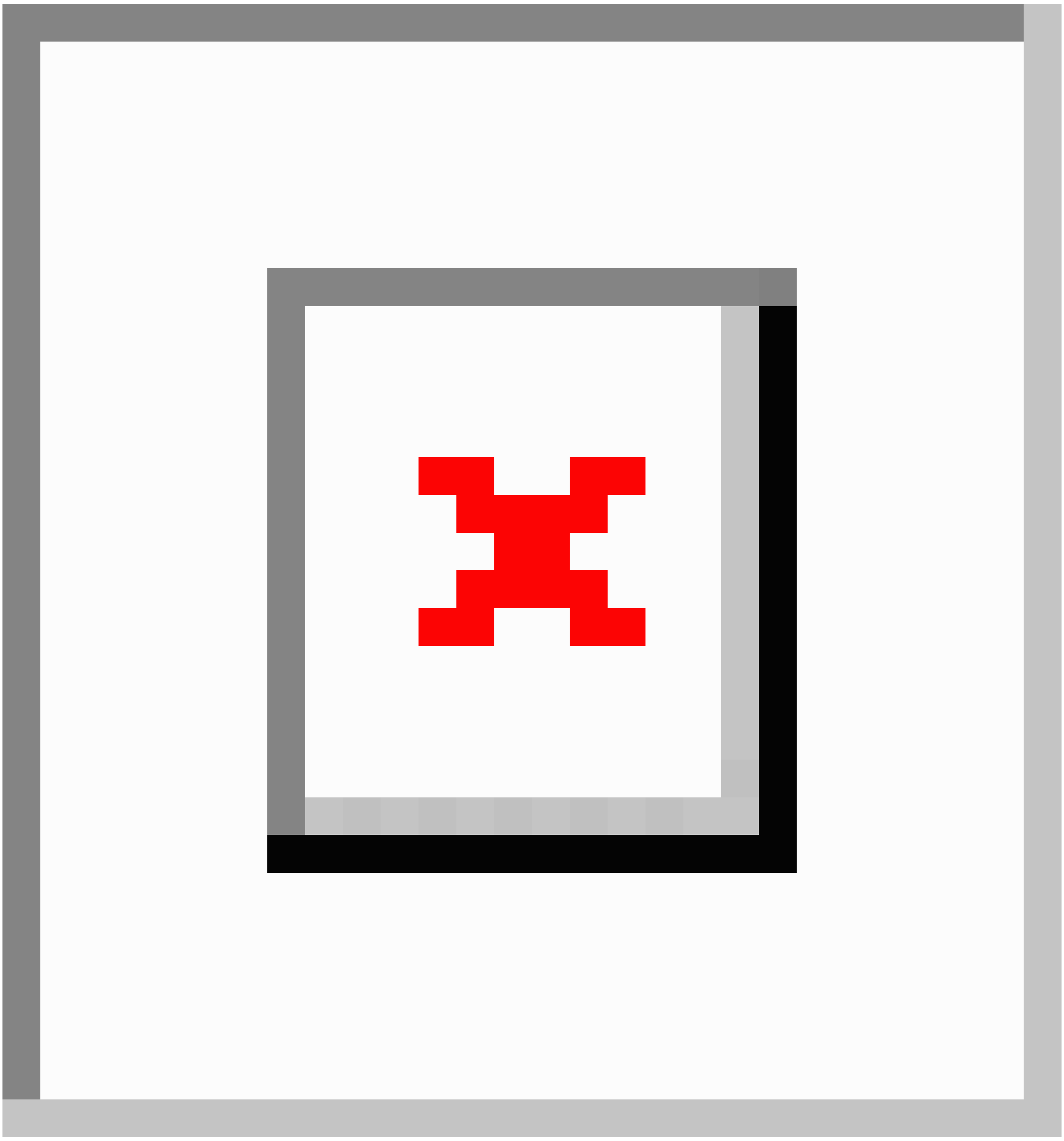
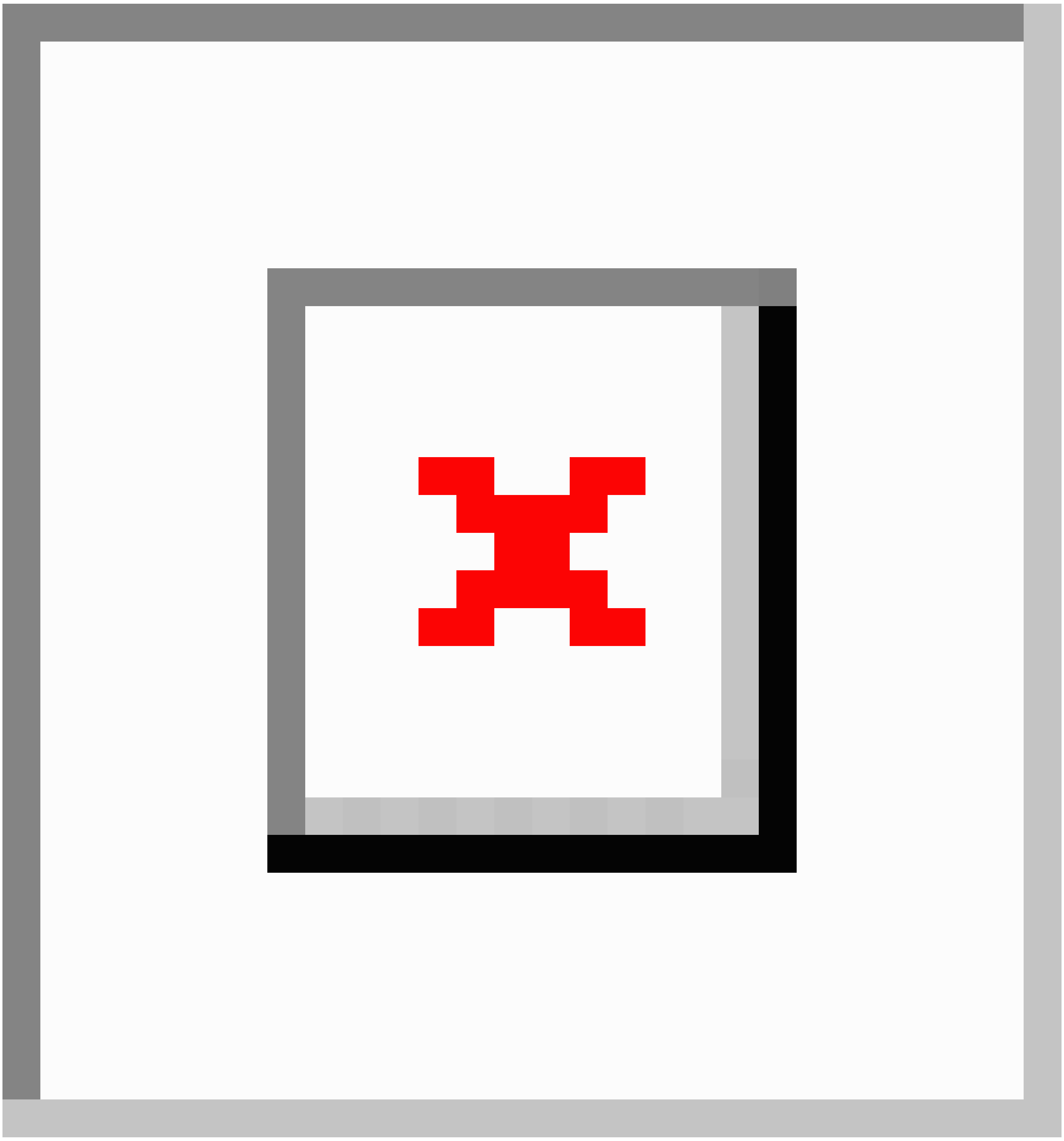


Figure 3. Combing one's hair scenario.



Study Design

This qualitative study was guided by the phenomenological methodology; it focused on investigating the perceived experience of using VR rehabilitation and then used an inductive approach and the Colaizzi [25] interview data analysis method to identify themes from the VR rehabilitation experiences.

Ethics Approval

Ethical approval for this study was granted by the Human Trials Ethics Committee of Chang Gung Memorial Hospital, Taiwan (202001617A3C501). All the participants provided written informed consent to participate in this study.

Setting

This study was conducted at a medical center in Taiwan that receives up to 20 new patients with breast cancer per month.

Participants and Recruitment

The criteria for participation in this study were as follows: (1) a new diagnosis of breast cancer and being hospitalized for surgical treatment; (2) being aged 20 to 65 years; (3) being able to read and speak Mandarin Chinese or Taiwanese; (4) being willing to use a wearable VR HMD and having sufficient cognitive ability to understand and follow instructions, as well as to interact with the VR environment; (5) having sufficient physical ability (self-reported) to use the VR equipment in the VR rehabilitation program; and (6) being willing to be interviewed. After surgical treatment for breast cancer, patients were invited to a rehabilitation consultation to ask about their willingness to try VR rehabilitation and were recruited to participate in this study. The recruitment period ran from September to December 2021, with 18 patients interviewed.

Data Collection

First, the participants filled out a questionnaires on demographic data (including age, gender, and type of breast surgery), the Distress Thermometer (DT) questionnaire, and the Chinese Health Questionnaire–12 (CHQ-12). The DT is a visual analog scale for routine screening for distress in all patients with cancer; scores range from 0 (indicating no distress) to 10 (indicating extreme distress). A cutoff score of ≥7 is optimal for distress in patients with a new diagnosis [26]. The CHQ-12 was modified from the General Health Questionnaire–12, a widely used psychometric test for screening mental problems, with a score of ≥4 indicating psychological distress [27].

After completing the questionnaires, a face-to-face interview was conducted with each participant for approximately 30 minutes. Interview guides were developed by the research team, which included an expert in qualitative studies and a clinical nurse specialist in breast cancer. After the content of the first interview was analyzed, the interview questions were discussed and modified with the research team.

During the interview process, the researchers actively listened and continuously clarified any perspectives and meanings that the participants wanted to express. Participants were allowed to freely express their views. All interviews were audiorecorded. After the data analysis, the researchers reviewed the data with the patients to ensure data accuracy and completeness.

Inductive Content Analysis

According to Kyngäs [28], an inductive approach, which is common in qualitative research, should be used when no prior research has addressed a specific phenomenon or if existing knowledge is fragmented. Because prior studies had not investigated the experiences of patients with breast cancer to evaluate the effectiveness of VR-based rehabilitation, our study used an inductive approach to content analysis, which involved data reduction, data grouping, and, finally, concept formation. The content was analyzed with the Colaizzi [25] phenomenological analysis method, which involves a 7-step procedure for content analysis. Specifically, 3 researchers analyzed the data and grouped them into themes to enhance comprehension. First, the interview transcript was read to obtain an idea of its overall meaning. Second, a researcher extracted significant statements relevant to the purpose of the study. Third, each statement considered meaningful was recorded for accurate representation of the research data. Fourth, the researcher classified significant statements into categories based on their meaning. Fifth, the researcher grouped similar categories together and organized them into themes. Sixth, all themes presented were combined to produce an exhaustive description. Seventh, the results were presented to each participant for feedback and to verify their accuracy.

Trustworthiness

To enhance the trustworthiness of our study, we referred to the Guba and Lincoln [29] quality criteria of credibility, transferability, dependability, and confirmability.

To establish confidence in the credibility of the findings, triangulation was used to examine the consistency of the data. Researchers presented their summaries of the categories and themes from data analysis and discussed any uncertainty about any category or theme by replaying and reanalyzing the interview recording.

High transferability was ensured through the interview process, including the stability of observations and the interrater reliability. As an interviewer, the researcher was attentive to the interviewees' experiences and avoided bias arising from their values and beliefs. The interviewer was conscious about avoiding misperceptions about what the interviewees were saying to ensure the stability of observations. There was only 1 interviewer, hence interrater reliability was not measured.

The study had high dependability, as the data collecting procedures were approved by an expert in qualitative methodology, and the themes and subthemes that emerged were consistent among all researchers after the coding and recoding procedure during the data analysis.

Finally, during the data analysis stage, 3 researchers who were not part of the study team confirmed the collected data by distinguishing between the units of the main themes and subthemes to verify the main themes' theoretical saturation and thus achieved confirmability.

Results

Characteristics of Participants

Table 2 shows the participants' characteristics. The mean age of the participants was 46.66 (SD 8.05) years; 17 participants were female (95%) and 1 was male (5%); 11 were married (61%); 11 (61%) were full-time workers; and 7 (39%) were housekeepers or had no formal employment. Regarding the types of surgery received, 9 participants (50%) underwent mastectomy and sentinel lymph node biopsy, 4 (23%) underwent mastectomy and axillary lymph node dissection, 2 (11%) underwent mastectomy and reconstruction, and 3 (16%) underwent modified radical mastectomy. All participants had DT scores less than 7, indicating no distress. In the CHQ-12 items, some participants reported having shaking or numbness of limbs (n=3, 17%), losing a great deal of sleep because of worry (n=4, 22%), and losing confidence in themselves (n=4, 22%) (items 4, 5, and 8). However, all participants had CHQ-12 scores less than 4, indicating no mental problems. Finally, 7 (39%) and 11 (61%) of the participants reported moderate (4-6) and mild (1-3) wound pain on the Numeric Rating Scale for Pain, respectively.

Table . Distribution of sociodemographic variables of participants (N=18).

Variables		Participants, n (%)
Sex		
	Female	17 (95)
	Male	1 (5)
Age group (years)		
	35-44	6 (33)
	45-54	8 (44)
	55-64	4 (3)
Educational level		
	Senior high school	9 (50)
	College	9 (50)
Distress		
	Yes (DT[a] score ≥7)	0 (0)
	No (DT score <7)	10 (100)
Mental problems		
	Yes (CHQ-12[b] score ≥4)	0 (0)
	No (CHQ-12 score <4)	18 (100)
Numerical Rating Scale for Pain score		
	1-3	11 (61)
	4-6	7 (39)
Marital status		
	Married or in a domestic partnership	11 (61)
	Single	5 (28)
	Separated or divorced	2 (11)
Employment status		
	Employed	11 (61)
	Unable to work	7 (39)
Type of breast surgery		
	Mastectomy with sentinel lymph node biopsy	9 (50)
	Mastectomy with axillary lymph node dissection	4 (23)
	Mastectomy with reconstruction	2 (11)
	Modified radical mastectomy	3 (16)

[a]DT: Distress Thermometer.

[b]CHQ-12: Chinese Health Questionnaire–12.

Themes

The following 3 themes were inductively extracted from the interview data: (1) VR was powerful in facilitating rehabilitation, (2) early and repetitive upper limb movements were an advantage of VR rehabilitation, and (3) extensive VR use had challenges to be overcome. Tables 3-5 present these themes in greater detail alongside their subthemes and interviewee statements.

Table . Subthemes and interviewee quotes related to theme 1: virtual reality (VR) was powerful in facilitating rehabilitation.

Subthemes	Quotes
Obtaining knowledge	• "It turns out that some actions are unsuitable [because of the drainage tube]; after use, I have a better idea of the angle of movement" (participant 2). • "After practicing with this system [VR] and being discharged from the hospital, I know that some actions are not suitable for me to do, so I will not do them randomly" (participant 3). • "I was anxious, thinking that not moving my arms was suitable for wound healing" (participant 13). • "I can adjust my range of motion according to my tolerance, and I will not feel frustrated" (participant 17).
Companionship	• "This way of rehabilitation [VR] is like a kind of support for me when exercising because thephysiotherapist does not always accompany me" (participant 9). • "The usage process is very simple. I think it is well executed in the hospital and I can concentrate on practicing" (participant 5). • "The physiotherapist was very busy and left after teaching me the rehabilitation movement, so I had to do it independently" (participant 11). • "I didn't understand what this was at first, and I was a little scared, but it turned out that I was doing rehabilitation while playing games" (participant 1).

Table . Subthemes and interviewee quotes related to theme 2: early and repetitive upper limb movements are an advantage of virtual reality (VR) rehabilitation.

Subthemes	Quotes
Reduced fear of movement	• "The VR rehab game method is more fun than the traditional rehab method because it lets me know I need to do rehab exercises; I think I can do it a few more times" (participant 1). • "My time in the hospital was short, and it was a pity that I didn't do VR many times" (participant 6). • "I experienced no pain or discomfort during the VR rehab; I want to do it a few more times" (participant 7). • "After playing Whac-A-Mole, I attempted to move my arm" (participant 11). • "While using it, I temporarily forgot that I had surgery" (participant 14). • "Initially, I was terrified of moving my arms, but now, I'm having so much fun that I don't realize I'm doing rehab" (participant 18).
Enhanced motivation to engage in rehabilitation	• "The VR program can be performed as soon as possible; I have been performing VR for seven days since surgery. In the execution of the Whac-A-Mole and wiping the table exercises, I did not have too much discomfort or pain, and I am doing better and better" (participant 15). • "Although there is no VR at home, I can continue to do the VR movements, and I will not worry about my arm's limited range of motion in the future" (participant 16). • "After using VR for rehabilitation, I found that my condition was not as bad as I thought" (participant 7).

Table . Subthemes and interviewee quotes related to theme 3: extensive virtual reality (VR) use has challenges to be overcome.

Subthemes	Quotes
Safety of the real-world environment	• "I was initially scared because I had never used VR" (participant 4). • "Wearing a VR HMD, I can't see the outside world, so I'm worried about hitting something" (participant 5).
Cost of virtual reality rehabilitation equipment is too high to extend to home use	• "VR rehabilitation equipment is too expensive to buy and can only be used in hospitals" (participant 5). • "Without equipment, there is no way to use VR for rehabilitation" (participant 8). • "VR rehabilitation can only be used in hospitals; at home, we use traditional rehabilitation methods" (participant 17). • "I hope to rent VR equipment for my home so I can use it anytime" (participant 18). • "Although there is the Switch, finding a suitable sports game for rehabilitation is difficult" (participant 9).
Motion sickness	• "After wiping the table, I felt a little dizzy" (participant 10). • "My eyes constantly roll when wiping the table, which is a little uncomfortable" (participant 9).

VR Was Powerful in Facilitating Rehabilitation

Overview

A key concept of this inductive theme was the "power" obtained from knowledge and companionship. Related subthemes and interviewee statements are presented in Table 3.

Obtaining Knowledge

A total of 12 (70%) of the participants described how the VR content helped them learn the most helpful rehabilitation movements and the angular limitations of their exercises and thus enabled them to perform these exercises safely and relatively early in the rehabilitation process.

Companionship

Most of the participants stated that playing games enabled them to feel as though they were interacting with health care professionals and experiencing companionship and support.

Early and Repetitive Upper Limb Movements: An Advantage of VR Rehabilitation

Overview

After surgical treatment for breast cancer, rehabilitative exercises must be performed early. In this study, wound pain and discomfort during such exercises were the main concerns of the participants, leading to fear of movement. Related subthemes and interviewee statements are presented in Table 4.

Reduced Fear of Movement

A total of 14 (80%) of the participants stated that the gamified design of VR rehabilitation made it more interesting than traditional rehabilitation, while 7 of the patients experienced moderate pain after surgery. However, during the use of VR for rehabilitation, these patients reported no wound pain or discomfort. In addition, their willingness to engage in rehabilitation was enhanced by the second day after surgery.

Enhanced Motivation to Engage in Rehabilitation

Most of the participants described an immersive virtual environment that facilitated the performance of early rehabilitative exercises. In addition, the participants were subsequently able to continue performing the aforementioned 3 rehabilitation movements at home without VR.

Extensive VR Use Has Challenges to Be Overcome

Overview

Most of the patients had safety concerns, mainly because while wearing the VR HMD, they were unable to pay attention to the environment around them. Consequently, some of the patients initially felt apprehensive, fearing collisions with nearby objects. In addition, high equipment costs limited the accessibility of the equipment for home use, and many of the patients expressed disappointment that they were able to use the system only in hospital. Related subthemes and interviewee statements are presented in Table 5.

Safety of the Real-World Environment

While the patients were wearing the VR HMD and engaging in VR-based exercise, they were unable to pay attention to the environment around them. Thus, patients had the not-unfounded fear that they might collide with nearby objects and were initially apprehensive. Therefore, patients should be in a clutter-free environment when engaging in VR rehabilitation.

Cost of VR Rehabilitation Equipment Is Too High to Extend to Home Use

The cost of VR equipment for rehabilitation is high, making it relatively inaccessible for home use. Each new technology faces the problem of high equipment cost in its initial stage, and similarly, the main problem for most of the participants in this study was that they wanted to continue using the rehabilitation equipment at home. They indicated that they thought it was a pity that the system could be used only in hospital.

Motion Sickness

We found that 3 of the participants reported the side effect of motion sickness after performing the wiping table exercise.

Discussion

Principal Findings

To the best of our knowledge, this study is the first phenomenological study to investigate the feasibility of an immersive VR experience for facilitating postoperative rehabilitation in patients with breast cancer. Through inductive thematic analysis, our study found VR to be an effective and suitable intervention for our patient population. Tables 3 and 4 address feasibility in relation to 2 of the 3 inductive themes of this study: "VR was powerful in facilitating rehabilitation" and "early and repetitive upper limb movements are an advantage of VR rehabilitation." Table 5 reveals the key finding that the high costs of VR hardware and software limit its at-home use. These results further support the recommendations of the VR CORE group [20]. However, the high cost of VR equipment remains a concern.

VR Was Powerful in Facilitating Rehabilitation

This study found that "power" was an essential concept and that power was cultivated through knowledge and companionship. One key finding related to this inductive theme was the knowledge obtained by the individual. In our study, most of the participants were skeptical about their early postoperative rehabilitation, and some had no idea how to proceed with it; however, through VR, these patients gained information about rehabilitation, which then led to behavioral changes. A previous study noted the importance of people's need to know which approach they should adopt to improve their health [30]. In addition, most patients with breast cancer desire information concerning recovery details; in particular, they want real-time, transparent information regarding appropriate rehabilitative measures [21]. Our findings were consistent with those of the 2 aforementioned studies. Another present finding was related to companionship; specifically, the participants reported that they felt as if a medical care team was with them in their rehabilitation in the VR world. This finding is notable because not every patient with breast cancer receives a rehabilitation exercise consultation after surgery. Our interactive VR system provides clear information regarding the commencement and completion of rehabilitation exercises, and its prompts align with how the clinical rehabilitation program operates (eg, 20 prompts for each motion). The literature notes that an ideal user experience must include interaction with VR applications in order to encourage and support users to complete clinical activities; this assertion is consistent with our findings [31].

Early and Repetitive Upper Limb Movements: Advantage of VR Rehabilitation

VR elicited engagement in early postoperative rehabilitation in most of this study's patients, seemingly stemming from their reduced fear of movement and their enhanced motivation. Early postoperative rehabilitation in patients who have undergone breast cancer surgery aims to improve shoulder mobility, arm mobility, or both. VR is believed to be as effective as conventional physiotherapy in improving upper limb function to facilitate the maintenance of activities of daily living [32]. Some of the participants in this study were initially apprehensive to undertake postoperative rehabilitation primarily because of fear related to insufficient knowledge of the appropriate degree of movement. Nevertheless, our VR rehabilitation system was based on the concept of gamification; the participants who used the system were attracted to the game screen, and consequently, their negative feelings—including fear, pain, and discomfort—were temporarily alleviated. In this manner, the VR system considerably enhanced the participants' motivation to engage in rehabilitation; this finding was consistent with those of previous studies [18,33]. In addition, the game design can be adjusted according to the needs of the patient to prevent patients from giving up on movements that are either difficult or so simple that the desired training outcomes cannot be achieved by performing them.

Challenges of Extensive VR Use to Be Overcome

The participants in this study were required to wear HMDs; thus, they could not see their surroundings during rehabilitation. Related studies that have recommended VR have stated that when using such technology, a space of at least 70 cm × 70 cm is required; however, in our study, the participants took a seated position to minimize the risk of tripping and thus needed only to pay attention to whether the seat back was affecting exercise safety. Although research has shown that the cost of HMDs has fallen in recent years [15], VR devices remain unaffordable for many people in Taiwan. Therefore, even if patients with breast cancer are interested in using this emerging technology for rehabilitation, doing so without an HMD is difficult. In addition, the development and maintenance costs of VR software and hardware are high, and thus, the cost of VR equipment remains a challenge for large-scale implementation in postoperative rehabilitation.

Motion sickness (MS) is a common physiological response to VR immersion [34]. Previous studies have shown that behavioral and dietary strategies and physical therapy, such as listening to music or chewing gum [35], are effective in alleviating MS symptoms. Although each movement in our VR rehabilitation system was accompanied by music with a brisk tempo, 3 of the patients in this study still experienced MS due to unsuitable seat height. A previous study demonstrated that positive emotions can relieve MS [36]; we found that most patients have positive emotions while using VR for rehabilitation. Accordingly, this paper suggests that the use of VR should be combined with emotional assessments to effectively prevent MS. With regard to feasibility, we found that the present VR system met the patients' needs, conformed to their values, and satisfied their expectations. The system also provided an engaging and interactive rehabilitation environment that stimulated the patients' motivation to engage in early and continuous rehabilitation and, in turn, reduced their risk of developing defensive functional impairments. Nevertheless, the accessibility of VR proved to be a challenge because of the difficulty of using VR technology at home. Thus, such interventions may tend to be limited to in-hospital use because not every patient can afford to purchase VR equipment and software.

Study Limitations

This study has some limitations. First, we did not conduct a pilot study of the interview guide. However, after the content of the first interview was analyzed, the interview questions were discussed and modified with the research team to identify and address any potential issues in the data collection methods. It may be recommended to incorporate a pilot study in future research to ensure validity and robustness. Second, the development of VR technology to facilitate postoperative rehabilitation in patients with breast cancer is still in its infancy. In this study, only 3 rehabilitation exercises were developed and clinically tested. The facilitators of and barriers to rehabilitation were investigated. To determine the clinical efficacy of the proposed VR system, a large-scale randomized controlled trial is needed.

Conclusion

Feasibility research related to health care is crucial for generating novel ideas and improvements to provide more effective care for patients. From the perspective of this study, rehabilitation is an iterative, active, and educational process of problem-solving that must focus on a patient's behavior. VR technology was effective in facilitating rehabilitation by providing knowledge and companionship to participants. It helped reduce fear of movement and enabled early and repetitive upper limb movements. However, safety concerns, high equipment costs, and potential side effects like MS were reported as challenges to extensive VR use.

Acknowledgments

We are grateful to Mei-Ling Li and Su-Fang Li in for their informative comments during the development of the virtual reality rehabilitation system. This research was partially supported by Kaohsiung Chang Gung Memorial Hospital, part of the Chang Gung Medical Foundation, Taiwan (grant CMRPG8L0081), for the development of the virtual reality system and partially supported by the China Medical University Project (grant CMU108-MF-21).

Authors' Contributions

SCW and CWC were responsible for the study design and coordination. CWC, HHS, and CFL extracted the data. SCW and WCL drafted the manuscript. All authors have contributed to, read, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Taiwan cancer registry. Health Promotion Administration. 2015. URL: https://www.hpa.gov.tw/EngPages/Detail.aspx?nodeid=1061&pid=6069 [accessed 2024-02-20]
2. DiSipio T, Rye S, Newman B, Hayes S. Incidence of unilateral arm lymphoedema after breast cancer: a systematic review and meta-analysis. Lancet Oncol 2013 May;14(6):500-515. [doi: 10.1016/S1470-2045(13)70076-7] [Medline: 23540561]
3. Lin CN, Lee KT, Chang SM, Wang JD. Cost-effectiveness evaluation of mammography screening program in Taiwan: adjusting different distributions of age and calendar year for real world data. J Formos Med Assoc 2022 Mar;121(3):633-642. [doi: 10.1016/j.jfma.2021.06.013] [Medline: 34246512]
4. Coleman MP, Quaresma M, Berrino F, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). Lancet Oncol 2008 Aug;9(8):730-756. [doi: 10.1016/S1470-2045(08)70179-7] [Medline: 18639491]
5. da Luz CM, Deitos J, Siqueira TC, Palú M, Heck APF. Management of axillary web syndrome after breast cancer: evidence-based practice. Rev Bras Ginecol Obstet 2017 Nov;39(11):632-639. [doi: 10.1055/s-0037-1604181] [Medline: 28701024]
6. Pedersen AB, Horváth-Puhó E, Ehrenstein V, Rørth M, Sørensen HT. Frozen shoulder and risk of cancer: a population-based cohort study. Br J Cancer 2017 Jun 27;117(1):144-147. [doi: 10.1038/bjc.2017.146] [Medline: 28524156]
7. Akezaki Y, Nakata E, Kikuuchi M, et al. Risk factors for early postoperative psychological problems in breast cancer patients after axillary lymph node dissection. Breast Cancer 2020 Mar;27(2):284-290. [doi: 10.1007/s12282-019-01020-y] [Medline: 31679113]
8. Arm/shoulder problems common after breast cancer surgery. Breastcancer.org. 2008 Sep 9. URL: https://www.breastcancer.org/research-news/20080909 [accessed 2024-01-31]
9. Scaffidi M, Vulpiani MC, Vetrano M, et al. Early rehabilitation reduces the onset of complications in the upper limb following breast cancer surgery. Eur J Phys Rehabil Med 2012 Dec;48(4):601-611. [Medline: 22510674]
10. De Groef A, Van Kampen M, Vervloesem N, et al. Myofascial techniques have no additional beneficial effects to a standard physical therapy programme for upper limb pain after breast cancer surgery: a randomized controlled trial. Clin Rehabil 2017 Dec;31(12):1625-1635. [doi: 10.1177/0269215517708605] [Medline: 28523988]
11. Bendz I, Olsén MF. Evaluation of immediate versus delayed shoulder exercises after breast cancer surgery including lymph node dissection–a randomised controlled trial. Breast 2002 Jun;11(3):241-248. [doi: 10.1054/brst.2001.0412] [Medline: 14965674]

12. Richmond H, Lait C, Srikesavan C, et al. Development of an exercise intervention for the prevention of musculoskeletal shoulder problems after breast cancer treatment: the prevention of shoulder problems trial (UK PROSPER). BMC Health Serv Res 2018 Jun 18;18(1):463. [doi: 10.1186/s12913-018-3280-x] [Medline: 29914494]
13. Malicka I, Stefańska M, Pawłowska K, Woźniewski M. Function of muscles of flexors and extensors of the elbow joint in women after treatment of breast cancer. Ortop Traumatol Rehabil 2009;11(2):111-119. [Medline: 19502668]
14. Shamley D, Lascurain-Aguirrebeña I, Oskrochi R. Clinical anatomy of the shoulder after treatment for breast cancer. Clin Anat 2014 Apr;27(3):467-477. [doi: 10.1002/ca.22267] [Medline: 23836595]
15. Garrett B, Taverner T, Gromala D, Tao G, Cordingley E, Sun C. Virtual reality clinical research: promises and challenges. JMIR Serious Games 2018 Oct 17;6(4):e10839. [doi: 10.2196/10839] [Medline: 30333096]
16. Yang Z, Rafiei MH, Hall A, et al. A novel methodology for extracting and evaluating therapeutic movements in game-based motion capture rehabilitation systems. J Med Syst 2018 Nov 7;42(12):255. [doi: 10.1007/s10916-018-1113-4] [Medline: 30406430]
17. Jimenez YA, Cumming S, Wang W, Stuart K, Thwaites DI, Lewis SJ. Patient education using virtual reality increases knowledge and positive experience for breast cancer patients undergoing radiation therapy. Support Care Cancer 2018 Aug;26(8):2879-2888. [doi: 10.1007/s00520-018-4114-4] [Medline: 29536200]
18. Feyzioğlu Ö, Dinçer S, Akan A, Algun ZC. Is Xbox 360 Kinect-based virtual reality training as effective as standard physiotherapy in patients undergoing breast cancer surgery? Support Care Cancer 2020 Sep;28(9):4295-4303. [doi: 10.1007/s00520-019-05287-x] [Medline: 31907649]
19. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS One 2016;11(3):e0150205. [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
20. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]
21. Brennan L, Kessie T, Caulfield B. Patient experiences of rehabilitation and the potential for an mHealth system with biofeedback after breast cancer surgery: qualitative study. JMIR Mhealth Uhealth 2020 Jul 29;8(7):e19721. [doi: 10.2196/19721] [Medline: 32687476]
22. Farič N, Yorke E, Varnes L, et al. Younger adolescents' perceptions of physical activity, exergaming, and virtual reality: qualitative intervention development study. JMIR Serious Games 2019 Jun 17;7(2):e11960. [doi: 10.2196/11960] [Medline: 31210135]
23. Zhang T, Booth R, Jean-Louis R, et al. A primer on usability assessment approaches for health-related applications of virtual reality. JMIR Serious Games 2020 Oct 28;8(4):e18153. [doi: 10.2196/18153] [Medline: 33112243]
24. Jerdan SW, Grindle M, van Woerden HC, Boulos MNK. Head-mounted virtual reality and mental health: critical review of current research. JMIR Serious Games 2018 Jul 6;6(3):e14. [doi: 10.2196/games.9226] [Medline: 29980500]
25. Colaizzi PF. Psychological research as the phenomenologist views it. In: Valle RS, King M, editors. Existential-Phenomenological Alternatives for Psychology: Oxford University Press; 1978.
26. van Amstel FKP, Tol J, Sessink KH, van der Graaf WTA, Prins JB, Ottevanger PB. A specific distress cutoff score shortly after breast cancer diagnosis. Cancer Nurs 2017;40(3):E35-E40. [doi: 10.1097/NCC.0000000000000380] [Medline: 27135753]
27. Chiou YJ, Lee CY, Li SH, Chong MY, Lee Y, Wang LJ. Screening for psychologic distress in Taiwanese cancer inpatients using the national comprehensive cancer network Distress Thermometer: the effects of patients' sex and chemotherapy experience. Psychosomatics 2017;58(5):496-505. [doi: 10.1016/j.psym.2017.03.016] [Medline: 28549574]
28. Kyngäs H. Inductive content analysis. In: Kyngäs H, Mikkonen K, Kääriäinen M, editors. The Application of Content Analysis in Nursing Science Research: Springer Cham; 2020:13-21 URL: http://link.springer.com/10.1007/978-3-030-30199-6 [accessed 2024-02-28] [doi: 10.1007/978-3-030-30199-6]
29. Lincoln YS, Guba EG. Naturalistic Inquiry, 1st edition: Sage Publications Inc, Newbury Park; 1985.
30. Sligo FX, Jameson AM. The knowledge-behavior gap in use of health information. J Am Soc Inf Sci 2000 May 18;51(9):858-869. [doi: 10.1002/(SICI)1097-4571(2000)51:9<858::AID-ASI80>3.0.CO;2-Q]
31. Moore N, Yoo S, Poronnik P, Brown M, Ahmadpour N. Exploring user needs in the development of a virtual reality-based advanced life support training platform: exploratory usability study. JMIR Serious Games 2020 Aug 7;8(3):e20797. [doi: 10.2196/20797] [Medline: 32763877]
32. Pazzaglia C, Imbimbo I, Tranchita E, et al. Comparison of virtual reality rehabilitation and conventional rehabilitation in Parkinson's disease: a randomised controlled trial. Physiotherapy 2020 Mar;106:36-42. [doi: 10.1016/j.physio.2019.12.007] [Medline: 32026844]
33. Zasadzka E, Pieczyńska A, Trzmiel T, Hojan K. Virtual reality as a promising tool supporting oncological treatment in breast cancer. Int J Environ Res Public Health 2021 Aug 19;18(16):8768. [doi: 10.3390/ijerph18168768] [Medline: 34444513]
34. Zhang LL, Wang JQ, Qi RR, Pan LL, Li M, Cai YL. Motion sickness: current knowledge and recent advance. CNS Neurosci Ther 2016 Jan;22(1):15-24. [doi: 10.1111/cns.12468] [Medline: 26452639]

35. Kaufeld M, De Coninck K, Schmidt J, Hecht H. Chewing gum reduces visually induced motion sickness. Exp Brain Res 2022 Feb;240(2):651-663. [doi: 10.1007/s00221-021-06303-5] [Medline: 34997261]
36. Kaufeld M, Bourdeinik J, Prinz LM, Mundt M, Hecht H. Emotions are associated with the genesis of visually induced motion sickness in virtual reality. Exp Brain Res 2022 Oct;240(10):2757-2771. [doi: 10.1007/s00221-022-06454-z] [Medline: 36068308]

Abbreviations

CHQ-12: Chinese Health Questionnaire–12
CORE: Clinical Outcomes Research Experts
DT: Distress Thermometer
HMD: head-mounted display
MS: motion sickness
ROM: range of motion
VR: virtual reality

Edited by G Eysenbach, T Leung; submitted 03.11.22; peer-reviewed by S Schär, X Liu; revised version received 31.10.23; accepted 30.12.23; published 16.04.24.

Please cite as:
Wu SC, Chuang CW, Liao WC, Li CF, Shih HH
Using Virtual Reality in a Rehabilitation Program for Patients With Breast Cancer: Phenomenological Study
JMIR Serious Games 2024;12:e44025
URL: https://games.jmir.org/2024/1/e44025
doi:10.2196/44025

Original Paper

Digital Interventions for Stress Among Frontline Health Care Workers: Results From a Pilot Feasibility Cohort Trial

Caroline W Espinola[1,2*], MSc, MD; Binh Nguyen[3*], BSc; Andrei Torres[4*], BArch, MEng; Walter Sim[2], BSc; Alice Rueda[2], PhD; Lindsay Beavers[5,6], BSc, MPT; Douglas M Campbell[5,7,8,9], MSc, MD; Hyejung Jung[10], MSc; Wendy Lou[10], PhD; Bill Kapralos[4], PhD; Elizabeth Peter[11], PhD; Adam Dubrowski[4], PhD; Sridhar Krishnan[3], PhD; Venkat Bhat[1,2], MSc, MD

[1]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[2]Interventional Psychiatry Program, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada
[3]Department of Electrical, Computer, and Biomedical Engineering, Toronto Metropolitan University, Toronto, ON, Canada
[4]maxSIMhealth Group, Ontario Tech University, Oshawa, ON, Canada
[5]Allan Waters Family Simulation Program, Unity Health Toronto, Toronto, ON, Canada
[6]Department of Physical Therapy, University of Toronto, Toronto, ON, Canada
[7]Neonatal Intensive Care Unit, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada
[8]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, ON, Canada
[9]Department of Pediatrics, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[10]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[11]Lawrence S. Bloomberg, Faculty of Nursing, University of Toronto, Toronto, ON, Canada
[*]these authors contributed equally

Corresponding Author:
Venkat Bhat, MSc, MD
Department of Psychiatry
University of Toronto
250 College Street, 8th floor
Toronto, ON
Canada
Phone: 1 416 360 4000 ext 76404
Email: venkat.bhat@utoronto.ca

Abstract

Background: The COVID-19 pandemic has challenged the mental health of health care workers, increasing the rates of stress, moral distress (MD), and moral injury (MI). Virtual reality (VR) is a useful tool for studying MD and MI because it can effectively elicit psychophysiological responses, is customizable, and permits the controlled study of participants in real time.

Objective: This study aims to investigate the feasibility of using an intervention comprising a VR scenario and an educational video to examine MD among health care workers during the COVID-19 pandemic and to use our mobile app for longitudinal monitoring of stress, MD, and MI after the intervention.

Methods: We recruited 15 participants for a compound intervention consisting of a VR scenario followed by an educational video and a repetition of the VR scenario. The scenario portrayed a morally challenging situation related to a shortage of life-saving equipment. Physiological signals and scores of the Moral Injury Outcome Scale (MIOS) and Perceived Stress Scale (PSS) were collected. Participants underwent a debriefing session to provide their impressions of the intervention, and content analysis was performed on the sessions. Participants were also instructed to use a mobile app for 8 weeks after the intervention to monitor stress, MD, and mental health symptoms. We conducted Wilcoxon signed rank tests on the PSS and MIOS scores to investigate whether the VR scenario could induce stress and MD. We also evaluated user experience and the sense of presence after the intervention through semi–open-ended feedback and the Igroup Presence Questionnaire, respectively. Qualitative feedback was summarized and categorized to offer an experiential perspective.

Results: All participants completed the intervention. Mean pre- and postintervention scores were respectively 10.4 (SD 9.9) and 13.5 (SD 9.1) for the MIOS and 17.3 (SD 7.5) and 19.1 (SD 8.1) for the PSS. Statistical analyses revealed no significant pre-to postintervention difference in the MIOS and PSS scores (*P*=.11 and *P*=.22, respectively), suggesting that the experiment did not acutely induce significant levels of stress or MD. However, content analysis revealed feelings of guilt, shame, and betrayal, which relate to the experience of MD. On the basis of the Igroup Presence Questionnaire results, the VR scenario achieved an above-average degree of overall presence, spatial presence, and involvement, and slightly below-average realism. Of the 15 participants, 8 (53%) did not answer symptom surveys on the mobile app.

Conclusions: Our study demonstrated VR to be a feasible method to simulate morally challenging situations and elicit genuine responses associated with MD with high acceptability and tolerability. Future research could better define the efficacy of VR in examining stress, MD, and MI both acutely and in the longer term. An improved participant strategy for mobile data capture is needed for future studies.

Trial Registration: ClinicalTrails.gov NCT05001542; https://clinicaltrials.gov/study/NCT05001542

International Registered Report Identifier (IRRID): RR2-10.2196/32240

(JMIR Serious Games 2024;12:e42813) doi:10.2196/42813

KEYWORDS

virtual reality; simulation; mobile app; stress; moral distress; moral injury; COVID-19; mobile phone

Introduction

Background

The COVID-19 pandemic has exerted unprecedented strain on health care workers (HCWs) globally [1]. Frontline HCWs have been forced to make difficult medical decisions that are contrary to their moral and professional principles and to work in conditions where they cannot meet standards of quality care [2,3], which has put them at a greater risk of experiencing moral distress (MD) than possibly ever before [4,5]. Distressing situations such as being forced to deal with a shortage of personal protective equipment and having to prioritize who will receive life-sustaining treatment have become common during the pandemic. For HCWs, experiencing such situations may cause significant emotional burden and induce the phenomenon of MD [6-8]. MD is defined as distress stemming from the inability to enact actions believed to be morally right owing to external constraints [8,9]. Moral injury (MI), an extreme form of MD, can occur when individuals witness or perpetrate actions that violate deeply held moral beliefs, resulting in severe emotional reactions with long-lasting consequences [7]. However, further investigation is needed to enable a more precise distinction between MD and MI [7].

The first description of MI was made in the military context by Shay [10] and was defined as a betrayal of moral character, usually as a result of the actions of a person in a position of authority [10], leading to feelings of powerlessness, helplessness, and loss of faith in humanity [7,10]. Shay [11] argues that MI occurs when the following conditions are met: (1) there has been a betrayal of what is considered right (2) by someone holding legitimate authority and (3) in high-stakes situations. Litz et al [12] expanded the concept of MI to include “the lasting psychological, biological, spiritual, behavioral, and social impact of perpetrating, failing to prevent, or bearing witness to acts that transgress deeply held moral beliefs and expectations.” As part of the definition, the authors also defined potentially morally injurious events (PMIEs) as the acts of perpetrating, failing to prevent harm, or bearing witness to acts that transgress deeply held moral beliefs [12]. Experiencing a PMIE is frequently associated with feelings of betrayal, guilt, shame, and self-blame [13]. Furthermore, PMIEs may not only cause acute MD but can also have long-term consequences because MD and MI may develop weeks or months after a PMIE [14].

MI was originally associated with, and frequently co-occurs with, posttraumatic stress disorder (PTSD) [13], which has been conceptualized as a fear-related disorder [15,16]. However, MI has not yet been defined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition [16], and a PMIE does not necessarily fulfill the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition criterion A for PTSD. The concept of MI was conceived to encompass the following criteria, among others: (1) reexperiencing self-referential moral emotions (eg, anger, guilt, and shame); (2) strong negative beliefs about the self, the world, and others; and (3) self-destructive behaviors that inflict severe distress or functional impairment [17,18]. In addition, emerging literature has defined MI as being mechanistically different from PTSD [13,15]. A positron emission tomography study in veterans with PTSD showed that regional blood glucose metabolism differed according to the nature of traumatic exposure as follows: the group with PTSD owing to danger-based traumas (ie, life-threatening events) showed higher metabolism in the amygdalae; by contrast, the group with PTSD secondary to non–danger-based traumas (eg, MI by self or others) had increased metabolism in the precuneus [19], a region that has been associated with the processing of self-referential feelings (eg, shame and guilt) [15]. Therefore, further research is needed to determine the ecological validity of MI as an independent diagnostic category [13]. In addition, there is a need to investigate specific interventions for MI because it has been found to not generally respond to evidence-based treatments for PTSD [12,17]; for example, moral resilience training, the development of emotional intelligence skills, and strategies for promoting moral repair have already been proposed as specific treatments for MI and are currently under investigation [5,17,20].

Although MI has been largely studied in military contexts [17], it is also applicable to HCWs, particularly in light of the COVID-19 pandemic. However, MI and PMIEs are poorly understood in this context. Čartolovni et al [7] argue that MI occurs in HCWs when they experience PMIEs involving high-stakes situations that are beyond their control. To investigate MI in the COVID-19 context, Rushton et al [5] conducted a survey with frontline HCWs and reported an overall prevalence rate of 32% for MI, with nurses being the most affected (38%). Fewer years of experience were positively associated with MI, whereas religious affiliation or spirituality and higher levels of moral resilience were associated with lower MI scores. In addition, the study showed a moderate correlation between MI and various ethically challenging situations, such as experiencing negative consequences at work after expressing safety concerns, working with limited resources, and carrying out decisions of others which threaten one's own values [5].

Litam and Balkin [4] examined the relationship between MI and the professional quality of life in a convenience sample of HCWs during the COVID-19 pandemic. The authors reported that secondary traumatic stress was a strong predictor of MI in frontline HCWs, but the contribution of compassion, satisfaction, and burnout to MI scores was nonsignificant. Of note, nurses had significantly higher burnout scores than physicians. Zerach and Levi-Belz [21] conducted a survey to investigate the patterns of exposure to PMIEs in a sample of HCWs and social care workers during the COVID-19 pandemic. The prevalence rate of symptoms of MI was 40%, with betrayal events being the most frequent PMIEs with a prevalence rate of 62%. In general, exposure to PMIEs was positively related to perceived stress, depression, anxiety, and self-criticism, whereas it was negatively associated with self-compassion. Interestingly, the duration of care for patients with COVID-19 was not associated with MI [21].

To increase the ecological validity of MI as a diagnostic entity, the experiences of the MD-MI continuum should be examined using accurate methods [13]. To date, several measurement instruments have been developed to identify MI outside of military contexts, including the Moral Injury Symptom Scale–Healthcare Professionals version [22] and the Moral Injury Outcome Scale (MIOS) [18]. The MIOS is a self-rated scale, developed as an assessment tool to evaluate MI as a multidimensional outcome [18]. This scale comprises 10 binary (*yes* or *no*) questions and 15 five-point Likert scale questions about experiencing a PMIE and feelings associated with this event; higher scores indicate greater severity of MI symptoms. At the end, the MIOS has an additional 7-point Likert scale question that assesses the extent to which the experience of PMIEs has interfered in one's self-care or caused functional impairment (from *not at all* to *extremely*). The MIOS is in the final stage of development by the MIOS Consortium [18].

Conducting interventional studies to investigate the impact of PMIEs on mental health in real-world settings is challenging owing to operational constraints. This is especially true in health care, where limitations imposed by patient privacy regulations may make traditional clinical trials in MI impractical. Another important aspect to consider is the ethical implications of submitting an already strained workforce to moral stressors in an uncontrolled real-world environment such as an intensive care unit (ICU). A promising strategy to address these limitations is the use of virtual reality (VR) scenarios. VR is a powerful technology for examining mental health and the MD-MI continuum because it offers several advantages over traditional observational research in naturalistic environments. First, VR allows researchers to observe, monitor, and potentially support participants in fully controlled environments in real time [23]; therefore, it is safer and provides more accurate measures of one's reactions to ethically challenging situations compared with observational studies in naturalistic environments. Second, VR allows for the design of fully customizable scenarios [23], making it especially suitable to simulate real-world scenarios in health care that otherwise would be impractical to replicate. As traumatic events in both PTSD and MI are highly idiosyncratic, and treatment for PTSD requires exposure to individual cues, we assume that virtual environment customization should be a critical feature to provide personalized and effective interventions to treat MD and MI [24]. In addition, extensive evidence has demonstrated the effectiveness of VR-based interventions for PTSD [25-27]. Third, VR environments can effectively elicit real psychophysiological responses because individuals are immersed in virtual scenarios as if these were real events, with the advantage of enabling real-time data capture [23,24]. All these advantages make VR-based trials ideal to study the MD-MI phenomena in HCWs. However, no prior research has investigated the feasibility of VR interventions to examine MD and MI in the context of the COVID-19 pandemic.

Objectives

The overarching goal of this study was to determine the feasibility of using a compound VR intervention to examine MD and MI among HCWs during the COVID-19 pandemic. To achieve this, we designed a VR scenario in which HCWs faced a morally challenging situation in a midpandemic hospital environment while being monitored for acute psychological and physiological measures of stress. As outlined in our protocol paper [28], our aims were to (1) evaluate the feasibility of using a VR scenario to simulate the experience of a COVID-19–related morally challenging event by using measures of tolerability, dropouts, and suitability of the virtual scenario; (2) assess the potential of our VR scenario to elicit mild stress and MD, as measured by quantitative self-report questionnaires as well as qualitative analyses of semistructured interviews; and (3) investigate the feasibility of our novel mobile app (*DiiG App*) for longitudinal monitoring of stress and MD in naturalistic settings in the 8 weeks after the intervention.

On the basis of the findings with PTSD [25-27], we hypothesized that VR scenarios would be a feasible method for assessing MD and MI. Given the ability of VR to generate genuine responses, we additionally hypothesized that our virtual scenario would significantly increase stress levels and elicit feelings and symptoms associated with MD and MI. Finally, we hypothesized that our mobile app would successfully capture symptoms associated with stress and MD in the 8-week follow-up.

To the best of our knowledge, this pilot study is the first to assess the feasibility of using a VR scenario to simulate the experience of a morally challenging event related to the

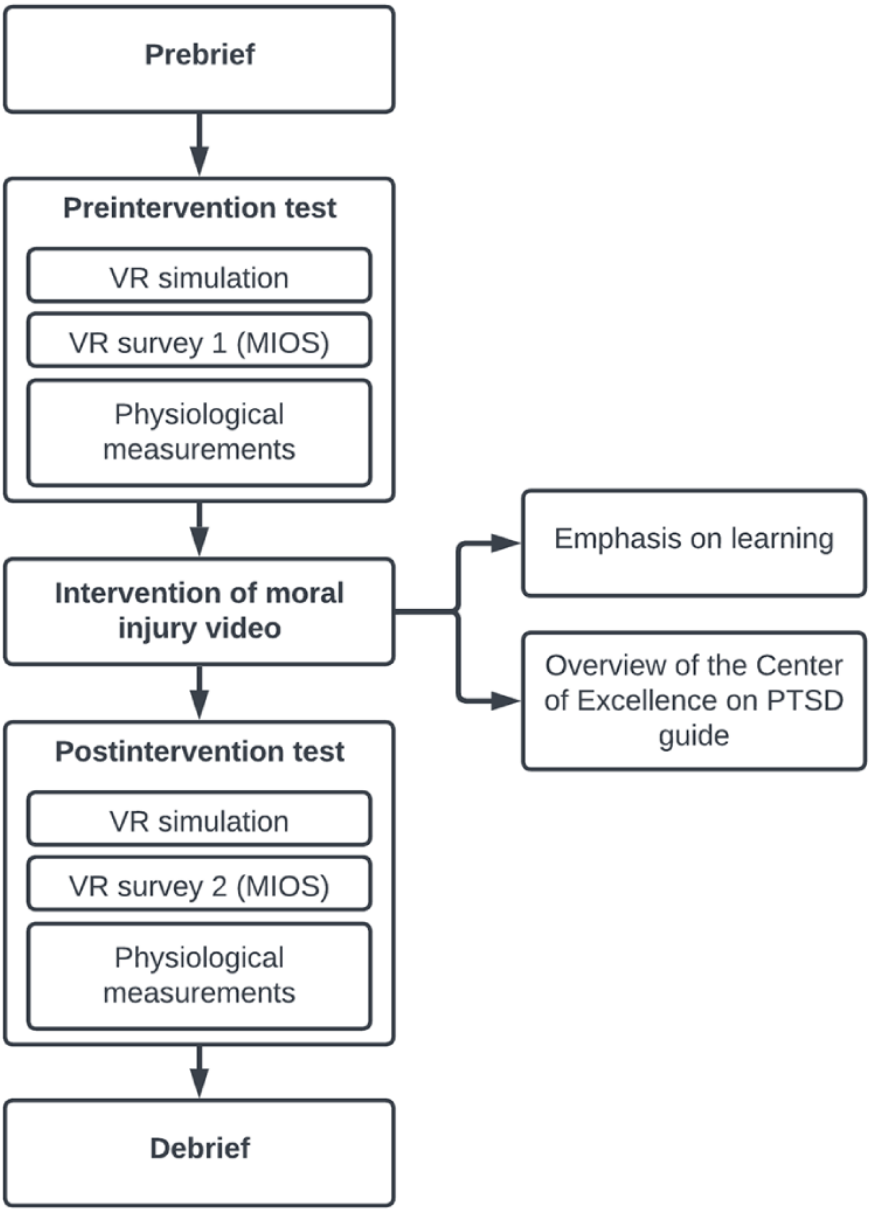
COVID-19 pandemic by HCWs while assessing its acute perceptual, psychological, and physiological effects in real time.

Methods

Study Design

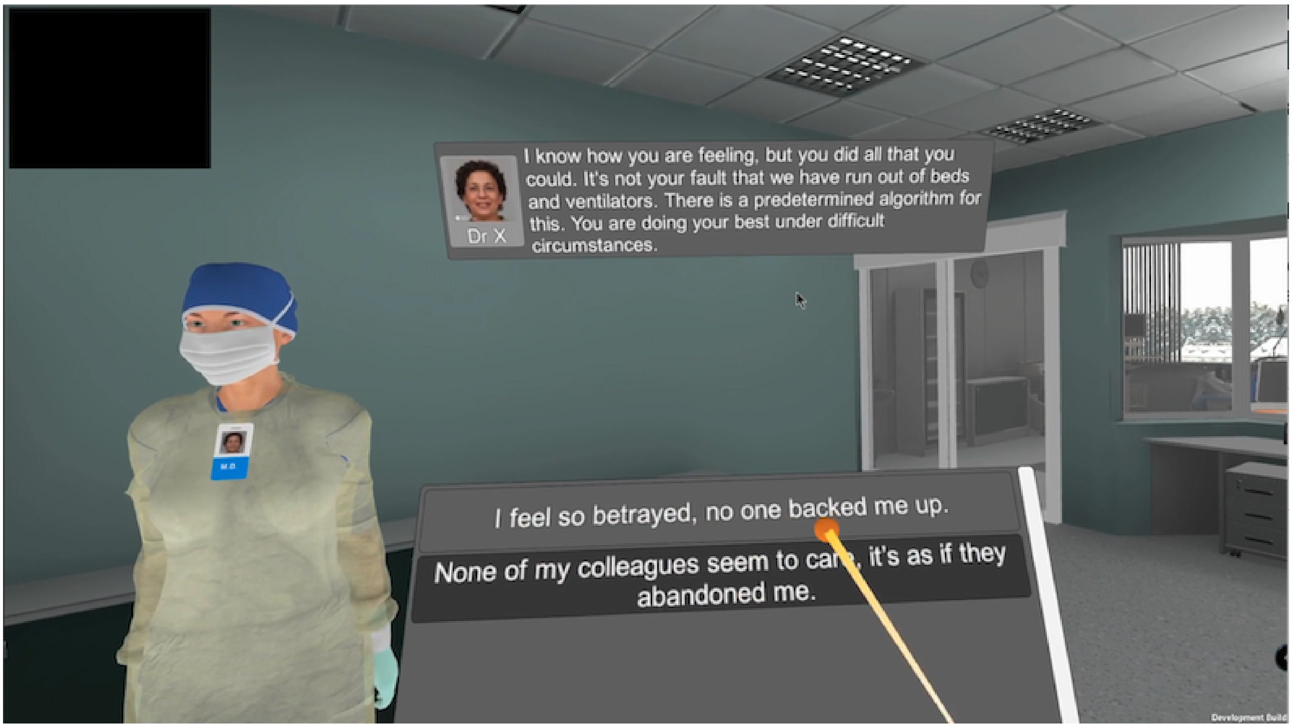
In this single-cohort pilot study (ClinicalTrials.gov: NCT05001542), we adopted a multimethod approach in a pretest-posttest design to develop a compound intervention consisting of three successive parts: (1) a VR scenario to simulate a morally complex situation, (2) an educational video on MI and appropriate mitigation strategies, and (3) a repetition of the VR scenario. The intervention was followed by longitudinal data collection of mental health and MI surveys using our mobile app. The MI educational video was based on the Center of Excellence on PTSD guide [29] that summarized the causes and identifiers of MI and potential interventions to mitigate MD. The effectiveness of the VR-based educational intervention was assessed using the MIOS [18], the Perceived Stress Scale (PSS) [30], and the Igroup Presence Questionnaire (IPQ) [31]. The PSS is a self-reported measure of stress, whereas the IPQ evaluates the experience of presence during the VR scenario. As previously mentioned, the MIOS is a self-rated scale that was developed as an assessment tool to evaluate MI. For the purposes of this pilot study, we adopted a brief version of the MIOS (hereinafter referred to as *the MIOS*), which comprises 10 five-point Likert scale questions and 4 binary (*yes* or *no*) questions [32]. During the VR scenario, respiratory impedance, electrocardiography (ECG), galvanic skin response, and photoplethysmography were continuously collected. In addition to the original signals, we extracted the derivation of these signals, including ECG pulse rate, ECG RR interval, respiratory rate, and elevated respiratory rate. A visualization of the VR experimental flow can be seen in Figure 1. Further details on the intervention and data collection have been explained and outlined in the paper by Nguyen et al [28].

Figure 1. Flowchart of the virtual reality (VR) experiment. MIOS: Moral Injury Outcome Scale; PTSD: posttraumatic stress disorder.



The experimental session was divided into prebrief, preintervention test, intervention video, postintervention test, and debrief components (Figure 1). The preintervention test and postintervention test were conducted in VR, whereas the prebrief and debrief occurred outside the virtual environment. The MIOS was performed at 4 time points as follows: as a paper-based version for the prebrief and debrief and in the virtual scenario for the preintervention test and postintervention test. The PSS was performed twice, at prebrief and debrief. The MIOS and the PSS focus on symptoms of MD and stress, respectively, over the last month. However, when answering these scales, participants were told to rate symptoms at that exact moment. The goal of the prebrief was to explain how the physiological data would be collected and prepare the participant for the VR scenario; it consisted of an orientation to the virtual space and equipment, safety precautions, and the expected outcome of the study. During the preintervention test, participants were immersed in the VR scenario where they took on the role of a physician in an ICU during the COVID-19 pandemic. To experience the VR scenario, participants used a VR headset and 2 wireless controllers that tracked their head and hand movements, mapping it to an avatar. Semitranslucent panels were displayed as spatial elements in the VR scenario (Figure 2), providing information to the participant in the form of the dialogue panel (which displayed the current nonplayable character's photograph, name, and the text version of the dialogue being spoken) and the interaction panel (which displayed a list of available choices and responses for the participant to choose from).

Figure 2. User interface displaying the dialogue and interaction panels.



In the scenario, a shortage of life-saving equipment resulted in the decision to move a ventilator from 1 patient to another patient who had a greater chance of survival. After being informed of this, the participant's avatar appeared in the next scene, where they had to communicate this decision to the first patient's family and respond to the family's reactions of frustration and anger. After completing the preintervention test and while still immersed in the VR scenario, participants watched a brief 2D educational video comprising key concepts of MD and MI and adaptive behaviors to cope with morally complex situations at the individual, team, and organizational levels. Participants then completed the postintervention test, which consisted in a repetition of the VR scenario played in the preintervention test. Finally, in the debrief, participants were asked open-ended questions to encourage them to describe their experiences in the virtual setting, followed by an exit survey [28].

After the experiment, participants were instructed to use our mobile app [33] to collect passive and active data for distress monitoring during the following 8 weeks. As MI may have a delayed onset, such data collection allows for longer-term monitoring of emotions associated with MD, offering insights into the distress experienced in real time.

Participants

Participants were recruited and enrolled between May 2021 and August 2021 from the 3 affiliated hospitals at Unity Health Toronto. Participants were enrolled if they were an HCW currently providing health care at their respective hospital of employment, aged ≥18 years, and owned a mobile phone (an Android mobile phone with operating system version 6.0 or above or an iPhone with operating system version 11.0 or above).

Statistical Analysis

As this was a pilot feasibility trial, we summarized dropout rates, easiness of use, tolerability, acceptability, and utility using counts and percentages. Continuous data were summarized using range, mean and SD, and median and IQR. To assess the effect of the VR scenario on symptoms of MI, we compared MIOS scores across the 4 time points using a Friedman test. In addition, follow-up MIOS scores were compared with the score at prebrief using Wilcoxon signed rank tests with Bonferroni correction (.05/3=.0167) to adjust for multiple comparisons. As PSS scores were collected only at 2 time points (ie, at prebrief and debrief), a Wilcoxon signed rank test was used to compare the difference in the PSS scores between these 2 time points. A *P* value of <.05 was considered significant unless otherwise specified. We performed statistical analysis using SAS 9.4 (SAS Institute Inc).

Quantitative Analysis

Stress and MD Analysis

In this feasibility study, we piloted the application of the MIOS to assess MD both acutely and longitudinally. As mentioned in the *Study Design* section, MIOS was administered during the prebrief, preintervention test, postintervention test, and debrief. Participants were also prompted to complete MIOS on the mobile app in the 8 weeks after the intervention for a longitudinal assessment of MD and MI. All questionnaires used in the mobile app (eg, the MIOS and the PSS) are available in the appendices of the study by Nguyen et al [28].

IPQ Assessment

To objectively assess user experience within the VR scenario, we adopted the IPQ, which is a questionnaire for measuring the sense of presence experienced in a virtual environment [31]. Composed of 14 questions (answered on a 6-point Likert scale), the IPQ has a high reliability (Cronbach α=.87) and outputs four items (1 general item, not belonging to a subscale, and 3 subscales): (1) general presence (sense of being there), (2) spatial presence (the sense of being physically present in the virtual environment), (3) involvement (measuring the attention devoted to the virtual environment), and (4) experienced realism (measuring the subjective experience of realism in the virtual environment).

Hereinafter, the 4 outputs will be referred to as *IPQ components*. More information about the construction and structure of the scale and the IPQ's reliability analysis is available on the Igroup project consortium website [34,35].

Mobile Data Analysis

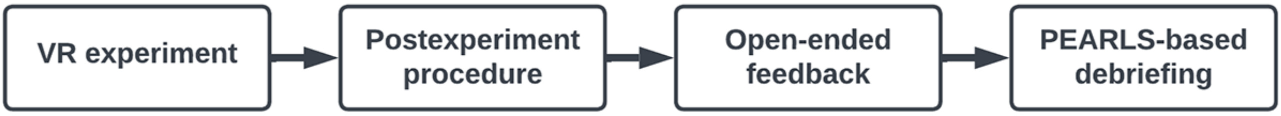
After participating in the intervention, participants were instructed by our research staff to download and regularly use our mobile app to answer surveys in the 8-week follow-up. Participants received push notifications on the mobile app 3 times weekly to answer short versions of the scales related to depression (2-item Patient Health Questionnaire), anxiety (2-item Generalized Anxiety Disorder), stress (4-item PSS) MI (4-item MIOS), and loneliness (3-item University of California Los Angeles Loneliness Scale). With the exception of the 3-item University of California Los Angeles Loneliness Scale, participants were also asked to answer the full version of these scales once weekly. Short versions of the scales were used on weekdays to minimize participant burden. The mobile app also had the option of collecting passive data from built-in smartphone sensors (GPS and accelerometer) from participants who provided in-app consent to gather information on distance traveled and activity patterns. Details on the mobile data collection were previously overviewed in the study by Nguyen et al [28]. We used in-app automated survey reminders to promote app use.

Qualitative Analysis

Content Analysis

We performed a content analysis on the data collected from the scenario debriefing conducted immediately after the compound intervention. Qualitative content analysis is a method to interpret meaning from text data and draw conclusions from words, themes, or concepts that occur in the text, in reference to their context, so that research questions can be answered [36]. We used inductive category development by becoming immersed in the data and allowing insights on categories to emerge from the data [37]. The scenario debriefing consisted of a semistructured interview that allowed participants to answer open-ended questions about their overall experience, followed by a semistructured debriefing methodology (the interview guide is included in Multimedia Appendix 1). The researchers (BN and AT) who collected the VR data were trained using the Promoting Excellence and Reflective Learning in Simulation (PEARLS) health care debriefing tool [38], a simulation debriefing framework to help learners assess their experience within a safe environment. A flow diagram of the debriefing can be seen in Figure 3. After completion of the intervention, we conducted a postexperiment procedure, which consisted of removing the VR headset from the participant but keeping the physiological sensors attached. In addition, we confirmed with the participant that they were able to continue with the debriefing.

Figure 3. Flow diagram of the debriefing. PEARLS: Promoting Excellence and Reflective Learning in Simulation; VR: virtual reality.



During the open-ended feedback part of the debrief, we asked participants to speak freely about their experience with the experiment. We specifically asked the following questions:

1. "What suggestions or feedback would you give to improve the scenarios? Please comment on what can be improved, what can be more realistic, and any deviation from real-life applications."
2. "Could you share something that you have learned about moral injury today? How might this apply to your clinical practice?"

The research questions we sought to answer with our content analysis from this feedback were as follows:

1. "How can the VR scenario be improved?"

2. “How accessible and relevant was our intervention?”

We subsequently conducted scenario-based debriefing using the PEARLS [38] methodology, which involved an exploration of the following predetermined topics: participant experience with the technology used, decision-making during the scenario, and emotions elicited during the scenario. The research questions we sought to answer with the content analysis from the scenario-based debriefing were as follows:

1. “What is the overall user experience of participants with the VR technology?”
2. “What were the determining factors for the decisions that participants made in the scenario?”
3. “How did the scenario make the participant feel?”

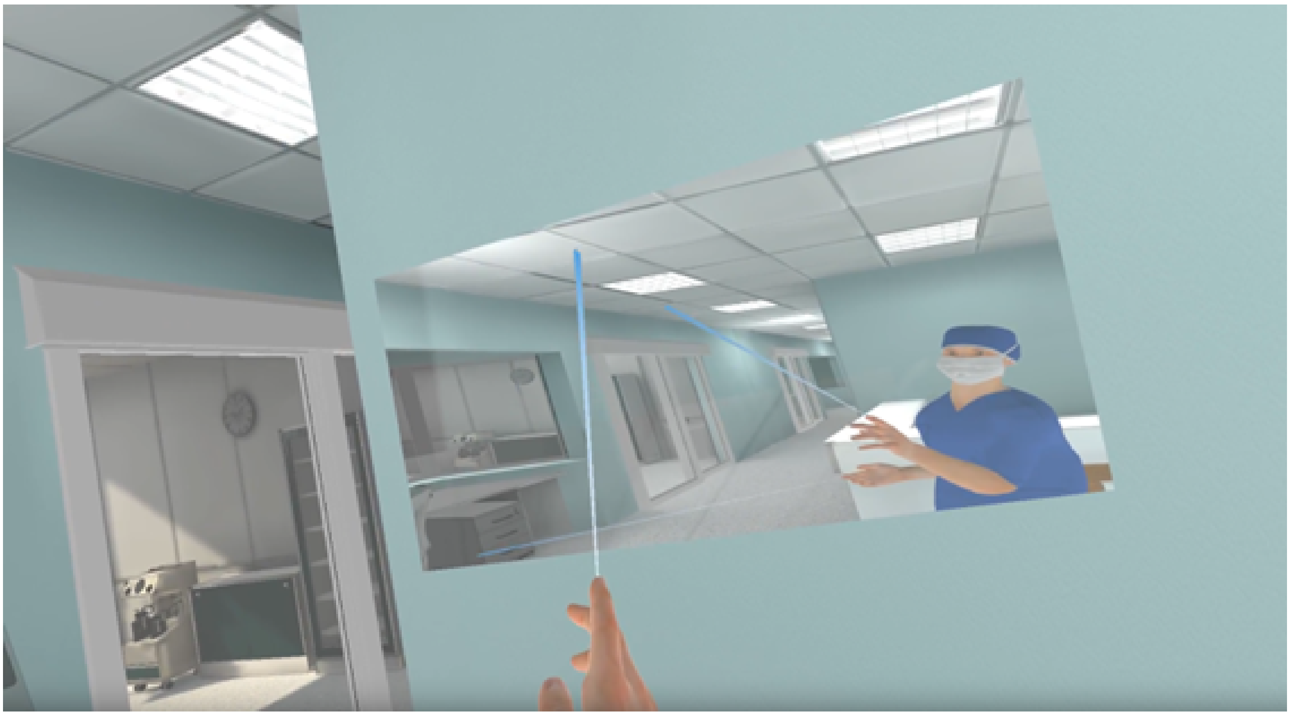
The PEARLS structure is a well-validated debriefing tool that is typically used to provide introspection on performance for a simulation participant [38]. It has been used extensively in the simulation literature, including a recent user qualitative study with patient-led simulations [39]. A PEARLS debrief integrates 4 main segments: setting the scene, eliciting reactions, description and analysis of the experience, and summary or reflections.

After the debrief, participants were asked to answer a debrief feasibility questionnaire with 3 five-point Likert questions answered on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) about the relevance and utility of the psychoeducational content on MD for real-life situations as well as the ability of the VR scenario to elicit emotions (Multimedia Appendix 1).

User Experience

To assess the user experience within the VR scenario, we evaluated the dropout rate, the feasibility questionnaire, and the qualitative responses provided during the debrief. During the VR scenario, participants had their head and hand movements tracked by the VR headset and controllers, and all movements were mapped into a virtual avatar (Figure 4). To help improve the sense of body ownership (ie, making the users recognize the virtual body as their own) [40], the preintervention test started with a tutorial that had the participants looking at a mirror and moving their head and hands to visualize that their virtual avatar actions reflected their own.

Figure 4. Snapshot of the virtual reality scenario showing the participant’s avatar reflected in a mirror. The blue beam indicates the cursor used to interact with the virtual environment.



Ethics Approval

Ethics approval was obtained from the research ethics board at St. Michael’s Hospital before starting any study activities (21-066).

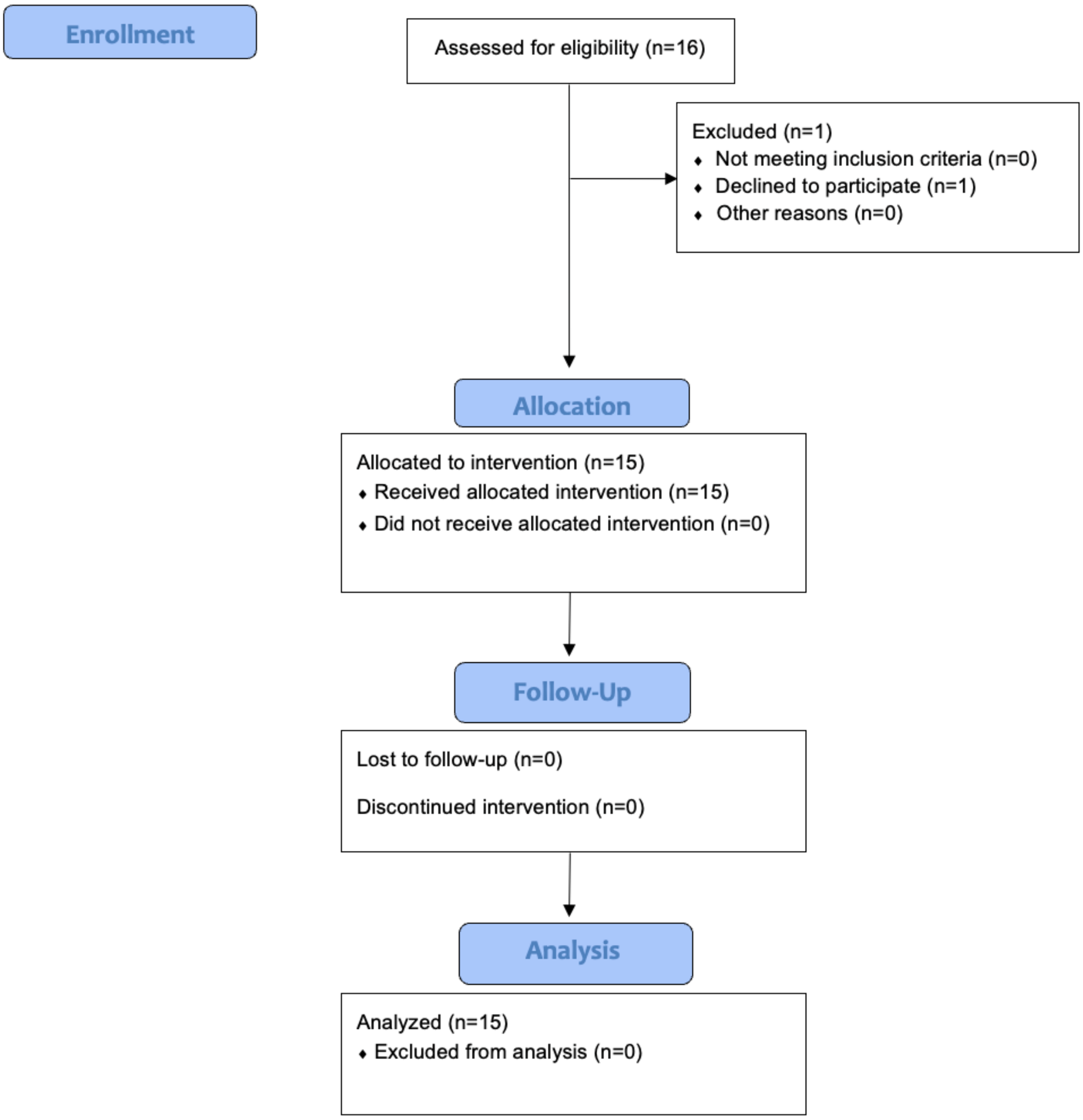
Results

Participants

Participant Flow

A total of 16 participants were assessed for eligibility; 1 (6%) declined to participate, and therefore 15 (94%) participants were allocated to the intervention. All 15 participants received the intervention. No participants were lost to follow-up, and data from all 15 participants were analyzed. Information on participant flow is presented in Figure 5.

Figure 5. Flowchart of participant enrollment and attrition.



Baseline Data

Our sample consisted of 15 HCWs (female participants: n=11, 73%; male participants: n=4, 27%). The participants had a mean age of 32.7 (SD 9.5) years; the male participants had a mean age of 34.3 (SD 4.9) years, whereas the female participants had a mean age of 32.2 (SD 10.9) years. Among the 15 participants, the most common occupations were nursing (n=7, 47%) and medicine (n=3, 20%); other professions included mental health research staff (n=2, 13%), physician assistant (n=1, 7%), educator (n=1, 7%), and graduate student (n=1, 7%). At the time of the experiment, none of the 15 participants had a prior or current COVID-19 infection; however, 4 (27%) had a prior family history of COVID-19 infection. The VR experiments were conducted between May 2021 and August 2021.

Data Analyzed

For 15 participants, MIOS, PSS, IPQ, and mobile data were analyzed. The data of 14 participants were analyzed for the content analysis.

Quantitative Analysis

Stress and MD Analysis

The average MIOS scores for the prebrief, preintervention test, postintervention test, and debrief were 10.4 (SD 9.9), 12.9 (SD 6.9), 12.6 (SD 7.1), and 13.5 (SD 9.1), respectively, with a difference between the debrief and prebrief (between after the intervention and before the intervention) of 3.1 (SD 6.8; Table 1). There was no statistical difference in the MIOS scores at the 5% level when comparing all 4 scores using the Friedman test (Q=4.61; *P*=.20). Using Bonferroni correction (.05/3=.0167), the results showed no significant difference between the prebrief scores and any follow-up score: preintervention test (*P*=.30), postintervention test (*P*=.32), and debrief (*P*=.11). The MIOS is a new scale that is still under development by the MIOS Consortium and has not yet been established for the assessment of MI [18,41].

Table 1. Wilcoxon signed rank test comparing Moral Injury Outcome Scale follow-up scores at preintervention test, postintervention test, and debrief with the prebrief score (n=15)[a].

	Values, mean (SD)	Values, median (IQR; range)	*P* value
Prebrief score	10.4 (9.9)	12 (0 to 17; 0 to 28)	N/A[b]
Preintervention test score	12.9 (6.9)	13 (6 to 17; 3 to 27)	.30[c]
Postintervention test score	12.6 (7.1)	13 (8 to 17; 1 to 28)	.32[c]
Debrief score	13.5 (9.1)	14 (5 to 18; 0 to 32)	.11[c]
Difference (debrief − prebrief)	3.1 (6.8)	1 (−1 to 7; −8 to 18)	.11[c]

[a]There was no statistical difference in the Moral Injury Outcome Scale scores at the 5% level when comparing all 4 scores using the Friedman test (Q=4.61; *P*=.20).

[b]N/A: not applicable.

[c]Follow-up scores were compared with the preintervention test score using the Wilcoxon signed rank test; Bonferroni correction was used (.05/3=.0167), that is, significance at 1.67% was applied.

PSS scores were only collected at 2 time points: at prebrief and debrief. The average PSS scores during the prebrief and the debrief were 17.3 (SD 7.5) and 19.1 (SD 8.1), respectively, with a postintervention test–preintervention test difference of 1.8 (SD 6.0; Table 2). Similar to the MIOS scores, the prebrief and debrief PSS scores were not statistically different (*P*=.22). Tables 1 and 2 summarize the analysis for the MIOS and PSS scores.

Table 2. Wilcoxon signed rank test of the Perceived Stress Scale prebrief and debrief scores (n=15).

	Values, mean (SD)	Values, median (IQR; range)	*P* value
Prebrief score	17.3 (7.5)	15 (12 to 22; 4 to 33)	N/A[a]
Debrief score	19.1 (8.1)	19 (14 to 26; 4 to 33)	N/A
Difference (debrief − prebrief)	1.8 (6.0)	1 (−1 to 7; −11 to 11)	.22[b]

[a]N/A: not applicable.

[b]Wilcoxon signed rank test to test no difference in the distribution between the preintervention test and postintervention test scores.

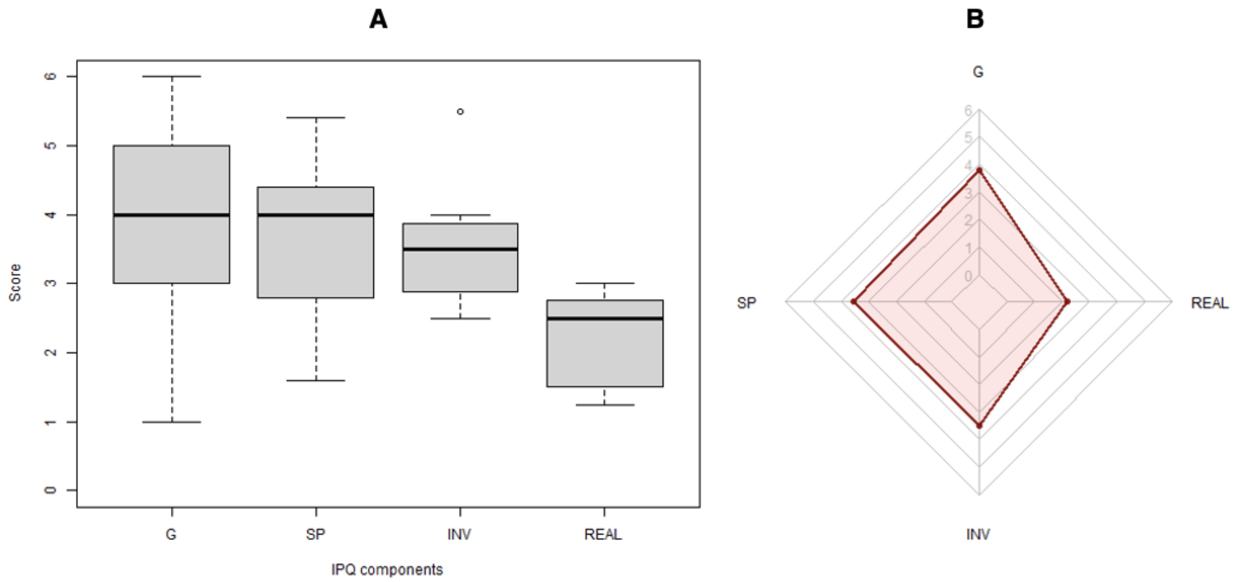
IPQ Assessment

On the basis of the data collected from the 15 participants, the VR scenario achieved an above-average degree of overall presence, spatial presence, and involvement, with slightly below-average realism (Table 3 and Figure 6). Considering that the presence component is influenced by the other 3 components, it makes sense that it has a higher variance and SD, which suggests an opportunity to improve the immersion of the VR scenario. The lowest scoring component was realism, with the lowest variance and SD. These findings are corroborated by the qualitative feedback provided during the debrief session, where only 5 (33%) of the 15 participants commented that the environment felt realistic and that they felt immersed in the experience. By contrast, 1 (7%) of the 15 participants stated that they found the environment more immersive than simulation with real people. The participants' feedback also highlighted other areas for future improvement, particularly regarding the realism component, such as having less restrictive dialogues, making the ICU environment more crowded, improving the voice-over acting features, and having the ICU equipment show patients' physiological data (eg, heart rate monitor).

Table 3. Igroup Presence Questionnaire data statistics.

	Values, mean (SD)	Values, median (IQR)	Variance
General presence	3.80 (1.47)	4.0 (2.0)	2.17
Spatial presence	3.53 (1.16)	4.0 (1.6)	1.34
Involvement	3.48 (0.78)	3.5 (1.0)	0.60
Experienced realism	2.20 (0.67)	2.5 (1.3)	0.45

Figure 6. (A) Box plot showing median, quartiles, extreme data points, and outlier. (B) Radar chart of the overall mean of the 4 Igroup Presence Questionnaire (IPQ) components. It uses a radial axis, with the center representing a score of 0 and the outer outline representing a score of 6. G: general presence; INV: involvement; REAL: experienced realism; SP: spatial presence.



Mobile Data Analysis

The dropout rates for the study app were very high. Of the 15 participants, 8 (53%) did not perform any survey, whereas 7 (47%) completed at least 1 questionnaire. Instead of answering surveys periodically, only 4 (27%) of the 15 participants had >1 set of survey results. There were not sufficient mobile data to provide informative analysis. In the future, an improved participant engagement strategy is needed to help optimize mobile data collection.

Post Hoc Sample Size Calculation

As an exploratory analysis, we calculated post hoc sample sizes using 2-tailed paired *t* tests with a significance level of .05 based on the mean differences in the MIOS and PSS scores observed between the respective scores at prebrief and debrief. The common SDs for each score were estimated using the observed larger SD among the 2 scores. The computed correlations between the 2 scores were used in this calculation. The sample sizes required to achieve a power of 80% were 41 and 95 participants based on the observed results for the MIOS and the PSS, respectively (Multimedia Appendix 2).

Qualitative Analysis

Content Analysis

Content analysis was performed for 14 (93%) of the 15 participants because technical issues compromised the speech recording of the 15th participant. Common references to real-life experiences were recognized in the content analysis, with the most frequent themes being the following: the virtual characters' choices during the experiment were too restrictive (10/14, 71%), feelings of some guilt or shame (8/14, 57%), no feelings of failure or being punished (7/14, 50%), no guilt or shame (6/14, 43%), need of organizational support to deal with the morally challenging situation presented in the experiment (7/14, 50%), numbness (5/14, 36%), and the VR scenario was immersive, real, or engaging (5/14, 36%). Of the 14 participants, 1 (7%; participant 13) provided contradictory responses to feelings of guilt and shame, once saying that they did experience these feelings and once saying that they did not. Furthermore, 2 (14%) of the 14 participants considered the learning experience about MD and MI valuable and useful to their daily practice. A complete summary of the content analysis is provided in Multimedia Appendix 3.

Participants also recommended some specific areas of improvement in the VR scenario; for example, the following suggestions were made by 1 (7%) of the 14 participants: the patient's vital signs were at a normal range although he was experiencing respiratory failure, the skin color should be consistent with that of the participant (all virtual characters were White), the scenario was unrealistic because other interventions apart from the ICU ventilator should have been portrayed, and photographs of the patient should have been added to better customize the character's appearance. Finally, 2 (14%) of the 14 participants reported not being able to relate to religious mentions of God in the VR scenario.

User Experience

Although only 3 (20%) of the 15 participants reported prior experience with VR headsets (Multimedia Appendix 4), there were no dropouts during the VR scenario (Figure 5). As we had expected that new VR users could potentially experience nausea or disorientation, participants were reminded multiple times during the prebrief that they could pause or stop the session at any moment. Having said that, of the 15 participants, 14 (93%) did not report any side effects; only 1 (7%) participant reported claustrophobia and slight anxiety at first, but these feelings quickly subsided, and the participant was able to complete the VR scenario without any further side effects or complaints. Finally, all participants agreed that the VR platform and scenario were easy to navigate (Multimedia Appendix 4). Regarding the debrief feasibility questionnaire, of the 15 participants, 6 (40%) agreed that they learned about MD and interventions, and 11 (73%) agreed that the knowledge about MD and interventions will help them perform better in real-life events (Multimedia Appendix 5). Although only 8 (53%) of the 15 participants agreed that the VR simulation managed to make them experience the same emotions as they would in a real-life event (Multimedia Appendix 5), during the qualitative debrief, common emotions cited included some guilt, shame, betrayal, and isolation, which are consistent with MD.

Discussion

Principal Findings

In this work, we developed a fully immersive VR scenario to emulate a real experience of a morally distressing situation by HCWs in a simulated ICU setting during the COVID-19 pandemic and assess its acute effects on physiological and psychological parameters as well as longer-term effects on MD. This was followed by an educational video on MD and MI and appropriate mitigation strategies for MD and finally a repetition of the VR scenario in a pretest-posttest design. Because of COVID-19 constraints that resulted in health care settings often being described as a *war zone* [42], HCWs have been particularly exposed to PMIEs in their work environment during this pandemic [4,7]. However, despite the attention it has gained over the last decade, the concept of MI remains poorly understood. VR is a promising strategy to investigate MI owing to its ability to provide highly controlled virtual environments, personalized and tailored experiences, and full control and monitoring of the participants by the research team. The VR scenario created by the research team involved a complex ethical problem that became unfortunately frequent owing to the strain of the pandemic: prioritizing which patients would receive vital support in the face of the shortage of essential equipment such as ventilators [6]. This situation may be considered morally distressing because participants may witness the transgression of some of their core moral values [12], but it is not considered severe enough to induce MI. To achieve our goals, we performed a thorough quantitative and qualitative analysis of the acceptability, easiness of use, tolerability, and utility of the VR technology using a head-mounted display. To the best of our knowledge, this study is the first to examine the feasibility of using an immersive VR scenario to investigate the psychobiological impacts of a moral stressor on HCWs, as well as to use physiological parameters to predict the severity of stress and symptoms of MD and MI.

The feasibility analysis showed high acceptability of the VR scenario among participants, with no dropouts occurring during the study. Although only one-fifth of the participants (3/15, 20%) had previously used VR, all participants reported that the VR technology was easy to use. Moreover, the tolerability was also high because only 1 (7%) of the 15 participants reported mild transient side effects (claustrophobia); no participants reported nausea, whereas other specific side effects (eg, headache and dizziness) were neither reported by participants nor inquired on by the research team. This finding aligns with the literature showing that the incidence rate of VR-induced side effects is low and ranges between 0.5% and 8% [43], with the most common side effects being nausea, eye strain, and dizziness [43]. Specifically, nausea is reported to have an incidence rate of 5.2% [44], whereas vomiting is considered a rare event with an incidence rate of approximately 2% [45]. These symptoms are defined as cybersickness, a form of motion sickness that may be experienced during immersive VR experiences [44]. In this study, we hypothesize that the lack of nausea and other symptoms of cybersickness may have been due to limited head motion during the VR scenario and to the relatively reduced duration (mean 26.3, SD 2.7, min) of the experiment [46].

Regarding the technical quality of our VR scenario, the IPQ results revealed that the scenario achieved a high degree of general presence and spatial presence, above-average involvement, and slightly below-average realism. Therefore, most of the participants felt immersed and involved in the virtual environment but reported that the experiment was not realistic enough (10/15, 67%). This lack of realism was corroborated by the content analysis, where only approximately one-third of the participants (5/14, 36%) felt that the scenario was immersive, real, or engaging. To improve the experience of realism in virtual hospital environments, future studies could address the limitations pointed out by participants in the qualitative debrief session, such as more realistic ICU settings with equipment displaying patients' vital parameters and having ethnically diverse virtual characters to be more representative virtual avatars of participants.

Content analysis of the debriefing revealed that feelings of guilt, shame, betrayal, isolation, and failure were commonly reported; these are impairing moral emotions consistent with MD [7,17,47] and might suggest a violation of moral beliefs. This finding suggests that the VR scenario could acutely induce real experiences of mild MD. Interestingly, numbness was mentioned by approximately a third of the participants (5/14, 36%). This feeling could be considered as a consequence of not having real power in relation to a real-world experience; it may also represent an emotional consequence of being exposed to a PMIE [12,18]. We assume that numbness could be related to potential signs of the erosion of moral agency, not in relation to our intervention but to previous real-world experiences of prolonged and repeated stressors and moral stressors. The content analysis revealed that most of the participants (8/14, 57%) reported guilt and shame, which are feelings consistently related to the experience of MD [7,17]. This finding suggests that the moral stressor experienced during the VR scenario could successfully induce some degree of MD. In addition, half of the participants (7/14, 50%) expressed the need for organizational support, an aspect frequently related to MD. Participants suggested that there could be a greater emphasis on organizational dimensions in future simulations, given the expressed need and the alignment with past research on MD [48]. The findings from the content analysis supported our hypothesis that a VR scenario can be successfully used to elicit and discuss real-life experiences and emotions related to MD.

In contrast to the qualitative results, the quantitative analysis did not show significant changes in the MIOS scores between before and after the experiment. The PSS scores showed the same trend and were not significantly different from baseline, which contradicts our hypothesis that the VR scenario would significantly increase stress levels. Both the MIOS and the PSS focus on symptoms developed over the last month. Although participants were instructed to rate their symptoms at that specific moment, these scales might not have enough sensitivity to capture acute changes in stress and MD symptoms. Alternatively, the changes in MD symptoms may have not been severe enough to induce significant changes in the MIOS scores acutely. Combining our findings from the qualitative and quantitative analyses, we assumed that some degree of MD was experienced by most participants, but we believe that these symptoms were not severe enough to induce MI. This is an important ethical aspect because the VR scenario was designed by specialists in MD and MI to minimize the risk of inducing significant MD in participants.

As MI may develop in the long term, we additionally attempted to use a mobile app to monitor participants for stress and MD and offer psychological support during an 8-week follow-up. Unfortunately, a longitudinal analysis of MD during the follow-up was not possible owing to very low app compliance. It is possible that participants might have developed additional symptoms of MD during follow-up that otherwise could not be captured by our analysis. However, we believe that this is unlikely because no participants requested the psychological support offered in the study. Alternatively, the brief version of the MIOS might not have been sensitive enough to detect slight but important changes in MD that would otherwise be detected by its complete version or by another MD scale. Having said that, this study is a feasibility study with a small sample size, and such an implication is beyond the scope of this work. Finally, the MIOS is still under development; hence, future studies are needed to assess the validity of the MIOS and its brief version.

Mobile app retention proved to be challenging because more than half of the participants (8/15, 53%) did not use the study app, and less than one-third (4/15, 27%) completed at least 1 set of surveys. Our app engagement strategy was based solely on in-app automated reminders and was insufficient to promote participant retention. This finding is supported by recent literature that recommends a combination of different engagement strategies to optimize app use [49,50]. In addition, another possible explanation for the low compliance is that a user-centered design process was not adopted during app development; therefore, the study app may not be particularly targeted to HCWs as the end users [51,52]. Nevertheless, our results are in line with previous research that demonstrates that retention is frequently a great challenge in mobile health studies in both clinical and nonclinical samples [50,53].

Post hoc sample size calculations indicate that a 3-fold and 6-fold sample size is required to reach a power of 80% for the MIOS and the PSS, respectively. With a sample of only 15 participants, our results were underpowered, which may at least in theory explain the nonsignificance of our quantitative findings and the discrepancy between the qualitative and quantitative results. This study was developed during a critical period of the COVID-19 pandemic, with recruitment occurring between May 2021 and August 2021, when contact restrictions were very strict. As the VR intervention required in-person data collection, recruitment proved to be very challenging. Nevertheless, our sample size of 15 participants is appropriate for a preliminary analysis, considering previous VR studies published in PTSD and other mental health disorders [54-57]. Our post hoc sample size calculations may be useful to guide the design of future adequately powered studies using VR in the context of MD and MI.

Limitations

This study has several limitations that must be considered. First, it is a pilot feasibility study with a single arm and a small sample size; thus, the results should be interpreted with due caution. Additional studies with a controlled design are necessary to assess the safety and effectiveness of VR interventions in the assessment of MD and MI. Second, stratification analysis by demographic variables was not possible owing to the reduced sample size; therefore, we were unable to compare symptoms of MI among different subpopulations (eg, nurses and physicians). In addition, our experiments were performed on a purposive sample of only HCWs, thus limiting the generalizability of our findings to other populations. Third, the debriefing methodology used may have also provided a different lens than a traditional qualitative interview or focus group. Fourth, the MIOS and the PSS were used outside of their time frame scope; additional studies should include assessments that focus on acute symptoms of stress and MD. Fifth, a standardized cybersickness scale to assess the side effects within the VR scenario, such as the Virtual Reality Sickness Questionnaire [58], was not used and might have caused underreporting of side effects in this study. Sixth and last, the low app engagement found during the 8-week follow-up hindered an analysis of any potential long-term consequences of the experiment related to MD. Considering that the symptoms of MI may have a late onset, this represents an important limitation to our findings.

Conclusions

The COVID-19 pandemic has challenged the mental health of HCWs, with increased rates of distress, anxiety, and depression being reported. During patient care, ethically difficult situations became common and put frontline HCWs at risk of MD and MI. VR-based interventions are a promising method to address these limitations because they allow for the possibility of developing experiments in safe, personalized, and highly controlled environments. This pilot study investigated the feasibility of using a VR scenario to simulate the experience of a mild morally challenging event for HCWs during the COVID-19 pandemic and to examine participants' physiological reactions to making morally difficult decisions in a virtual environment. Our results suggest the feasibility of using a VR scenario to simulate real experiences of morally stressful events and elicit genuine responses associated with MD with high acceptability and tolerability. In addition, our VR-based intervention demonstrated utility as a pedagogical tool for teaching possible ways to prevent and mitigate MD. Future studies should be conducted to further validate our findings in a larger sample.

Acknowledgments

This work was funded by Innovation for Defence Excellence and Security (IDEaS), Competitive Projects, Department of National Defence, Canada. The authors are grateful to Dr Deborah Kenny, Ms Kristen Sampson, and the Unity Health Toronto Simulation Program for their contribution and support. The financial support of the Ontario Trillium Scholarship program is gratefully acknowledged by AT.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

The study was conceptualized by VB along with SK, EP, and AD. CWE was involved in the interpretation of results, manuscript write-up, and revision. BN and AT were involved in software design of the virtual reality scenario, data collection, data analysis, and the writing of methods and quantitative results. WS was involved in data collection and logistics. AR performed quantitative data analysis. EP was involved in scenario development and qualitative data analysis. LB and DMC provided support with logistics and the debrief component. HJ conducted statistical analysis under the supervision of WL. BK and AD supervised AT for the creation of the virtual reality scenario. SK and VB supervised CWE, BN, and AT on all their tasks.

Conflicts of Interest

VB is supported by an Academic Scholar Award from the Department of Psychiatry, University of Toronto, and has received research support from the Canadian Institutes of Health Research, the Brain & Behavior Research Foundation, Ministry of Health Innovation Funds, the Royal College of Physicians and Surgeons of Canada, the Department of Defence (Canada), and an investigator-initiated trial from Roche Canada. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Postintervention debrief interview guide.
[DOCX File , 11 KB - games_v12i1e42813_app1.docx]

Multimedia Appendix 2
Post hoc sample size calculations.
[DOCX File , 8 KB - games_v12i1e42813_app2.docx]

Multimedia Appendix 3
Individual summary of the most common themes in the content analysis of data of 14 participants.
[DOCX File , 12 KB - games_v12i1e42813_app3.docx]

Multimedia Appendix 4
Results of the virtual reality scenario feasibility questions.
[DOCX File , 9 KB - games_v12i1e42813_app4.docx]

Multimedia Appendix 5
Scores from the debrief feasibility questionnaire.
[DOCX File , 9 KB - games_v12i1e42813_app5.docx]

References

1. Greenberg N. Mental health of health-care workers in the COVID-19 era. Nat Rev Nephrol 2020 Aug 19;16(8):425-426 [FREE Full text] [doi: 10.1038/s41581-020-0314-5] [Medline: 32561871]
2. Travers M. COVID-19 poses a major quandary for the field of moral psychology. Forbes. URL: https://www.forbes.com/sites/traversmark/2020/04/15/covid-19-poses-a-major-quandary-for-the-field-of-moral-psychology/ [accessed 2021-06-21]
3. Asken MJ. Now it is moral injury: the COVID-19 pandemic and moral distress. Medical Economics. URL: https://www.medicaleconomics.com/view/now-it-moral-injury-covid-19-pandemic-and-moral-distress [accessed 2021-06-21]
4. Litam SD, Balkin RS. Moral injury in health-care workers during COVID-19 pandemic. Traumatol 2021 Mar;27(1):14-19. [doi: 10.1037/trm0000290]

5. Rushton CH, Thomas TA, Antonsdottir IM, Nelson KE, Boyce D, Vioral A, et al. Moral injury and moral resilience in health care workers during COVID-19 pandemic. J Palliat Med 2022 May 01;25(5):712-719 [FREE Full text] [doi: 10.1089/jpm.2021.0076] [Medline: 34678091]
6. Han JJ, Luc JG, Pak E. Ethical dilemmas associated with the COVID-19 pandemic: dealing with the unknowns and unanswerables during training. J Am Coll Cardiol 2020 Sep 08;76(10):1266-1269 [FREE Full text] [doi: 10.1016/j.jacc.2020.07.041] [Medline: 32883420]
7. Čartolovni A, Stolt M, Scott PA, Suhonen R. Moral injury in healthcare professionals: a scoping review and discussion. Nurs Ethics 2021 Aug 11;28(5):590-602 [FREE Full text] [doi: 10.1177/0969733020966776] [Medline: 33427020]
8. Williams RD, Brundage JA, Williams EB. Moral injury in times of COVID-19. J Health Serv Psychol 2020 May 02;46(2):65-69 [FREE Full text] [doi: 10.1007/s42843-020-00011-4] [Medline: 32363349]
9. Jameton A. Nursing Practice: The Ethical Issues. Upper Saddle River, NJ: Prentice-Hall; 1984.
10. Shay J. Achilles in Vietnam: Combat Trauma and The Undoing of CharacterDisabil Rehabil. Southern Pines, NC: Scribner Publishing; 2010.
11. Shay J. Casualties. Daedalus 2011 Jul;140(3):179-188. [doi: 10.1162/daed_a_00107] [Medline: 21898967]
12. Litz BT, Stein N, Delaney E, Lebowitz L, Nash WP, Silva C, et al. Moral injury and moral repair in war veterans: a preliminary model and intervention strategy. Clin Psychol Rev 2009 Dec;29(8):695-706. [doi: 10.1016/j.cpr.2009.07.003] [Medline: 19683376]
13. Neria Y, Pickover A. Commentary on the special issue on moral injury: advances, gaps in literature, and future directions. J Trauma Stress 2019 Jun 04;32(3):459-464. [doi: 10.1002/jts.22402] [Medline: 31162730]
14. Dean W, Jacobs B, Manfredi RA. Moral injury: the invisible epidemic in COVID health care workers. Ann Emerg Med 2020 Oct;76(4):385-386 [FREE Full text] [doi: 10.1016/j.annemergmed.2020.05.023] [Medline: 32586634]
15. Barnes HA, Hurley RA, Taber KH. Moral injury and PTSD: often co-occurring yet mechanistically different. J Neuropsychiatry Clin Neurosci 2019;31(2):A4-103. [doi: 10.1176/appi.neuropsych.19020036] [Medline: 31012825]
16. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. New York, NY: American Psychiatric Association; 2013.
17. Griffin BJ, Purcell N, Burkman K, Litz BT, Bryan CJ, Schmitz M, et al. Moral injury: an integrative review. J Trauma Stress 2019 Jun 28;32(3):350-362. [doi: 10.1002/jts.22362] [Medline: 30688367]
18. Yeterian JD, Berke DS, Carney JR, McIntyre-Smith A, St Cyr K, King L, Members of the Moral Injury Outcomes Project Consortium. Defining and measuring moral injury: rationale, design, and preliminary findings from the moral injury outcome scale consortium. J Trauma Stress 2019 Jun 04;32(3):363-372. [doi: 10.1002/jts.22380] [Medline: 30947372]
19. Ramage AE, Litz BT, Resick PA, Woolsey MD, Dondanville KA, Young-McCaughan S, STRONG STAR Consortium. Regional cerebral glucose metabolism differentiates danger- and non-danger-based traumas in post-traumatic stress disorder. Soc Cogn Affect Neurosci 2016 Feb;11(2):234-242 [FREE Full text] [doi: 10.1093/scan/nsv102] [Medline: 26373348]
20. Lambert S. Role of emotional intelligence in effective nurse leadership. Nurs Stand 2021 Dec 01;36(12):45-49. [doi: 10.7748/ns.2021.e11782] [Medline: 34514759]
21. Zerach G, Levi-Belz Y. Moral injury and mental health outcomes among Israeli health and social care workers during the COVID-19 pandemic: a latent class analysis approach. Eur J Psychotraumatol 2021;12(1):1945749 [FREE Full text] [doi: 10.1080/20008198.2021.1945749] [Medline: 34367526]
22. Mantri S, Lawson JM, Wang Z, Koenig HG. Identifying moral injury in healthcare professionals: the moral injury symptom scale-HP. J Relig Health 2020 Oct 17;59(5):2323-2340 [FREE Full text] [doi: 10.1007/s10943-020-01065-w] [Medline: 32681398]
23. Bell IH, Nicholas J, Alvarez-Jimenez M, Thompson A, Valmaggia L. Virtual reality as a clinical tool in mental health research and practice . Dialogues Clin Neurosci 2020 Jun;22(2):169-177 [FREE Full text] [doi: 10.31887/DCNS.2020.22.2/lvalmaggia] [Medline: 32699517]
24. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400 [FREE Full text] [doi: 10.1017/S003329171700040X] [Medline: 28325167]
25. Gonçalves R, Pedrozo AL, Coutinho ES, Figueira I, Ventura P. Efficacy of virtual reality exposure therapy in the treatment of PTSD: a systematic review. PLoS One 2012;7(12):e48469 [FREE Full text] [doi: 10.1371/journal.pone.0048469] [Medline: 23300515]
26. Rizzo A, Shilling R. Clinical Virtual Reality tools to advance the prevention, assessment, and treatment of PTSD. Eur J Psychotraumatol 2017;8(sup5):1414560 [FREE Full text] [doi: 10.1080/20008198.2017.1414560] [Medline: 29372007]
27. Deng W, Hu D, Xu S, Liu X, Zhao J, Chen Q, et al. The efficacy of virtual reality exposure therapy for PTSD symptoms: a systematic review and meta-analysis. J Affect Disord 2019 Oct 01;257:698-709. [doi: 10.1016/j.jad.2019.07.086] [Medline: 31382122]
28. Nguyen B, Torres A, Sim W, Kenny D, Campbell DM, Beavers L, et al. Digital interventions to reduce distress among health care providers at the frontline: protocol for a feasibility trial. JMIR Res Protoc 2022 Feb 16;11(2):e32240 [FREE Full text] [doi: 10.2196/32240] [Medline: 34871178]
29. A Guide to Moral Injury. Centre of Excellence on PTSD. 2020. URL: http://moralinjuryguide.ca/ [accessed 2022-06-29]

30. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385. [doi: 10.2307/2136404]
31. igroup presence questionnaire (IPQ) overview. igroup. URL: http://www.igroup.org/pq/ipq/index.php [accessed 2022-03-02]
32. Litz B, Plouffe RA, Nazarov A, Murphy D, Phelps A, Coady A, Moral Injury Outcome Scale Consortium. Defining and assessing the syndrome of moral injury: initial findings of the moral injury outcome scale consortium. Front Psychiatry 2022;13:923928 [FREE Full text] [doi: 10.3389/fpsyt.2022.923928] [Medline: 35873252]
33. Digital therapeutics. St. Michael's Hospital. 2022. URL: https://stmichaelshospitalresearch.ca/digital-therapeutics [accessed 2022-03-11]
34. igroup presence questionnaire (IPQ) scale construction. igroup. URL: http://www.igroup.org/pq/ipq/construction.php [accessed 2022-03-02]
35. igroup presence questionnaire (IPQ) factor analysis. igroup. URL: http://www.igroup.org/pq/ipq/factor.php [accessed 2022-03-02]
36. White MD, Marsh EE. Content analysis: a flexible methodology. Libr Trends 2006;55(1):22-45. [doi: 10.1353/lib.2006.0053]
37. Mayring P. Qualitative content analysis. Forum Qual Sozialforschung 2007 Oct 15;55(6):251-257 [FREE Full text] [doi: 10.17169/FQS-1.2.1089]
38. Bajaj K, Meguerdichian M, Thoma B, Huang S, Eppich W, Cheng A. The PEARLS healthcare debriefing tool. Acad Med 2018 Feb;93(2):336. [doi: 10.1097/ACM.0000000000002035] [Medline: 29381495]
39. Orkin AM, Charles M, Norris K, Thomas R, Chapman L, Wright A, et al. Mixed methods feasibility study for the surviving opioid overdose with naloxone education and resuscitation (SOONER) trial. Resusc Plus 2021 Jun;6:100131 [FREE Full text] [doi: 10.1016/j.resplu.2021.100131] [Medline: 34223388]
40. González-Franco M, Pérez-Marcos D, Spanlang B, Slater M. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In: Proceedings of the 2010 IEEE Virtual Reality Conference. 2010 Presented at: VR '10; March 20-26, 2010; Boston, MA p. 111-114 URL: https://ieeexplore.ieee.org/document/5444805 [doi: 10.1109/vr.2010.5444805]
41. Forchuk CA, Nazarov A, Plouffe RA, Liu JJ, Deda E, Le T, et al. Well-being of Canadian armed forces veterans and spouses of veterans during the COVID-19 pandemic: protocol for a prospective longitudinal survey. JMIR Res Protoc 2022 Jan 11;11(1):e34984 [FREE Full text] [doi: 10.2196/34984] [Medline: 34935624]
42. Gooch K. "War zones": healthcare workers share front-line accounts of fighting COVID-19. Becker's Hospital Review. URL: https://www.beckershospitalreview.com/workforce/war-zones-healthcare-workers-share-front-line-accounts-of-fighting-covid-19.html [accessed 2022-06-29]
43. Smith V, Warty RR, Sursas JA, Payne O, Nair A, Krishnan S, et al. The effectiveness of virtual reality in managing acute pain and anxiety for medical inpatients: systematic review. J Med Internet Res 2020 Nov 02;22(11):e17980 [FREE Full text] [doi: 10.2196/17980] [Medline: 33136055]
44. Gold JI, Mahrer NE. Is virtual reality ready for prime time in the medical space? A randomized control trial of pediatric virtual reality for acute procedural pain management. J Pediatr Psychol 2018 Apr 01;43(3):266-275. [doi: 10.1093/jpepsy/jsx129] [Medline: 29053848]
45. Mazloumi Gavgani A, Walker FR, Hodgson DM, Nalivaiko E. A comparative study of cybersickness during exposure to virtual reality and "classic" motion sickness: are they different? J Appl Physiol (1985) 2018 Dec 01;125(6):1670-1680 [FREE Full text] [doi: 10.1152/japplphysiol.00338.2018] [Medline: 30284516]
46. Davis S, Nesbitt K, Nalivaiko E. A Systematic Review of Cybersickness. In: Proceedings of the 2014 Conference on Interactive Entertainment. 2014 Presented at: IE '14; December 2-3, 2014; Newcastle, Australia p. 1-9. [doi: 10.1145/2677758.2677780]
47. Jinkerson JD. Defining and assessing moral injury: a syndrome perspective. Traumatol 2016 Jun;22(2):122-130. [doi: 10.1037/trm0000069]
48. Nelson KE, Hanson GC, Boyce D, Ley CD, Swavely D, Reina M, et al. Organizational impact on healthcare workers' moral injury during COVID-19: a mixed-methods analysis. J Nurs Adm 2022 Jan 01;52(1):57-66 [FREE Full text] [doi: 10.1097/NNA.0000000000001103] [Medline: 34910709]
49. Cao W, Milks MW, Liu X, Gregory ME, Addison D, Zhang P, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. JMIR Mhealth Uhealth 2022 Mar 02;10(3):e29415 [FREE Full text] [doi: 10.2196/29415] [Medline: 35234655]
50. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. J Med Internet Res 2022 Apr 26;24(4):e35120 [FREE Full text] [doi: 10.2196/35120] [Medline: 35471414]
51. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS One 2020 Aug 19;15(8):e0237910 [FREE Full text] [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
52. Saparamadu AA, Fernando P, Zeng P, Teo H, Goh A, Lee JM, et al. User-centered design process of an mHealth app for health professionals: case study. JMIR Mhealth Uhealth 2021 Mar 26;9(3):e18079 [FREE Full text] [doi: 10.2196/18079] [Medline: 33769297]

53. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
54. Jetly CR, Meakin LC, Sinitski EH, Blackburn L, Menard J, Vincent M, et al. Multi-Modal virtual-reality based treatment for members with combat related posttraumatic stress disorder: Canadian Armed Forces pilot study. In: Proceedings of the 2017 International Conference on Virtual Rehabilitation. 2017 Presented at: ICVR '17; May 19-22, 2017; Montreal, QC p. 1-2 URL: https://ieeexplore.ieee.org/document/8007474 [doi: 10.1109/icvr.2017.8007474]
55. Gamito P, Oliveira J, Rosa P, Morais D, Duarte N, Oliveira S, et al. PTSD elderly war veterans: a clinical controlled pilot study. Cyberpsychol Behav Soc Netw 2010 Feb;13(1):43-48. [doi: 10.1089/cyber.2009.0237] [Medline: 20528292]
56. Kaussner Y, Kuraszkiewicz AM, Schoch S, Markel P, Hoffmann S, Baur-Streubel R, et al. Treating patients with driving phobia by virtual reality exposure therapy - a pilot study. PLoS One 2020 Jan 7;15(1):e0226937 [FREE Full text] [doi: 10.1371/journal.pone.0226937] [Medline: 31910205]
57. van 't Wout-Frank M, Shea MT, Larson VC, Greenberg BD, Philip NS. Combined transcranial direct current stimulation with virtual reality exposure for posttraumatic stress disorder: feasibility and pilot results. Brain Stimul 2019 Jan;12(1):41-43 [FREE Full text] [doi: 10.1016/j.brs.2018.09.011] [Medline: 30266416]
58. Kim HK, Park J, Choi Y, Choe M. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl Ergon 2018 May;69:66-73. [doi: 10.1016/j.apergo.2017.12.016] [Medline: 29477332]

Abbreviations

ECG: electrocardiography
HCW: health care worker
ICU: intensive care unit
IPQ: Igroup Presence Questionnaire
MD: moral distress
MI: moral injury
MIOS: Moral Injury Outcome Scale
PEARLS: Promoting Excellence and Reflective Learning in Simulation
PMIE: potentially morally injurious event
PSS: Perceived Stress Scale
PTSD: posttraumatic stress disorder
VR: virtual reality

Edited by G Eysenbach, T Leung; submitted 21.09.22; peer-reviewed by A Garavand, M Attridge, M Cozad, R Weibel; comments to author 14.04.23; revised version received 03.08.23; accepted 30.09.23; published 09.01.24.

Please cite as:
Espinola CW, Nguyen B, Torres A, Sim W, Rueda A, Beavers L, Campbell DM, Jung H, Lou W, Kapralos B, Peter E, Dubrowski A, Krishnan S, Bhat V
Digital Interventions for Stress Among Frontline Health Care Workers: Results From a Pilot Feasibility Cohort Trial
JMIR Serious Games 2024;12:e42813
URL: https://games.jmir.org/2024/1/e42813
doi:10.2196/42813
PMID:38194247

Original Paper

A Serious Game to Train Rhythmic Abilities in Children With Dyslexia: Feasibility and Usability Study

Francois Vonthron[1], MSc; Antoine Yuen[1], MSc; Hugues Pellerin[2], MSc; David Cohen[2,3], MD, Prof Dr; Charline Grossard[2,3], PhD

[1]Poppins, Palaiseau, France

[2]Service de Psychiatrie de l'Enfant et de l'Adolescent, Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique–Hôpitaux de Paris, Paris, France

[3]Institut des Systèmes Intelligents et Robotiques (ISIR, CNRS UMR7222), Sorbonne Université, Paris, France

Corresponding Author:
Francois Vonthron, MSc
Poppins
73 rue Leon Bourgeois
Palaiseau, 91120
France
Phone: 33 669515961
Email: francois.vonthron@gmail.com

Abstract

Background: Rhythm perception and production are related to phonological awareness and reading performance, and rhythmic deficits have been reported in dyslexia. In addition, rhythm-based interventions can improve cognitive function, and there is consistent evidence suggesting that they are an efficient tool for training reading skills in dyslexia.

Objective: This paper describes a rhythmic training protocol for children with dyslexia provided through a serious game (SG) called Mila-Learn and the methodology used to test its usability.

Methods: We computed Mila-Learn, an SG that makes training remotely accessible and consistently reproducible and follows an educative agenda using Unity (Unity Technologies). The SG's development was informed by 2 studies conducted during the French COVID-19 lockdowns. Study 1 was a feasibility study evaluating the autonomous use of Mila-Learn with 2500 children with reading deficits. Data were analyzed from a subsample of 525 children who spontaneously played at least 15 (median 42) games. Study 2, following the same real-life setting as study 1, evaluated the usability of an enhanced version of Mila-Learn over 6 months in a sample of 3337 children. The analysis was carried out in 98 children with available diagnoses.

Results: Benefiting from study 1 feedback, we improved Mila-Learn to enhance motivation and learning by adding specific features, including customization, storylines, humor, and increasing difficulty. Linear mixed models showed that performance improved over time. The scores were better for older children ($P<.001$), children with attention-deficit/hyperactivity disorder ($P<.001$), and children with dyslexia ($P<.001$). Performance improved significantly faster in children with attention-deficit/hyperactivity disorder ($\beta=.06$; $t_{3754}=3.91$; $P<.001$) and slower in children with dyslexia ($\beta=-.06$; $t_{3816}=-5.08$; $P<.001$).

Conclusions: Given these encouraging results, future work will focus on the clinical evaluation of Mila-Learn through a large double-blind randomized controlled trial comparing Mila-Learn and a placebo game.

(JMIR Serious Games 2024;12:e42733) doi:10.2196/42733

KEYWORDS

serious game; rhythm; dyslexia; musical abilities; design framework; reading skills; children; digital health intervention; attention-deficit/hyperactivity disorder; ADHD; child development; mobile phone

Introduction

Background

Music training and music-based interventions are becoming increasingly popular for developing brain and cognitive functions in children [1-5]. Building on brain plasticity induced by learning music and the tight link between musical and cognitive skills [6-8], music interventions have been used as training tools in neurodevelopmental disorders (NDDs) such as dyslexia [9-13]. Musical skills, especially when developed in childhood, are associated with enhanced cognitive abilities in various domains, such as attention, processing speed [3], executive functions [14], or speech and language [15-17]. Improvements in cognitive skills induced by musical training have been attributed to structural and functional brain changes in areas that support both music processing and cognition [6-8,18,19].

Recent studies have focused specifically on the relationship between rhythmic skills, such as the capacity to discriminate musical rhythms or synchronize with a beat [20,21], and cognition during development. Tierney and Kraus [22] showed that correlations exist between synchronization with a metronome and attentional and reading skills in typically developing adolescents. In children, rhythm production accuracy is associated with both phonological awareness and reading [23]. Rhythm perception is also related to reading performance [24,25]. Language and music processing may rely on common timing mechanisms that allow for the extraction of temporal information, which is crucial to accurately perceive sequences of events [7,20,26-28]. This hypothesis is supported by neurofunctional evidence as temporal processing involved in music and language recruits partially overlapping neuronal pathways that include the auditory cortex, dorsal premotor cortex, cerebellum, basal ganglia, and thalamus [29,30].

Further evidence of the link between rhythmic skills and cognitive abilities comes from the observation that rhythmic skills are disrupted in NDDs that also affect cognition. Notably, rhythmic deficits have been extensively reported in individuals with dyslexia. Children and adults with dyslexia exhibit inaccurate rhythm perception [25,31] as well as increased variability in motor tapping tasks [32]. These observed rhythmic deficits have given rise to theories (eg, the temporal sampling framework; Goswami [26]) that postulate that poor predictive temporal sampling and coding of events explain reading difficulties in those with dyslexia [26,33].

Building on the importance of rhythmic skills in development, music-based training protocols for children have been developed in recent decades. Studies have shown that children with dyslexia who participate in music-based interventions display better reading and phonological abilities [10-12]. In addition, the effect of music-based programs was extended to typically developing children, who showed significant improvements in speech processing skills and verbal intelligence [14]. However, these encouraging preliminary data have not reached the recommended quality for evidence-based studies owing to methodological limitations such as limited sample size, lack of blind assessment, and potentially inconsistent delivery of interventions [34]. In addition, access to these interventions is still too limited, with inequalities remaining because of significant disparities according to social background and place of residence [35]. For instance, children in poor and remote urban areas, who are more likely to develop an NDD [36], often have less access to care. Furthermore, these traditional music-based interventions usually require in-person instruction, which can be challenging under certain circumstances such as during the COVID-19 pandemic or in areas with limited access to specialized resources. More research is needed to determine whether written language skills can improve in children with dyslexia after training with more accessible and scalable music-based interventions.

To address these limitations, serious games (SGs) designed for educational and training purposes provide a more standardized, scalable, and accessible format for delivering music-based interventions through information and communications technologies. This approach allows for the delivery of the same training to a large sample regardless of geographic location or in-person resource availability. The number of SGs developed for educational and training purposes has increased over the last decade [37], primarily because of the expansion of information and communications technologies such as mobile technologies and telehealth systems. As most households, including those in low-income brackets and rural areas, are now equipped with at least 1 tablet, smartphone, or computer, these SGs can be broadly accessible [35]. Furthermore, a meta-analysis revealed that, across domains, learning is improved with SGs compared with conventional methods [38]. In addition to motivation, several preliminary findings have supported another exciting alternative hypothesis that playing an SG fosters electrical brain activity related to memory, emotions, and concentration [39], providing a possible neuronal explanation for the beneficial effect of SGs. SGs have been used in typically developing populations [40] and in children with NDDs [41,42]. Notably, SGs have been used to deliver rhythm-based training to healthy young adults [43]. Recently, interest in using computer-based interventions to train rhythm skills has been explored in people with dyslexia [44]. One SG named “Jellys” was developed for this purpose in a usability study and showed that children with dyslexia positively engaged with this type of remediation [45]. However, although some studies seem to support the effectiveness of using SGs as a treatment for people with NDDs, the methodological quality of these studies is limited, and further research is needed [46].

Objectives

In this study, our goal was to evaluate the usability of Mila-Learn, an SG aimed at training rhythmic abilities in children with dyslexia. The methodological design of the game was developed in user participatory pilot studies, allowing the children and their families to provide feedback to shape human-machine interactions. We report on 2 studies conducted during the French COVID-19 lockdowns. The first was a feasibility study to assess the children’s engagement through gameplay frequency and collected their feedback. After modifying the SG according to the feedback by adding specific features such as customization, storyline, humor, or increasing difficulty, we report a usability study that addressed the children's performance on the latest version of the SG when played autonomously at home and according to declared diagnoses.

Methods

Overview

Mila-Learn is an SG that delivers rhythm-based exercise designed for children with dyslexia (called "specific learning disorders in the field of reading" in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*). This SG involves a rich musical universe aiming to lead the child to spontaneously come back and engage with the instrument with their parents. It consists of two main elements: (1) a mobile app that offers rhythmic, sensory-motor, and cognitive tasks in the form of musical activities; and (2) secure servers that allow for data analysis. It enables real-time evaluations to understand children's difficulties and improve the effectiveness of Mila-Learn. In this section, we first describe the SG from its initial beta version. We then detail the methodology of the 2 exploratory studies conducted during the COVID-19 lockdown.

Mila-Learn Description

From an Initial Prototype (2018) as a Progressive Web Application to a Version (2019-2020) Developed for Tablets

The first version of Mila-Learn included five tasks:

1. Dance With Your Hands is an auditory-motor coordination exercise. It involves performing a movement following the tempo of a piece of music. These are pieces with a 4/4 signature that easily allow the child to have rhythmic stability and associate a motor action with the rhythm. This action includes gestures such as clapping, silencing, and raising the arms in the air.
2. Play the Drums is a rhythmic memory game. A drum appears on the child's screen, and a sequence is played. The child then presses on the drum elements to reproduce the initially played sequence.
3. Rhythmic Vitamins is an exercise in singing and repetition [47]. An initial recorded vocal sequence consisting of syllables and phonemes is played by the software. The child must reproduce it using the rhythm, pronunciation, and pitch of the initial sequence.
4. Following the Tempo requires recognizing and reproducing different rhythmic structures. In this exercise, the child is asked to mark the strong beats of music using the space key on the keyboard.
5. Musical Pitch is an exercise of association between the pitch and its representation. A sound sequence composed of 3 sounds is played (high, medium, and low), and then a visual representation is displayed composed of lines (low, medium, and high). The child has to judge whether the graphic production of the sound is correct.

This prototype version (Figure 1 [48]) was offered to a small group of children end users with dyslexia. We asked them to provide feedback on the design and players. The data collection method was centered on gathering children's feedback at the end of 15-minute game sessions. A total of 14 children were invited to respond, interact, and provide feedback on the first version of Mila-Learn in the form of a progressive web application [49]. In total, 3 sessions per week over a period of 3 months were conducted.

Figure 1. Landing page of the prototype with 3 different tasks each day [56].

This participatory design phase led to a framework for the development of Mila-Learn enhanced by a literature review on SG playability to increase players' motivation. Although some studies have proposed a framework to develop SGs for people with NDDs, very few have focused specifically on dyslexia [46]. To improve the design of our game, we expanded our research to the use of SGs in the typical population and in children with NDDs, especially children with attention-deficit/hyperactivity disorder (ADHD), because of the high co-occurrence between dyslexia and ADHD [50].

Design Framework

We integrated new gaming features to increase players' motivation and engagement. First, we developed a storyline that allowed us to include the different tasks within a larger story. The use of a storyline supported the engagement of a player in the games [51,52]. The storyline was intended to not be complex to prevent the child from losing the main goal of the game [53]. The story in Mila-Learn was designed to correspond to the interests of children aged between 7 and 14 years. The story is inspired by *shonen* manga, which is manga inspired by the cartoon universe. This type of manga is based on a storyline that involves a young hero who starts without knowledge and becomes increasingly powerful during the progression of the story. Some crucial values such as friendship and perseverance are typically present in the story. Most of the time, the first opponent of the hero becomes his friend during the story. In Mila-Learn, the player embodies a little monster who meets another character, a little blue monster named "Blue" who asks for their help—some villagers are held captive by the villain Diabolus, another character. The player has to learn rhythm skills to challenge Diabolus and free the villagers. Over the course of the game, the player discovers that there is a larger villain who holds "Rosa," the Diabolus' scooter. At the end, the player must win a large rhythm tournament to finish the game. The story is divided into 12 chapters containing 6 tasks (see the following section). Each task lasts between 1 minute, 20 seconds and 1 minute, 30 seconds, with rare songs playing for 1 minute, 40 seconds to maintain consistency with the music. In this way, we considered the attention capacity of children [53,54].

As recommended in the literature, we created evolving tasks, gradually increasing the level of difficulty in each task and from one task to another [51,53,54]. The tasks must be challenging but accessible. In total, 6 tasks are used in the second version of Mila-Learn. They are introduced progressively to allow the player to practice a task 2 to 3 times before introducing another one. Once all the tasks are known by the player, they increase in difficulty with progression throughout the game. First, within each task, the rhythm displayed corresponds to each beat of a measure. Then, the rhythm changes to correspond to eighth notes (meaning that the rhythm is clapping 3 times in 2 beats) or slows down to be marked only once every 2 beats. Moreover, the songs are played at an increased speed to challenge the player. At higher levels, the marked rhythms can change during the task.

The SG was built to provide clear instructions to the player [51,53]. The instructions are given orally and with visual support, notably by imitating one or more nonplayer characters. Before each task, a quick tutorial allows the player to repeat the movement they have to perform during the game 3 times (ie, clap their hands, touch the screen, and move the tablet). At the beginning of a task, the character played by the user is clearly identified with an arrow. Moreover, for each task, the player is always placed in the same location.

The visual environment is thought to be easily navigated by children. The graphics are pleasant but minimalistic [53,55]. The visuals are thought to be pleasant for children aged between 7 and 14 years and are inspired by the cartoon universe. During the tasks, the background is mostly static, allowing the child to focus on the goal of the task. The characters only move to the rhythm of the music, with repetitive and predictable movements.

We differentiate between short- and long-term goals [51,53,54]. In each task, there is only 1 clear goal (ie, touch the screen to the rhythm) that is clearly differentiated from the long-term goal of a chapter (ie, complete the chapter to challenge Diabolus; Figure 2). Feedback is provided throughout the different tasks using visual cues [51,53,54]. These cues allow the player to know whether they are performing the exercise properly. The feedback for each task is described in the following section. As rewards have been described as a main feature of SGs [51,54], players obtain a reward of 1 to 3 stars at the end of each task depending on their accuracy during the exercise. Personalization has also been described as an important key to enhancing the motivation of the player [51-54]. As in the first version, players have to pick a name for their character at the beginning of the game and modify its color. Finally, we introduced new songs to work on in this version of the game. We added some famous songs known by most children (ie, songs from Disney movies) to increase the motivation of the players. For some tasks such as Fruity Jump, Karate Fruit, and Sing Lab, the predetermined structures of these songs did not make their use possible. We specifically composed songs to fit with the requirements of these tasks.

Figure 2. Example of a chapter menu introducing the upcoming adventure.

Description of the Tasks

All the tasks (Figure 3) were designed to work on rhythm, which was the main and explicit goal of each task. However, each task requires the mobilization of other skills such as attention, inhibition, working memory, and motor skills, which are also often impaired in children with dyslexia [50,56].

Follow me aims to introduce rhythm to the player. The child first sees a little monster clapping hands to the rhythm of a song and then has to touch the screen to the rhythm by imitating the monster. Then, the character stops clapping, and the player has to keep going alone without the support of the monster. This task allows the player to work on maintaining regularity in rhythm but also sustained attention.

In Clap Trap, 2 characters and the player appear on the screen. The first 2 characters clap one after another, giving a tempo to the player, who has to complete the sequence by clapping their hands to the rhythm at the right time. The first character claps on the first beat of a 4-time measure of the song played. The second character claps on the second beat, and the player has to clap their hands on the third beat. The microphone records the child’s clap. In this task, the child has to anticipate and adapt to the rhythm. It was designed to train inhibition skills as the child has to wait until the right moment to clap their hands.

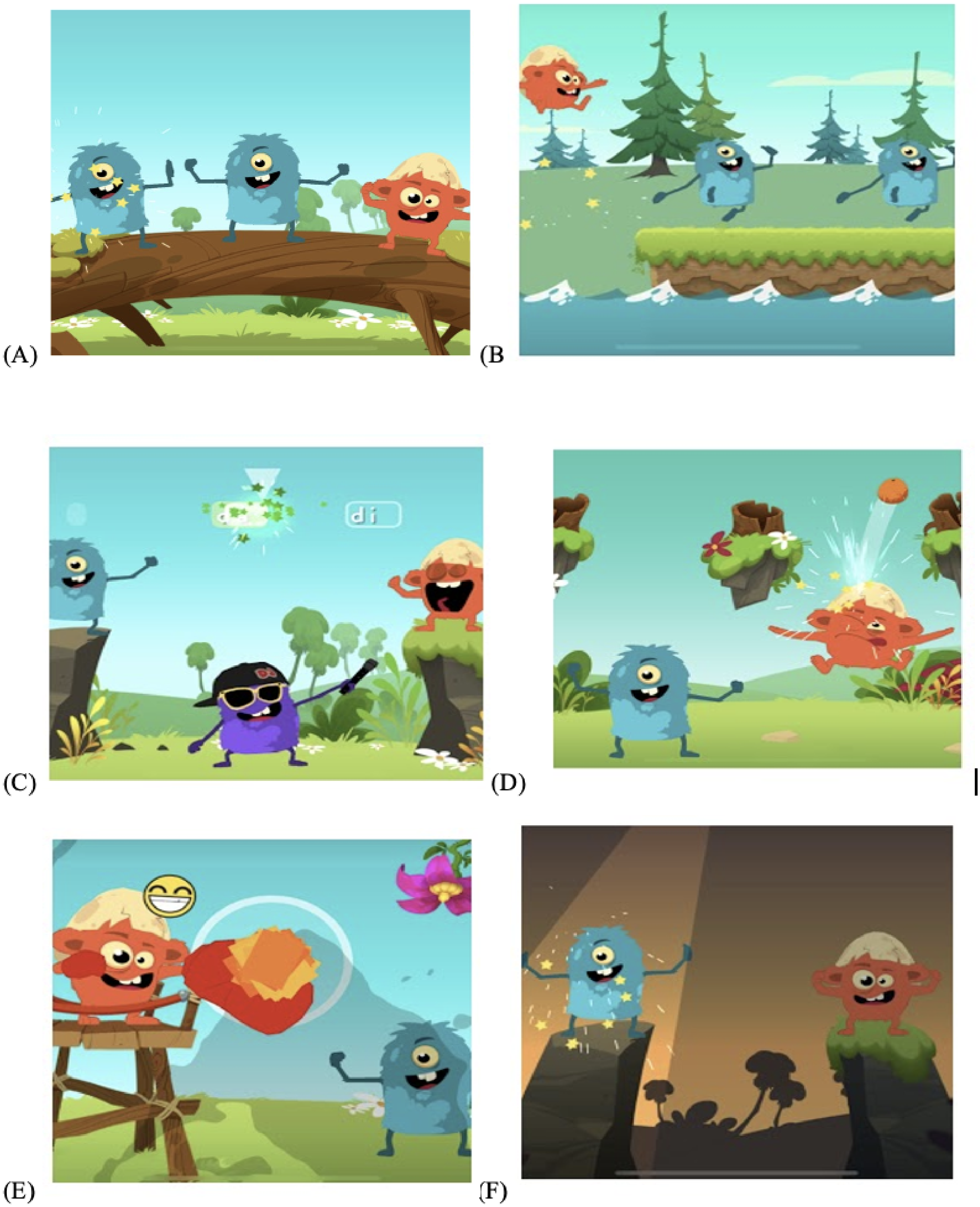
In River Splash, the player is placed behind 2 other characters who run next to the water and sometimes have to jump across the river to the rhythm. The first character jumps on the first beat of a 4-time measure of a song. The second character jumps on the second beat, and the player jumps on the third beat. The player has to shake the tablet quickly to jump. In addition to rhythm perception, this task was designed to train inhibition skills similarly to the Clap Trap task.

In Sing Lab, the first character produces a sequence of phonemes or syllables at a particular tempo. The player has to reproduce this sequence with particular attention to the pattern and duration of the phonemes. Phonemes or syllables pass across the top of the screen, visually represented by gauges that the player has to fill. If the child sings at the right time, the gauge starts filling. When the duration of the note is complete, the gauge changes color from white to green. In this task, the phonological loop is involved in correctly memorizing the sequence. We used specific music constructed for this task that allowed us to add phonemes or syllables to sing at particular moments and for as long as we wanted.

In Fruity Jump, a character reproduces a rhythmic sequence. The player has to memorize this sequence and then reproduce it correctly by tapping the screen at a good tempo. The tempo is visually indicated by fruits falling from a tree. If the player claps at the right time, the character jumps and hits a fruit with its head to throw it to another tree. If the player misses the fruit, it crashes on the ground. If the player jumps at another time (ie, during the demonstration), their character jumps, and nothing special happens. Working memory and intermodality are particularly engaged in this task.

Finally, in Karate Fruits, the player has to hit fruits that appear based on the rhythm. To hit the fruit, the player has to put the tablet on the floor and extend their arms above it. Each time the camera detects the arm, the player’s character punches. If the player punches at the right time, the fruit explodes, and a smiley face appears. If the player misses the fruit, the fruit goes off the screen, and a smiley face with an annoyed head appears. If the player punches at another time, the character punches, and nothing else happens.

Figure 3. Examples of screens in Mila-Learn: (A) Clap Trap, (B) River Splash, (C) Sing Lab, (D) Fruity Jumps, (E) Karate Fruits, and (F) Follow Me.



Scoring Player Performance

Scoring of player performance is based on rhythmic synchronization through multiple modalities of interaction (sometimes in combination) as rhythmic synchronization is a requirement for all games. Player responses are captured through accelerometers, microphones, webcams, and pressure-sensitive screens, as shown in Table 1.

By assessing the audiomotor synchronization of the child with the rhythmic instruction, we define (1) a time T that corresponds to the exact moment when the player's input is expected (regardless of the interaction mode) and (2) tolerance thresholds ($t_{Perfect}<t_{Good}<t_{Correct}$).

The different intervals allow for judging the quality of the answer with 4 levels of acceptance. An input is considered acceptable when it is in the interval [$T - t_{Correct}$; $T + t_{Correct}$] and not acceptable otherwise. An input of better quality, either in the interval [T – t_{Good}; T + t_{Good}] or in the interval [T – $t_{Perfect}$; T + $t_{Perfect}$], results in different visual and audio feedback for the child.

In the second improved version, which was a modified version based on the first pilot study, a simplified calculation was performed by considering the ratio of acceptable inputs to total inputs as the main measure. This final score is presented to the child in the form of stars depending on their performance: no stars if the child has an average of <50%, 1 star if ≥50% of inputs are acceptable, 2 stars for ≥75%, and 3 stars for ≥90%. In addition, this architecture allows for the storage of all the child's inputs for retro-analysis purposes.

Table 1. Players' recorded responses and game parameters in the second version of Mila-Learn.

Task	Type	Interaction	Capture technology	Songs	Tolerance threshold
Follow Me[a]	Continuous tapping	Tapping	Contact pressure	Commercial	• $t_{Perfect}$: 0.1 s before or after the beat • t_{Good}: 0.15 s before or after the beat • $t_{Correct}$: 0.25 s before or after the beat
Clap Trap	Last beat	Clapping hands	Microphone	Commercial	• $t_{Perfect}$: 0.1 s before or after the beat • t_{Good}: 0.15 s before or after the beat • $t_{Correct}$: 0.25 s before or after the beat
River Splash	Last beat	Shaking tablet	Accelerometer	Commercial	• $t_{Perfect}$: 0.1 s before or after the beat • t_{Good}: 0.15 s before or after the beat • $t_{Correct}$: 0.25 s before or after the beat
Sing Lab	Call and response	Singing	Microphone	Commercial+built in-house	• $t_{Pattern}$: 0.15 s before or after the beat and up to 30% of the note duration • Song duration: the note must be sung at least 60% of the time
Fruity Jump	Call and response	Tapping	Contact pressure	Built in-house	• $t_{Perfect}$: 0.2 s before or after the beat • t_{Good}: 0.25 s before or after the beat • $t_{Correct}$: 0.3 s before or after the beat
Karate Fruits	Last beat	Punching	Webcam	Built in-house	• $t_{Perfect}$: 0.08 s before or after the beat • t_{Good}: 0.14 s before or after the beat • $t_{Correct}$: 0.3 s before or after the beat

[a]Name of the task.

Feasibility Study

We conducted a feasibility study to evaluate whether children with NDDs involving reading deficits could use Mila-Learn autonomously at home. Our main objective in assessing Mila-Learn's autonomous use was to monitor both the time users spent on the SG and their accuracy in each game played. In the context of the unprecedented health crisis caused by COVID-19, participants were recruited by the French Federation for Learning Disorders (FFDys), a national association that aggregates all regional associations of people with learning disabilities. The FFDys communicated to its members the possibility of testing an app and managed the information and consent of participants. Families were informed that Mila-Learn was an SG for performing rhythmic tasks at home and that we believed this practice might be beneficial for learning to read. In total, 2500 children downloaded Mila-Learn. The analyses were conducted on a subsample of 21% (525/2500) of these children, who spontaneously played at least 15 games. To improve the usability of Mila-Learn, we also asked users (both children and families) to provide feedback on the games and information on the children's impairments. This information was provided freely and was not compulsory to obtain Mila-Learn. In addition, we systematically collected through phone interviews all the problems that the children and their families encountered regarding the computing and web performance of the SG. Finally, we conducted a phone survey of 200 users, which is provided in Multimedia Appendix 1 [56]. The questions asked were designed to gain insights into the families' perceptions of the benefits of the tool, the improvements and difficulties of use they encountered, and their desire to continue using the game in the future; in addition, room was left for unstructured testimony. The data analysis for the feasibility study was limited to descriptive statistics.

Usability Study

This usability study was considered a continuation of study 1 and was conducted under the same ethical rules. In the usability study carried out in a real-life setting over 6 months, our primary focus was 2-fold following modifications to Mila-Learn based on study 1 feedback: first, to ensure that the computational architecture and final version of Mila-Learn were free of computer bugs and, second, to track player progress using Mila-Learn's scoring system over an extended duration. As part of the second lockdown because of the COVID-19 crisis, the final version of Mila-Learn was made available again starting on October 10, 2020, on National Learning Disabilities Day. Benefiting from the large amount of feedback received during the first lockdown, very few technical problems occurred, resulting in a game with much better fluidity that provided higher-quality data. A total of 3337 children had access to Mila-Learn for a total of 84,682 games that were played. As in study 1, at the time of registration, the patients' families were given the opportunity to complete the profile of the children, including information such as the children's diagnoses. A total of 304 diagnoses were reported by the parents. Finally, the children and their families had the option of linking the game character to the reported clinical profile. This option was exercised by 2.94% (98/3337) of the children, for whom we had both their reported diagnosis and game performance over time. These 2.94% (98/3337) of the children completed 3922 games.

To assess how children performed with Mila-Learn, we defined and computed the following variables:

1. "Time" is an incremental value representing the number of levels played by a player since the beginning of the experiment. Time is 1 at the beginning of the experiment and represents the total number of levels played by the player at the end of the experiment.
2. "Delta_tap" is the delay between the date of the played input (as defined in Table 1) and the date of the expected input.
3. "Threshold" is a delay defined for each game that was used to construct the performance score.
4. "Performance score" is a variable bounded between 0 and 100 that was created to quantify performance from delta_tap and normalize performance across games. We used the following formulas: performance score = (–100/threshold) × abs(delta_tap) + 100 for abs(delta_tap)≤threshold and performance score=0 for abs(delta_tap)>threshold.

We conducted several linear mixed models. To assess children's progress over time, we tested whether players improved their performance through the progression across the games using a linear mixed model with the following formula: performance score ~ time + (1|PlayerID/GameID/LevelName).

To assess whether a declared diagnosis was associated with the average performance of the children, we also conducted a linear mixed model using the following formula: performance score ~ age + dyscalculia + dysgraphia + dyslexia + dysphasia + ADHD + ExecutiveFunction impairment + (1|GameID/LevelName).

Finally, we also tested whether progress over time was moderated by a declared diagnosis using the following formula: performance score ~ time + diagnosis + time × diagnosis + (1|GameID/LevelName).

Ethical Considerations

Under French legislation, we did not need the approval of a Comité de Protection des Personnes (Committee for the Protection of Persons). However, as the pilot study was conducted in line with the creation of large databases, we obtained the approval of the Commission Nationale de l'Informatique et des Libertés (National Commission for Informatics and Freedoms) under number 2222283.

Results

Feasibility Study

Between April 2020 and June 2020, a total of 2500 children had access to Mila-Learn. Families reported the child's diagnosis in 60% (1500/2500) of cases. As children were recruited through the FFDys, they were diagnosed with an NDD in almost all cases, but only 23% (575/2500) were declared as having dyslexia. The other children had developmental coordination disorders (dyspraxia), dyscalculia, and communication disorders of oral language (dysphasia). In addition, 18% (450/2500) declared a diagnosis of ADHD.

Data regarding the use of Mila-Learn by each user were recorded as time spent on the SG and accuracy in each game played. The average use was 3.5 sessions per week. To ensure the significance of the data, we only kept the data of players who participated over a sufficient period (>15 games). Duration was expressed as the number of games played. We considered the number of games played inside the SG over the number of played sessions as the number of games played in 1 session could vary widely. In total, 21% (525/2500) of players aged 6 to 14 years played at least 15 games, with an overall mean of 54.77 and a median of 42 games played. The average number of games played was similar across ages (no main effect of age). No effect of age was found on the mean score. In addition, no floor or ceiling effects were observed (Multimedia Appendix 2).

It should be noted that several technical issues occurred during the first 2 weeks owing to the wide variety of tablet operating systems. This situation resulted in the deployment of corrective patches, but owing to the correction delay, it may have differentially altered one child's experience relative to another's. To improve the user experience, phone calls were systematically conducted to interview families, determine potential problem areas, and gather feedback for improvement. Parents consistently highlighted the recreational side of the game and its impact on the children's self-confidence. A survey of 200 users, provided in Multimedia Appendix 1, also indicated that 96% (192/200) wished to continue using Mila-Learn after the COVID-19 pandemic. However, they also provided significant feedback (164/200, 82%) to improve the game. Multimedia Appendix 3 [57] presents the most significant feedback with a frequency of ≥10 occurrences. We classified it according to the criteria by Morville [56], which distinguish 7 dimensions: usefulness, usability, findability (the ease of locating a feature or a piece of context), credibility, accessibility design, attractiveness, and value [48]. Usability was questioned in several comments, such as "the detection of movements should be improved," "sound detection needs to be improved," and "the game needs to be better adapted to the child's difficulty profile." Accessibility was also questioned as several parents indicated that "the writing could benefit from being larger and the display of dialogues slower."

Mila-Learn Description Adjustments Following Study 1

Design Framework

On the basis of the feedback obtained during study 1, we made several modifications to Mila-Learn. To improve accessibility, the first modification was to offer the player the choice between several fonts, including OpenDyslexic. This choice is reversible throughout the game. We also improved sound and movement detection. A second significant choice was to distinguish the children's pathways according to their predominant disorders to facilitate their entry into the game and usability. For example, a player who indicated that they had dyspraxia at the time of registration was offered more moderate motor exercises (ie, Sing Lab), allowing them to enter the adventure before training on River Splash or Karate Fruits that are more challenging in terms of motor abilities. In contrast, a child with dyslexia could be offered River Splash from the beginning, with Sing Lab exercises being offered only afterward as Sing Lab involves the phonological loop.

To increase motivation and interest in the game, we provided new possibilities of personalization for the character: the player could choose the gender of the avatar, their color, and the color of the hat. We then increased the storyline with the help of a screenwriter. This modification improved the consistency of the story and made it more inclusive by adding new characters that could help the player during the game. A new companion named "Mila" appeared, who is a fairy representing the planet "Mila" where the story takes place. These modifications also influenced (1) the dialogues, which were shortened with the language adapted to children; and (2) the appearance of the notion of "rhythmic," which was introduced as a martial art based on rhythm to clarify the main goal of the game during the adventure.

Through this expansion, we created 6 new chapters. We maintained the same concept as the preceding version and gradually increased the level of difficulty during the progression of the game by increasing the speed of the rhythm and varying the type of rhythm clapped (ie, clapping notes, then eighth notes). We also created daily missions. These 4 daily tasks allowed the child to revisit games on which they had practiced in the past and where they encountered difficulties. This allowed us to directly address the tendency to forget what has been learned and allowed for longer practice with Mila-Learn.

Description of the Tasks

Finally, we made structural modifications to the proposed tasks to ensure the game's fluidity and improve motor interactions. First, we changed the way children had to answer during Clap Trap. Instead of clapping both hands, which was recorded using the microphone, we changed the child's interaction with the SG to synchronously tapping both hands on the screen (and, therefore, we used a touch recording). Second, Follow me was extensively modified to be more understandable and involve the child more on a motor level. The interaction was changed from a passive mode (one contact pressure) to a more active hand clapping measured using the microphone. The child did not perform the task all at once but interacted with the character, who gave them instructions that the child reproduced on the principle of call and response. Specific music was created for the game. As a consequence, Follow me was renamed Clap Hero. Finally, we modified the way children had to answer during Fruity Jump—children's interaction with Mila-Learn changed from tapping to shaking the tablet, which was measured using an accelerometer. Table 2 summarizes the changes made in the final version of Mila-Learn.

Table 2. Players' recorded responses and game parameters in the final version of Mila-Learn.

Task	Type	Interaction	Capture technology	Songs	Tolerance threshold
Clap Hero	*Call and response* [a]	*Clapping hands*	*Microphone*	*Customized*	• Perfect: 0.1 s before or after the beat • Good: 0.15 s before or after the beat • Correct: 0.25 s before or after the beat
Clap Trap	Last beat	*Tapping on the left and right side of the screen*	*Touch*	Commercial	• Perfect: 0.1 s before or after the beat • Good: 0.15 s before or after the beat • Correct: 0.25 s before or after the beat
River Splash	Last beat	Shaking tablet	Accelerometer	Commercial	• Perfect: 0.1 s before or after the beat • Good: 0.15 s before or after the beat • Correct: 0.25 s before or after the beat
Sing Lab	Call and response	Singing	Microphone	*Customized*	• Pattern: 0.15 s before or after the beat and up to 30% of the note duration • Song duration: the note must be sung at least 60% of the time
Fruity Jump	Call and response	*Shaking tablet*	*Accelerometer*	Customized	• Perfect: 0.2 s before or after the beat • Good: 0.25 s before or after the beat • Correct: 0.3 s before or after the beat
Karate Fruits	Last beat	Punching	Webcam	Customized	• Perfect: 0.08 s before or after the beat • Good: 0.14 s before or after the beat • Correct: 0.3 s before or after the beat

[a]Italics indicate game and functional changes that were introduced compared with the Mila-Learn second version summarized in Table 1.

Usability Study

This usability study focused on a sample of 98 children (mean age 9.05; SE 2.4 years), and we had both their reported diagnoses and game performance over time. These 98 children completed 3922 games. The linear mixed models yielded the following significant results. First, we found that the performance of the children significantly improved over time (β=.02; t_{3268}=2.68; *P*=.007). That is, there was an increase in the performance score by an average of 5 points after 250 levels were played.

Second, we explored whether declared diagnosis and age influenced the average performance of the children. Table 3 summarizes the results. We found that older children performed better than younger children. One year of age increased the normalized performance score by 1.08 points (meaning 1.1% of the maximal range). In addition, children with dyslexia and ADHD performed significantly better than those with other diagnoses (performance improved significantly faster in children with ADHD, β=.06; t_{3754}=3.91; *P*<.001, and slower in children with dyslexia, β=−.06; t_{3816}=–5.08; *P*<.001). Having dyslexia increased the normalized performance score by 2.81 points (meaning 2.8% of the maximal range) compared with children without dyslexia, whereas having ADHD increased the normalized performance score by 4.16 points (meaning 4.2% of the maximal range) compared with children without ADHD. In contrast, children with executive function impairment and dysgraphia performed significantly worse than those with other diagnoses. Having dysgraphia decreased the normalized performance score by 2.06 points (meaning 2.1% of the maximal range) compared with children without dysgraphia, whereas having executive function impairment decreased the normalized performance score by 3.26 points (meaning 3.3% of the maximal range) compared with children without executive function impairment.

Finally, we also tested whether progress over time statistically interacted with the declared diagnosis. We found that children with ADHD progressed faster over time than those with other diagnoses (β=.06; t_{3754}=3.91; *P*<.001) and that children with dyslexia (β=−.06; t_{3816}=–5.08; *P*<.001) and executive dysfunction (β=−.03; t_{3805}=–2.09; *P*=.04) improved less over time than those with other diagnoses. We found no significant interaction between time and a diagnosis of dysphasia (β=−.01; t_{3816}=–0.68; *P*=.50), dyscalculia (β=.05; t_{3787}=1.46; *P*=.14), or dysgraphia (β=.00; t_{3816}=–0.1; *P*=.92).

Table 3. Average scores according to diagnosis during study 3 with the final version of Mila-Learn.

	Estimate (SE)	*t* test (*df*)	Pr(>|*t*|)
Intercept	37.83 (4.60)	8.23 (6.04)	<.001
Age	1.08 (0.15)	7.02 (2082.28)	<.001
Dyscalculia (yes)	−0.43 (1.03)	−0.42 (2081.02)	.68
Dysgraphia (yes)	−2.06 (0.79)	−2.60 (2084.61)	.009
Dyslexia (yes)	2.81 (0.70)	4.03 (2087.52)	<.001
Dysphasia (yes)	−0.51 (0.93)	−0.54 (2086.54)	.59
ADHD[a] (yes)	4.16 (0.61)	6.79 (2090.71)	<.001
Executive function impairment (yes)	−3.26 (1.36)	−2.40 (2086.78)	.02

[a]ADHD: attention-deficit/hyperactivity disorder.

Discussion

Principal Findings

The literature on SGs, especially when designed for a specific medical condition, is limited when it focuses on game design methodology or formal clinical validation [42]. In this paper, we described the process and empirical studies to address this issue for Mila-Learn, an SG based on rhythmic training for children with dyslexia. To do so, we placed the patient's experience at the center of the game construction while iterating with clinicians involved in treating children with dyslexia. In this paper, we described the different developmental phases that helped us design the game. We first constructed an initial prototype based on a literature review and with the help of clinicians specializing in learning disorders. Then, based on a first round of feedback from users and comments from professionals, we developed a first version of Mila-Learn for tablets.

In this version, we greatly improved the users' experience with the game by adding new gaming features to increase the motivation and engagement of players. We offered more possibilities for customization, created a storyline, and introduced humorous and friendly characters to align with children's interests [51,52]. Moreover, we adapted the difficulty of the game to enhance the learning possibilities of children by working on graphism and the instructions given to the children and by creating evolving tasks that gradually increased the level of difficulty [51,53,54]. With this second version, we adopted a user participatory design by inviting children, families, and professionals to test this version and send us feedback about their experience (feasibility study). User participatory design is a method that is currently gaining attention. Contrary to user-centered designs, which create games for a user, participatory design aims to construct the game with the users by collecting their experience and advice and then including them in the game [58]. It has been shown that participatory design promotes engagement of the user [52]. Indeed, collecting feedback both from families and children and from professionals is essential as professionals and families and children focus on different aspects of an SG and do not place the same importance on each feature [52]. We believe that this participative process helped us develop an SG that improved the experience within the game and the interest of families and children in Mila-Learn.

Regarding computational aspects, we also collected feedback that helped us resolve bugs and record the time spent on the game and the player's accuracy in each game. These features allowed us to follow children's interest in the game and their progression over time and demonstrate that progression occurred with Mila-Learn and was associated with age. Study 1 confirmed that children could engage with Mila-Learn for a rather long period and play at home without the need for an extra supporting person, suggesting that Mila-Learn was sufficiently motivating and adapted to this population. Children and their families appeared to be highly satisfied with the game.

Finally, following a third round of feedback from parents, children, and professionals, we developed a final version of Mila-Learn to improve accessibility and motivation for the player. We made structural modifications to the proposed tasks to ensure the fluidity of the game and improve motor interactions. We resolved most of the technical problems, which allowed us to conduct a real-life usability study of the Mila-Learn game during the second lockdown.

Comparison With Prior Work

In the usability study, we observed that children significantly improved their scores on the 6 games included in Mila-Learn. Although we cannot conclude that the rhythm abilities of the children improved based only on these results, we believe that the children learned how to use Mila-Learn and that they were increasingly accurate in responding to each game. However, the effect size was small, although it may have been underestimated as the difficulty in the games increased, which could have masked the children's progression. In addition, based on the diagnosis declared by the children's parents, we performed exploratory analyses to assess whether improvements over time were associated with the declared diagnoses. Linear mixed models showed that children's performance significantly increased over time, that scores were better for children with ADHD and dyslexia, and that performance improved significantly faster for children with ADHD and slower for children with dyslexia.

Regarding the average performance of children according to diagnosis, the results were very encouraging if we consider the relationship between reading impairments and diagnosis. On the basis of the literature, we expected reading impairments to be associated with dyslexia, attention deficit, and specific oral language impairment (dysphasia) [59,60]. In addition, we expected that severity would negatively influence performance. This is usually the case when children have dysphasia [61], executive function impairment [62], or multidimensional impairments [63]. The results were in line with these expectations. Children with dyslexia and ADHD showed a significantly better performance over time, whereas children with dysphasia, executive function impairment, and dysgraphia showed a worse performance. Of note, children with dysgraphia often have motor coordination disorders [64]. Finally, dyscalculia had no influence on Mila-Learn performance. In summary, the predictions according to diagnosis were in line with the hypothesis that Mila-Learn may improve performance in children with reading impairments. The fact that dyscalculia showed no specific effects and that the diagnoses associated with the highest severity (dysphasia and executive function impairment) showed less improvement followed our hypothesis [61-63,65]. We speculated that dysgraphia was associated with multidimensional impairments, including motor coordination disorder. This interpretation is based on the fact that recruitment for the study was based only on user reading impairments.

Regarding the average performance of children according to diagnosis over time (ie, the statistical interaction), the fact that performance improved significantly faster for children with ADHD and slower for children with dyslexia is not surprising as the perception of rhythm is impaired in children with dyslexia. In contrast, children with ADHD may have impairments in reading abilities but do not have specific deficits in rhythm and speech perception [57,66].

Strengths and Limitations

The exploratory studies presented in this paper have some limitations despite the promising results. On the one hand, some aspects of the game need to be improved.

First, we currently consider a "standard" latency of 20 ms, which corresponds to the estimated delay between the child's real input and the input processed by the operating system. In reality, each tablet may have unique differences. In the next version of Mila-Learn, we need to consider this unique latency to get as close as possible to the real performance of children. This adjustment might lead to more accurate measurements of children's interactions and, potentially, more tailored game experiences.

Second, the game gradually increases in difficulty with the progression of the player within the game. We integrated some specific pathways as a function of the difficulties that the children declared before starting the game (ie, children with motor difficulties do not start with games that require a high level of motor skills). However, the progression is predetermined and does not take into account the results of the player. In the next version of Mila-Learn, the difficulty of each game will automatically adapt based on the child's performance in the previous games using a specific algorithm [53], allowing for much better stimulation. By doing so, the game could offer a more individualized experience, potentially leading to more sustained engagement and greater benefits for the children.

Third, the age range of 7 to 14 years is wide as children's interests can vary greatly during these years. In a future version of Mila-Learn, the graphics and music will be adapted to the age specified by the child so that the game will be more suitable for their age. This may enhance the game's appeal to players across the entire age range, fostering increased engagement and learning.

However, our study was only exploratory in nature. First, even if the lockdown gave us the opportunity to have a large sample for exploratory studies, diagnoses were not clinically grounded and were only declared by the children's parents. Therefore, caution should be the rule when interpreting predictive models.

Second, there was no predefined design for the studies as the training was spontaneous and included no comparison with alternative treatment proposals. Therefore, the clinical interest of Mila-Learn for dyslexia cannot be established based on the results of the 2 exploratory studies presented in this paper.

Future Directions

To address the clinical relevance of Mila-Learn in relation to dyslexia, the next step will be to evaluate the effects of Mila-Learn in the context of a randomized controlled trial. Children with dyslexia based on objective clinical assessments will be randomized to Mila-Learn sessions or placebo game sessions that take place in the same universe but with different tasks. We will assess the evolution of reading skills from before to after training with the hope of greater improvements with Mila-Learn. On the basis of the exploratory studies, we calculated the number of patients per group that would ensure a statistical power of at least 85% for an effect size equal to 0.5 (moderate) when the changes in the experimental and control groups were compared. This calculation indicated that each group should have at least 73 children (ie, 146 children in total). This study started in September 2021 (Comité de Protection des Personnes registration 2021-A01709-32; ClinicalTrials.gov Identifier: NCT05154721).

Conclusions

We presented how we constructed Mila-Learn, an SG based on rhythm activities, to improve reading skills in children with dyslexia. We developed several versions of the game considering the literature, professionals' experiences, and users' feedback. We also conducted a usability and a feasibility study to evaluate each version of Mila-Learn. The results indicated that Mila-Learn was attractive and sustained the players' motivation and engagement for several months. Moreover, children were able to learn how to use the game, and their performance in the games improved with training. Future research will include (1) adapting to the latency of the electronic devices, (2) automatically adapting the games based on the player's performance, and (3) conducting a large randomized controlled trial to evaluate the impact of Mila-Learn on reading skills.

Acknowledgments

The authors would like to thank Valentin Begel for his counsel in the design of this project, the Le Kremlin-Bicêtre hospitals for helping them develop the first version of Mila-Learn, and the École Polytechnique for its financial support. The authors would also like to thank all the families, children, and professionals for their feedback and encouragement. Special thanks to the French Federation for Learning Disorders. This work was supported by the École Polytechnique (grant "Prix Gérondeau 2018"). The sponsors of the aforementioned study funding source were not involved in the study, writing of the report, or decision to submit the paper for publication.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

FV conceptualized the study with DC and CG, managed data curation, acquired funding, undertook the investigation, and contributed to software development alongside AY. AY participated in data curation, formal analysis, and software development. HP was responsible for formal analysis and visualization. DC further contributed to conceptualization and was involved in methodology, supervision, and the review and editing process. CG also took part in conceptualization, formal analysis, writing of the original draft, investigation, supervision, and the review and editing process.

Conflicts of Interest

FV reports a relationship with bMotion Technologies that includes equity or stocks. AY reports a relationship with bMotion Technologies that includes employment and equity or stocks.

Multimedia Appendix 1
User survey on Mila-Learn.
[DOCX File , 26 KB - games_v12i1e42733_app1.docx]

Multimedia Appendix 2
Average scores from all games between April 2020 and June 2020 according to the children's age.
[PNG File , 22 KB - games_v12i1e42733_app2.png]

Multimedia Appendix 3
Feedback classification based on the criteria by Morville [55].
[DOCX File , 27 KB - games_v12i1e42733_app3.docx]

References

1. Bilhartz T, Bruhn R, Olson J. The effect of early music training on child cognitive development. J Appl Dev Psychol 1999 Dec;20(4):615-636 [FREE Full text] [doi: 10.1016/s0193-3973(99)00033-7]
2. Habibi A, Damasio A, Ilari B, Elliott Sachs M, Damasio H. Music training and child development: a review of recent findings from a longitudinal study. Ann N Y Acad Sci (Forthcoming) 2018 Mar 06. [doi: 10.1111/nyas.13606] [Medline: 29508399]
3. Roden I, Könen T, Bongard S, Frankenberg E, Friedrich EK, Kreutz G. Effects of music training on attention, processing speed and cognitive music abilities—findings from a longitudinal study. Appl Cogn Psychol 2014 Jul 02;28(4):545-557 [FREE Full text] [doi: 10.1002/acp.3034]
4. Schlaug G, Norton A, Overy K, Winner E. Effects of music training on the child's brain and cognitive development. Ann N Y Acad Sci 2005 Dec;1060:219-230. [doi: 10.1196/annals.1360.015] [Medline: 16597769]
5. Wan CY, Schlaug G. Music making as a tool for promoting brain plasticity across the life span. Neuroscientist 2010 Oct;16(5):566-577 [FREE Full text] [doi: 10.1177/1073858410377805] [Medline: 20889966]
6. Hyde KL, Lerch J, Norton A, Forgeard M, Winner E, Evans AC, et al. The effects of musical training on structural brain development: a longitudinal study. Ann N Y Acad Sci 2009 Jul;1169:182-186. [doi: 10.1111/j.1749-6632.2009.04852.x] [Medline: 19673777]
7. Patel AD. Language, music, syntax and the brain. Nat Neurosci 2003 Jul;6(7):674-681. [doi: 10.1038/nn1082] [Medline: 12830158]
8. Zatorre R, McGill J. Music, the food of neuroscience? Nature 2005 Mar 17;434(7031):312-315. [doi: 10.1038/434312a] [Medline: 15772648]

9. Cancer A, Bonacina S, Antonietti A, Salandi A, Molteni M, Lorusso ML. The effectiveness of interventions for developmental dyslexia: rhythmic reading training compared with hemisphere-specific stimulation and action video games. Front Psychol 2020 Jun 03;11:1158 [FREE Full text] [doi: 10.3389/fpsyg.2020.01158] [Medline: 32581961]
10. Flaugnacco E, Lopez L, Terribili C, Montico M, Zoia S, Schön D. Music training increases phonological awareness and reading skills in developmental dyslexia: a randomized control trial. PLoS One 2015 Sep 25;10(9):e0138715 [FREE Full text] [doi: 10.1371/journal.pone.0138715] [Medline: 26407242]
11. Habib M, Lardy C, Desiles T, Commeiras C, Chobert J, Besson M. Music and dyslexia: a new musical training method to improve reading and related disorders. Front Psychol 2016 Sep 25;7:26 [FREE Full text] [doi: 10.3389/fpsyg.2016.00026] [Medline: 26834689]
12. Overy K. Dyslexia and music. From timing deficits to musical intervention. Ann N Y Acad Sci 2003 Nov;999:497-505. [doi: 10.1196/annals.1284.060] [Medline: 14681173]
13. Thomson JM, Leong V, Goswami U. Auditory processing interventions and developmental dyslexia: a comparison of phonemic and rhythmic approaches. Read Writ 2012 Feb 15;26(2):139-161 [FREE Full text] [doi: 10.1007/s11145-012-9359-6]
14. Moreno S, Bialystok E, Barac R, Schellenberg EG, Cepeda NJ, Chau T. Short-term music training enhances verbal intelligence and executive function. Psychol Sci 2011 Nov;22(11):1425-1433 [FREE Full text] [doi: 10.1177/0956797611416999] [Medline: 21969312]
15. Chobert J, Marie C, François C, Schön D, Besson M. Enhanced passive and active processing of syllables in musician children. J Cogn Neurosci 2011 Dec;23(12):3874-3887. [doi: 10.1162/jocn_a_00088] [Medline: 21736456]
16. François C, Chobert J, Besson M, Schön D. Music training for the development of speech segmentation. Cereb Cortex 2013 Sep;23(9):2038-2043 [FREE Full text] [doi: 10.1093/cercor/bhs180] [Medline: 22784606]
17. Jentschke S, Koelsch S, Friederici AD. Investigating the relationship of music and language in children: influences of musical training and language impairment. Ann N Y Acad Sci 2005 Dec;1060:231-242. [doi: 10.1196/annals.1360.016] [Medline: 16597770]
18. Bidelman GM, Weiss MW, Moreno S, Alain C. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. Eur J Neurosci 2014 Aug;40(4):2662-2673. [doi: 10.1111/ejn.12627] [Medline: 24890664]
19. Herholz SC, Zatorre RJ. Musical training as a framework for brain plasticity: behavior, function, and structure. Neuron 2012 Nov 08;76(3):486-502 [FREE Full text] [doi: 10.1016/j.neuron.2012.10.011] [Medline: 23141061]
20. Dalla Bella S, Białuńska A, Sowiński J. Why movement is captured by music, but less by speech: role of temporal regularity. PLoS One 2013;8(8):e71945 [FREE Full text] [doi: 10.1371/journal.pone.0071945] [Medline: 23936534]
21. Repp BH, Su Y. Sensorimotor synchronization: a review of recent research (2006-2012). Psychon Bull Rev 2013 Jun;20(3):403-452. [doi: 10.3758/s13423-012-0371-2] [Medline: 23397235]
22. Tierney A, Kraus N. The ability to move to a beat is linked to the consistency of neural responses to sound. J Neurosci 2013 Sep 18;33(38):14981-14988 [FREE Full text] [doi: 10.1523/JNEUROSCI.0612-13.2013] [Medline: 24048827]
23. Bonacina S, Krizman J, White-Schwoch T, Nicol T, Kraus N. Distinct rhythmic abilities align with phonological awareness and rapid naming in school-age children. Cogn Process 2020 Nov;21(4):575-581 [FREE Full text] [doi: 10.1007/s10339-020-00984-6] [Medline: 32607802]
24. Gordon RL, Fehd HM, McCandliss BD. Does music training enhance literacy skills? A meta-analysis. Front Psychol 2015 Dec 01;6:1777 [FREE Full text] [doi: 10.3389/fpsyg.2015.01777] [Medline: 26648880]
25. Huss M, Verney JP, Fosker T, Mead N, Goswami U. Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. Cortex 2011 Jun;47(6):674-689. [doi: 10.1016/j.cortex.2010.07.010] [Medline: 20843509]
26. Goswami U. A temporal sampling framework for developmental dyslexia. Trends Cogn Sci 2011 Jan;15(1):3-10. [doi: 10.1016/j.tics.2010.10.001] [Medline: 21093350]
27. Jung H, Sontag S, Park YS, Loui P. Rhythmic effects of syntax processing in music and language. Front Psychol 2015 Nov 23;6:1762 [FREE Full text] [doi: 10.3389/fpsyg.2015.01762] [Medline: 26635672]
28. Patel AD, Daniele JR. An empirical comparison of rhythm in language and music. Cognition 2003 Feb;87(1):B35-B45. [doi: 10.1016/s0010-0277(02)00187-7] [Medline: 12499110]
29. Chen JL, Penhune VB, Zatorre RJ. Listening to musical rhythms recruits motor regions of the brain. Cereb Cortex 2008 Dec;18(12):2844-2854. [doi: 10.1093/cercor/bhn042] [Medline: 18388350]
30. Kotz SA, Schwartze M. Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn Sci 2010 Sep;14(9):392-399. [doi: 10.1016/j.tics.2010.06.005] [Medline: 20655802]
31. Goswami U, Huss M, Mead N, Fosker T, Verney JP. Perception of patterns of musical beat distribution in phonological developmental dyslexia: significant longitudinal relations with word reading and reading comprehension. Cortex 2013 May;49(5):1363-1376. [doi: 10.1016/j.cortex.2012.05.005] [Medline: 22726605]
32. Leong V, Goswami U. Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. Front Hum Neurosci 2014;8:96 [FREE Full text] [doi: 10.3389/fnhum.2014.00096] [Medline: 24605099]

33. Tallal P. Improving language and literacy is a matter of time. Nat Rev Neurosci 2004 Sep;5(9):721-728. [doi: 10.1038/nrn1499] [Medline: 15322530]
34. Mayer-Benarous H, Benarous X, Vonthron F, Cohen D. Music therapy for children with autistic spectrum disorder and/or other neurodevelopmental disorders: a systematic review. Front Psychiatry 2021 Apr 09;12:643234 [FREE Full text] [doi: 10.3389/fpsyt.2021.643234] [Medline: 33897497]
35. Robinson LR, Holbrook JR, Bitsko RH, Hartwig SA, Kaminski JW, Ghandour RM, et al. Differences in health care, family, and community factors associated with mental, behavioral, and developmental disorders among children aged 2-8 years in rural and urban areas - United States, 2011-2012. MMWR Surveill Summ 2017 Mar 17;66(8):1-11 [FREE Full text] [doi: 10.15585/mmwr.ss6608a1] [Medline: 28301449]
36. Bitsko RH, Holbrook JR, Ghandour RM, Blumberg SJ, Visser SN, Perou R, et al. Epidemiology and impact of health care provider-diagnosed anxiety and depression among US children. J Dev Behav Pediatr 2018 Jun;39(5):395-403 [FREE Full text] [doi: 10.1097/DBP.0000000000000571] [Medline: 29688990]
37. Zhonggen Y. A meta-analysis of use of serious games in education over a decade. Int J Comput Games Technol 2019 Feb 03;2019:1-8 [FREE Full text] [doi: 10.1155/2019/4797032]
38. Wouters P, van Nimwegen C, van Oostendorp H, van der Spek ED. A meta-analysis of the cognitive and motivational effects of serious games. J Educ Psychol 2013 May;105(2):249-265 [FREE Full text] [doi: 10.1037/a0031311]
39. Mondéjar T, Hervás R, Johnson E, Gutiérrez-López-Franca C, Latorre JM. Analyzing EEG waves to support the design of serious games for cognitive training. J Ambient Intell Human Comput 2018 May 12;10(6):2161-2174 [FREE Full text] [doi: 10.1007/s12652-018-0841-0]
40. Lamb R, Annetta L, Firestone J, Etopio E. A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. Comput Human Behav 2018 Mar;80:158-167 [FREE Full text] [doi: 10.1016/j.chb.2017.10.040]
41. García-Redondo P, García T, Areces D, Núñez JC, Rodríguez C. Serious games and their effect improving attention in students with learning disabilities. Int J Environ Res Public Health 2019 Jul 11;16(14):2480 [FREE Full text] [doi: 10.3390/ijerph16142480] [Medline: 31336804]
42. Grossard C, Grynspan O, Serret S, Jouen AL, Bailly K, Cohen D. Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (ASD). Comput Educ 2017 Oct;113:195-211 [FREE Full text] [doi: 10.1016/j.compedu.2017.05.002]
43. Bégel V, Seilles A, Dalla Bella S. Rhythm workers: a music-based serious game for training rhythm skills. Music Sci 2018 Jan;1:1-16 [FREE Full text] [doi: 10.1177/2059204318794369]
44. Prato A, Puglisi GE, Astolfi A, Sacco T. Design of the rhythmic tasks of a computer-assisted phonological training to improve reading abilities in first-graders. In: Proceedings of the 11th European Congress and Exposition on Noise Control Engineering. 2018 Presented at: ECENCE '18; May 27-31, 2018; Crete, Greece p. 1-4 URL: https://www.euronoise2018.eu/docs/EURONOISE%202018%20Brochure_27072017.pdf
45. Ostiz-Blanco M, Lallier M, Grau S, Rello L, Bigham JP, Carreiras M. Jellys: towards a videogame that trains rhythm and visual attention for dyslexia. In: Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. 2018 Presented at: ASSETS '18; October 22-24, 2018; Galway, Ireland p. 395-403 URL: https://dl.acm.org/doi/10.1145/3234695.3241028 [doi: 10.1097/DBP.0000000000000571]
46. Valentine AZ, Brown BJ, Groom MJ, Young E, Hollis C, Hall CL. A systematic review evaluating the implementation of technologies to assess, monitor and treat neurodevelopmental disorders: A map of the current evidence. Clin Psychol Rev 2020 Aug;80:101870 [FREE Full text] [doi: 10.1016/j.cpr.2020.101870] [Medline: 32712216]
47. Bolduc J, Rondeau J. Rythmons les apprentissages !. Langage et pratiques 2015;56:12-22 [FREE Full text]
48. Horowitz-Kraus T. The role of executive functions in the reading process. In: Khateb A, Bar-Kochva I, editors. Reading Fluency: Currency Insights from Neurocognitive Research and Intervention Studies. Cham, Switzerland: Springer; 2016:51-63.
49. Fromageau M. Effets de la pratique du rythme musical sur le langage des enfants porteurs de troubles spécifiques du langage oral et écrit. Nantilus. 2019. URL: https://nantilus.univ-nantes.fr/vufind/Record/PPN241556023/Description [accessed 2023-11-15]
50. Lonergan A, Doyle C, Cassidy C, MacSweeney Mahon S, Roche RA, Boran L, et al. A meta-analysis of executive functioning in dyslexia with consideration of the impact of comorbid ADHD. J Cogn Psychol 2019 Oct 03;31(7):725-749 [FREE Full text] [doi: 10.1080/20445911.2019.1669609]
51. Powell L, Parker J, Harpin V, Mawson S. Guideline development for technological interventions for children and young people to self-manage attention deficit hyperactivity disorder: realist evaluation. J Med Internet Res 2019 Apr 03;21(4):e12831 [FREE Full text] [doi: 10.2196/12831] [Medline: 30942692]
52. Tang JS, Falkmer M, Chen NT, B lte S, Girdler S. Designing a Serious Game for Youth with ASD: perspectives from end-users and professionals. J Autism Dev Disord 2019 Mar;49(3):978-995. [doi: 10.1007/s10803-018-3801-9] [Medline: 30377883]
53. Carlier S, Van der Paelt S, Ongenae F, De Backere F, De Turck F. Empowering children with ASD and their parents: design of a serious game for anxiety and stress reduction. Sensors (Basel) 2020 Feb 11;20(4):966 [FREE Full text] [doi: 10.3390/s20040966] [Medline: 32054025]

54. Yusoff A, Crowder R, Gilbert L. Validation of serious games attributes using the technology acceptance model. In: Proceedings of the 2nd International Conference on Games and Virtual Worlds for Serious Applications. 2010 Presented at: ICGVWS '10; March 25-26, 2010; Braga, Portugal p. 45-51 URL: https://ieeexplore.ieee.org/abstract/document/5460149 [doi: 10.1109/vs-games.2010.7]
55. Dalton C. Interaction design in the built environment: designing for the 'universal user'. Stud Health Technol Inform 2016;229:314-323. [Medline: 27534322]
56. Morville P, Rosenfeld L. Information Architecture for the World Wide Web: Designing Large-Scale Web Sites, 3rd edition. Sebastopol, CA: O'Reilly Media; 2006.
57. Tiffin-Richards MC, Hasselhorn M, Richards ML, Banaschewski T, Rothenberger A. Time reproduction in finger tapping tasks by children with attention-deficit hyperactivity disorder and/or dyslexia. Dyslexia 2004 Nov;10(4):299-315. [doi: 10.1002/dys.281] [Medline: 15573962]
58. Sanders EB. From user-centered to participatory design approaches. In: Frascara J, editor. Design and the Social Sciences. Boca Raton, FL: CRC Press; 2002.
59. Frazier TW, Youngstrom EA, Glutting JJ, Watkins MW. ADHD and achievement: meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. J Learn Disabil 2007;40(1):49-65. [doi: 10.1177/00222194070400010401] [Medline: 17274547]
60. Isoaho P, Kauppila T, Launonen K. Specific language impairment (SLI) and reading development in early school years. Child Lang Teach Ther 2015 Aug 26;32(2):147-157 [FREE Full text] [doi: 10.1177/0265659015601165]
61. Bishop DV, Snowling MJ. Developmental dyslexia and specific language impairment: same or different? Psychol Bull 2004 Nov;130(6):858-886. [doi: 10.1037/0033-2909.130.6.858] [Medline: 15535741]
62. Locascio G, Mahone EM, Eason SH, Cutting LE. Executive dysfunction among children with reading comprehension deficits. J Learn Disabil 2010;43(5):441-454 [FREE Full text] [doi: 10.1177/0022219409355476] [Medline: 20375294]
63. Xavier J, Cohen D. Multidimensional impairments. Handb Clin Neurol 2020;174:159-169. [doi: 10.1016/B978-0-444-64148-9.00012-0] [Medline: 32977875]
64. Biotteau M, Danna J, Baudou, Puyjarinet F, Velay J, Albaret J, et al. Developmental coordination disorder and dysgraphia: signs and symptoms, diagnosis, and rehabilitation. Neuropsychiatr Dis Treat 2019;15:1873-1885 [FREE Full text] [doi: 10.2147/NDT.S120514] [Medline: 31371960]
65. Dyslexia dysorthography dyscalculia: review of the scientific data. INSERM Collective Expertise Centre. URL: https://www.ncbi.nlm.nih.gov/books/NBK10787/pdf/Bookshelf_NBK10787.pdf [accessed 2023-11-15]
66. Gooch D, Snowling M, Hulme C. Time perception, phonological skills and executive function in children with dyslexia and/or ADHD symptoms. J Child Psychol Psychiatry 2011 Feb;52(2):195-203 [FREE Full text] [doi: 10.1111/j.1469-7610.2010.02312.x] [Medline: 20860755]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
FFDys: French Federation for Learning Disorders
NDD: neurodevelopmental disorder
SG: serious game

Edited by G Eysenbach, T Leung; submitted 29.09.22; peer-reviewed by M Orr; comments to author 02.12.22; revised version received 26.01.23; accepted 09.10.23; published 11.01.24.

Please cite as:
Vonthron F, Yuen A, Pellerin H, Cohen D, Grossard C
A Serious Game to Train Rhythmic Abilities in Children With Dyslexia: Feasibility and Usability Study
JMIR Serious Games 2024;12:e42733
URL: https://games.jmir.org/2024/1/e42733
doi:10.2196/42733
PMID:37830510

Original Paper

A Novel Casual Video Game With Simple Mental Health and Well-Being Concepts (Match Emoji): Mixed Methods Feasibility Study

Russell Pine[1], PhD; James Mbinta[1], PhD; Lisa Te Morenga[2], PhD; Theresa Fleming[1], PhD

[1]School of Health, Victoria University of Wellington, Wellington, New Zealand

[2]Research Centre for Hauora and Health, Massey University, Wellington, New Zealand

Corresponding Author:
Russell Pine, PhD
School of Health, Victoria University of Wellington
Kelburn Parade, Kelburn
Wellington, 6012
New Zealand
Phone: 64 4 472 1000
Email: russell.pine@vuw.ac.nz

Abstract

Background: Adolescence is a crucial phase for early intervention and prevention of mental health problems. Casual video games are popular and have promise as a novel mechanism for reaching young people, but this potential has seldom been explored.

Objective: This study aimed to explore the acceptability, feasibility, and possible indicators of therapeutic changes after playing a purpose-built novel casual video game (Match Emoji) with simple mental health and well-being content among young adolescents.

Methods: We conducted a single-arm, nonrandomized trial of Match Emoji with 12- to 14-year-old school students (N=45; 26 [57%] New Zealand European, 12 [26%] Māori; 7 [15%] Asian or Pacific; 27 [60%] boys, 3 [6%] non-binary). Participants were invited to play Match Emoji for 15 minutes, 2-3 times a week over 2 weeks (a total of 60 minutes). Acceptability was assessed through the frequency and duration of use (analytics analyzed at the end of the 2-week intervention period and at weeks 4 and 6) and through participant reports. The Child and Adolescent Mindfulness Measure (CAMM), General Help-Seeking Questionnaire (GHSQ), Flourishing Scale (FS), and Revised Children's Anxiety and Depression Scale (RCADS) were assessed at baseline and week 2 to indicate possible effects. Focus groups were held in week 4.

Results: Most participants (n=39, 87%) used Match Emoji for at least 60 minutes over the 2-week intervention, with 80% (36/45) continuing to play the game after the intervention period. Mean change (from baseline to 2 weeks) on each measure was 1.38 (95% CI –0.03 to 2.79; *P*=.06) for CAMM; 0.8 (95% CI –2.71 to 4.31; *P*=.64) for GHSQ; –1.09 (95% CI –2.83 to 0.66; *P*=.21) for FS; and –3.42 (95% CI –6.84 to –0.001; *P*=0.49) for RCADS. Focus group feedback suggested that Match Emoji was enjoyable and helpful.

Conclusions: The casual video game with mental health content appeared to be acceptable and provided a promising indication of possible therapeutic effects. This approach is worthy of further investigation.

International Registered Report Identifier (IRRID): RR2-10.2196/31588

(JMIR Serious Games 2024;12:e46697) doi:10.2196/46697

KEYWORDS

adolescent; anxiety; casual video games; digital mental health interventions; gaming; mental health; micro interventions; serious game; teenage; video game; youth

Introduction

Mental distress and low well-being are common among adolescents [1-3] and appear to have increased over the past decade, at least in high-income nations [4-6]. Cognitive behavioral therapy (CBT) and psychotropic medications are recommended for young people experiencing mental health disorders [7,8]. Preventative programs that aim to buffer against higher levels of distress later in life also exist for young people

[9]. Nevertheless, structural and attitudinal barriers inhibit access to mental health support for many young people [10].

Digital mental health interventions (DMHIs) refer to specialized content, support, or therapy for mental health conditions delivered electronically to treat, alleviate, or manage symptoms [11]. DMHIs encompass various technologies, including computerized CBT programs, chatbots, virtual reality for mental health conditions, games for mental health, apps, and interactive web pages [12]. Systematic reviews have shown promising effects for specific DMHIs across various age groups [13,14], such as CBT therapies for anxiety and depression [15,16]. Quality DMHIs can address some of the challenges often impacting face-to-face treatments [17,18]. For example, well-designed DMHIs can be used by young people irrespective of their level of distress, and they can be scaled up at a low cost due to their reduced reliance on clinically trained professionals [19,20].

Although this method of delivering mental health content is promising, engagement with DMHIs outside trial settings is typically lower than in trials [16]. Even playful interventions, such as Pesky gNATs [21] and SPARX [22], designed to appeal to young people's interest in computer games, have had limited evidence of engagement [16]. In part, these findings may reflect mismatches between how end users engage with technology and the way tools are provided (eg, sessions approximating weekly face-to-face therapies may be a poor match with contemporary patterns of internet use) [16,23,24]. Moreover, a lack of appealing options, lack of trust, or uncertainty about digital tools for mental health purposes may create additional barriers [22-24]. Therefore, while DMHIs have a great capacity to address mental health needs, it is important to keep exploring new opportunities to improve engagement [19,25,26].

Casual video games (CVGs) refer to simple games that can be played in short bursts of time, require no specialized skills, are often used for relaxation [27] and distraction purposes [28], and are generally free or low-cost to download and play. Well-known CVGs include "Rise Up" and "Angry Birds." "Rise Up" has been downloaded over 10 million times on the Google Play Store worldwide, and "Angry Birds" is played for approximately 200 million minutes daily [29,30]. Given their popularity and potential therapeutic effects, CVGs may be an approach that could be explored for delivering mental health and well-being content [28].

We systematically reviewed the effects of CVGs on anxiety, depression, stress, and low mood [31]. We found that 12 of the 13 trials reported promising results on their respective outcome measures. Following this work, we developed simple prototypes of CVGs with mental health concepts based on the puzzle, word, and match-3 genres and reviewed these in focus groups and interviews with young adolescents [32]. Young people indicated interest in this idea, with a match-3–style CVG being preferred. Subsequently, a game designer was contracted to develop the first version of Match Emoji, a simple match-3 CVG that includes brief text-based mental health and well-being messages, which have been previously described [33]. In brief, this includes short "micromessages," which were developed using psychological well-being literature and were sometimes linked to gameplay, for example, "Great job focusing and matching the emojis!" and "Phew! Take a short breath to help focus again." Subsequently, we held think-aloud interviews [34] with a small group of young adolescents to refine components.

In this study, we aimed to conduct a small open trial to explore the feasibility of using Match Emoji to strengthen the mental health and well-being of adolescents in a school setting. Findings from the study can help develop the literature on this new possible method for delivering mental health and inform processes for a possible future randomized controlled trial.

Methods

Design

The recruitment procedures, sample size, and analyses differed from those planned and published in our protocol paper [35] due to COVID-19 pandemic–related restrictions. Each departure from protocol is documented in the relevant section below.

This feasibility study used a mixed methods design. Adolescents attending New Zealand intermediate and high schools were recruited to participate in this study. They were shown how to use Match Emoji and then asked to play for 15 minutes, 2-3 times a week over 2 weeks (a total of 60 minutes). Analysis of game use, analytic data, and focus groups were held with all participants to explore the acceptability of Match Emoji. The therapeutic potential of the game was assessed by changes in mental health and well-being, which were assessed by 4 validated instruments.

Recruitment

Before the onset of the COVID-19 pandemic, we developed a protocol to outline the guidelines for conducting the trial, including how participants would be recruited [35]. Initially, as 1 local secondary school had expressed interest in participating, we aimed to recruit students between the ages of 13 and 15 years from this school across 2-4 classrooms. However, several teachers had become ill during the recruitment phase, and the secondary school could no longer participate in the study. As such, we approached 2 secondary schools (students aged 12-18 years) and an intermediate school (students aged 12-14 years), which all expressed interest in participating in the study.

In the secondary school, we described the study to an assembly of over 400 students in years 9 and 10 (aged 12-14 years). Those interested in participating in the trial and with access to a smartphone or tablet were asked to take home information, an assent form, and a consent form for their parent or guardian. Of the 42 interested students, only 6 returned both forms. When recruiting participants in each intermediate school, the New Zealand government implemented restrictions on indoor face-to-face gatherings. At this time, indoor gatherings of up to 100 people were allowed. As such, instead of recruiting participants in an assembly, we delivered a 10- to 15-minute face-to-face presentation to students in each classroom, explaining the theory and research underpinning Match Emoji. In total, 39 returned the assent and parental consent forms. Given the primary aims, the inclusion criteria were students aged between 12 and 14 years who had access to a phone that could download Match Emoji and provided written consent from a parent or caregiver.

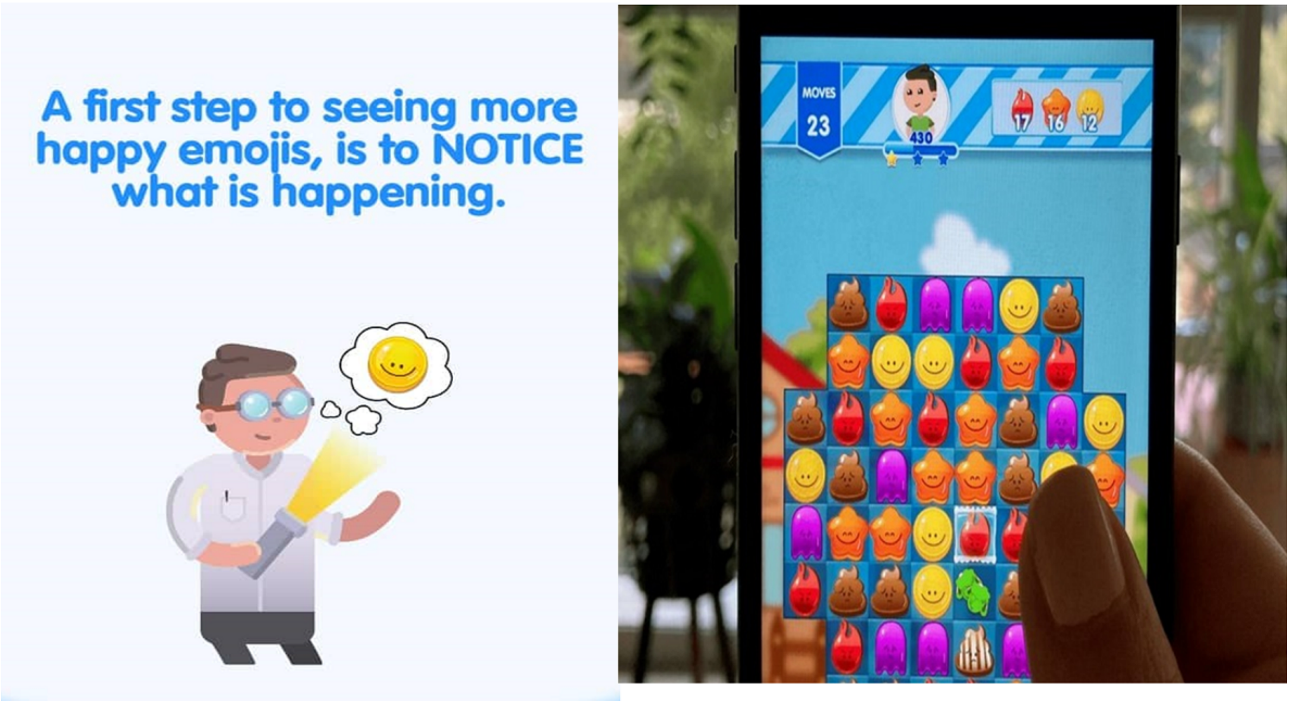
Study Procedure

Consenting participants completed the Child and Adolescent Mindfulness Measure (CAMM), General Help-Seeking Questionnaire (GHSQ), Flourishing Scale (FS), and Revised Children's Anxiety and Depression Scale (RCADS) at baseline. These assessments were completed in groups of 6 in the high school and 30 in the intermediate school. Students completed the questionnaires at their desks and were separated at least a meter apart from each other to protect privacy. Instructions on how to play and install the game were provided, and participants were given an opportunity to ask questions directly or email the primary researcher. Next, they were asked to play Match Emoji 2-3 times a week for a minimum of 15 minutes per session for 2 weeks (ie, a minimum of 60 minutes in total). Questions were repeated after the 2-week intervention period. All participants were invited to participate in focus groups held at each school 2 weeks later (4 weeks after the study began). After the study, *koha* (food and drink) was provided to acknowledge the student's effort. No financial incentives or gifts were provided.

The Intervention

The Match Emoji rationale, content, and processes have previously been described [33]. In brief, the micromessages in Match Emoji are based upon psychological well-being literature, specifically the *Five Ways to Wellbeing* [36]. As seen in Figure 1, these messages appear instead of in-game advertisements and function as prompts. For example, players are encouraged to read the message and practice skills including diaphragmatic breathing, noticing thoughts, or normalizing difficult emotions. In terms of the gameplay, users must identify and match 3 or more similar colored emojis together in rows or squares (a "match-3" game) to earn points. There are 6 different colored and shaped emojis, each representing an emotion or idea. The game has 99 levels, each designed to be completed within a few minutes, with a player advancing to the next level on completion of the current level. The gameplay becomes increasingly challenging as the player progresses.

Figure 1. Screenshot of the Match Emoji video game.



Measures and Outcomes

Demographic data were collected at baseline. Students who reported more than 1 ethnicity were categorized using the New Zealand Census ethnicity prioritization method [37].

Acceptability was assessed by the proportion of approached schools who agreed to participate, the number of participants who are able to download the game on their phone and those who fully participated in the study, and student feedback in focus groups. At week 4 (ie, after the intervention period), all participants were invited to take part in a 45-minute focus group at their school to explore their views of the intervention. Questions included (1) What parts of the game did you like? (2) What parts of the game could be improved? (3) What did you learn from playing the game? (4) Did you try and use any of the ideas from the game, and if so, which ones? and (5) Do you think you will continue to play Match Emoji? A general inductive approach was used to analyze the data from the focus groups [38]. The first author (RP) read participants' responses several times to identify emerging themes and categories from the raw data. A research assistant read through the raw data to ensure the themes reflected the essence of the category. Appropriate quotes that conveyed the key core themes were recorded and integrated into the results. Lastly, game analytics for minutes played and the number of sessions were recorded on the Unity platform [39]. Unity is a secure platform for creating and operating interactive games.

The secondary outcome measures assessing therapeutic potential were changes from the pre- to postintervention (baseline and 2 weeks) time period on mental health and well-being domains calculated from the CAMM, a 10-item instrument measuring acceptance and mindfulness for use with children and adolescents aged 10 between 17 years; the GHSQ, which measures formal help-seeking intentions for nonsuicidal and suicidal problems; the 8-item FS, which measures self-perceived success in important areas such as relationships, self-esteem, purpose, and optimism as a single psychological well-being score; and the RCADS, a 47-item youth self-report questionnaire with subscales, such as separation anxiety disorder and generalized anxiety disorder.

The specific mental health and well-being domains assessed were mindfulness derived from the CAMM, help-seeking from the GHSQ, psychological well-being from the FS, and overall anxiety and depression score from the RCAD. Pretest and posttest summary statistics (mean, median, range, and SD) were computed using the R software (R Foundation for Statistical Computing) developer package. Data were assessed for normality using the Shapiro-Wilk normality test. Since data were not normally distributed, the nonparametric Wilcoxon signed rank test was used to compare the means between pairs of values (pre and post).

Ethical Considerations

This study received ethics approval from the New Zealand Health and Disability Ethics Committee (21/NTA/34) on May 28, 2021. Data was de-identified and all participants provided informed consent. No financial compensation was provided to the study participants.

Results

Participants

Of the 45 adolescents who participated in the study (mean age 12.5, SD 0.33; range 12-14 years), 26 (57%) were New Zealand European, 12 (26%) were Māori, and 7 (15%) were Asian or Pacific. As seen in Table 1, the majority (n=27, 60%) were boys, while 15 (33%) were girls and 3 (6%) were nonbinary.

Table 1. Demographics of participants (N=45).

Characteristic	Value
Age (years)	
Mean (SD)	12.5 (0.33)
Range	12-14
Gender, n (%)	
Boy	27 (60)
Girl	15 (33)
Nonbinary	3 (6)
Ethnicity, n (%)	
Asian	3(6)
Māori	12 (26)
New Zealand European	26 (57)
Pacific	4 (9)

Acceptability

On average, each participant played 7.5 sessions for 24 minutes across the 2 weeks, comprising 180 minutes in total. In addition, data recorded from the focus groups suggested that, on average, participants completed 50 out of the 99 available levels during the 2-week duration of the study. Most participants in the focus groups said they would continue playing the game after completing the study. A total of 38 (84%) participants said they "would" continue to play the game, while 5 (11%) said they "might" continue to play. Only 2 (4%) participants said they would not continue to play the game after the trial. In addition, 36 (80%) reported playing Match Emoji after week 4, and 32 (71%) were still playing after week 6, according to game analytics from the Unity Platform. Findings from the focus groups suggested that participants enjoyed playing Match Emoji for several reasons. First, participants enjoyed the convenience of the game. For instance, many participants reported playing Match Emoji across multiple environments, including waiting rooms at the dentist, bus stops, and long car rides. As no internet connection was needed, participants could access the game whenever they wished. A participant explained, "I could play the game even when there was no Wi-Fi," while another said, "The game was really good when waiting for appointments (be)cause it could distract me for a bit and didn't use up data."

Second, many participants reported enjoyment from playing the game. They described this enjoyment as stemming from game features such as increasing levels of challenge, the variety of emojis, and clear goals: "it was fun (be)cause the game got harder, but you knew what you had to do." While there was some level of challenge, the simplicity of the game allowed students to bypass traditional barriers to CVGs, such as instructional videos. One participant described Match Emoji as a "super easy game to understand and play." A smaller group of participants also provided suggestions about game features. This group appeared to be more frequent users of CVGs, as they provided recommendations based on other games they had played. One participant suggested, “You could add more rewards or more characters and then get more power-ups like Fortnite,” while another recommended, “coins, customization, themed music, and bonus rounds... add stuff like they have in other casual games.”

In general, participants liked the subtle aspect of accessing mental health content. As 1 participant mentioned, “the messages are a nice way of getting mental health information out there that isn’t in your face.” There was a high consensus that they preferred micromessages over typical in-game advertisements. However, some were initially skeptical about their value, “the messages were cringe at first but got way better.”

Of the intermediate and secondary schools approached to participate in the research, only 3 (25%) of the 12 took part in the study. Only 3 (7%) out of the 45 participants could not download Match Emoji onto their phones. In each case, this was because their phone had limited capacity to download the necessary software. All participants completed baseline and follow-up assessments, but several needed clarifications on wording related to the RCADS questionnaire items.

Indicators of Possible Effects

As seen in Table 2, a small positive change was observed on the CAMM (mean difference 1.38, 95% CI –0.03 to 2.79) and on the RCADS (mean difference 3.42, 95% CI –6.84 to –0.001). In focus groups, when asked, “What did you learn from playing the game,” a number of participants answered that playing the game was helpful for their mental health and well-being: “I reckon playing the game for a bit of time was helpful for my mental health (be)cause it took my mind of stuff.” When asked, “Did you try and use any of the ideas from the game, and if so, which ones?” Several participants reported using specific skills suggested in Match Emoji: “Once when I started to think about annoying stuff, I tried the breathing thing, and it was actually pretty helpful,” and “I remember I got pretty mad at my brother and used the noticing a thought approach.”

Table 2. Changes in mental health and well-being indicators of adolescents aged between 12 and 14 years after 2 weeks of playing Match Emoji (N=45).

Outcome	Baseline, mean (SD)	Postintervention, mean (SD)	Mean differences (95% CI)	*P* value
CAMM[a]	22.44 (8.35)	23.82 (8.93)	1.38 (–0.03 to 2.79)	.06
GHSQ[b]	62.89 (21.96)	63.69 (23.30)	0.8 (–2.71 to 4.31)	.65
FS[c]	41.71 (11.58)	40.62 (12.07)	–1.09 (–2.83; 0.66)	.22
RCADS[d]	46.24 (26.39)	42.82 (26.49)	–3.42 (–6.84 to –0.001)	.049

[a]CAMM: Child and Adolescent Mindfulness Measure (mindfulness).
[b]GHSQ: General Help-Seeking Questionnaire (help-seeking).
[c]FS: Flourishing Scale (psychological well-being).
[d]RCADS: Revised Children’s anxiety and Depression Scale (overall anxiety and depression).

Discussion

Overview

In this study, we found that a CVG with psychological well-being concepts (Match Emoji) was a new and engaging mechanism of change that provided a promising indication of possible therapeutic impact. Most participants played more often and for a longer period than was requested for the study. Indeed, most participants continued to play in week 4. Small improvements in mindfulness assessed by CAMM and a small decrease in overall anxiety and depression assessed by RCADS were recorded. Given these promising changes, participants may have learned skills related to reducing their level of anxiety through playing Match Emoji. The findings of this small open feasibility trial indicate that the Match Emoji CVG was an acceptable way to support mental health and well-being in adolescents aged between 12 and 14 years.

Participants reported a high level of acceptability with Match Emoji, as evidenced by the game analytics, qualitative feedback, and the large portion of participants who were still playing the game in weeks 4 and 6. The percentage of participants who stated they continued to play Match Emoji even after week 6 of the study (n=32, 71%) is contrary to the poor retention rate typically found across the range of digital interventions. Real-world data on user engagement with popular mental health apps suggest that a small portion of users stay engaged with digital health interventions [16]. For example, once a health app is downloaded, approximately 4% of users continue to use the app after 15 days [24]. It is possible that the ongoing consultation with end users from the beginning of the development of Match Emoji, the simplicity with which CVGs can be played “on the go,” across environments with no Wi-Fi, and how playing CVGs fits with adolescents’ current behavior patterns may have been attributed to the high level of acceptability and engagement. That is, as many adolescents already play CVGs [32], there is less effort required to learn and change existing ways of engaging with technology. Data from the focus groups corroborated these findings. More specifically, participants mentioned they enjoyed playing for short periods across environments in comparison to computer games or those mobile phone games that require data to access. This is consistent with our previous work [32] and research [19], which suggests young adolescents tend to prefer brief therapeutic encounters. Moreover, Match Emoji enables large portions of the population to receive the same content irrespective of their proficiency with gaming or access to the internet, addressing a significant barrier to equity and engagement with DMHIs [11].

Our finding that most participants preferred micromessages over typical in-game advertisements is consistent with research assessing how in-game advertising in the form of short videos is distracting and can lead to disengagement, particularly among young people who often have a relatively short attention span [40,41]. Although paid versions can avoid advertisements, young people are reluctant to pay for them [41]. Thus, Match Emoji represents an opportunity for public health interventions to provide appealing free CVGs that replace the advertising with health-related micromessaging that is not distracting, annoying, or potentially harmful, as is the case with in-game advertising.

Similarly, diverse preferences were found with gamification elements of Match Emoji. The various preferences toward micromessages and game features among participants are consistent with the literature that suggests adolescents have different opinions about the type of DMHI they are attracted to [42,43]. Thus, while some adolescents may be frequent CVG users and interested in gamification elements, others may be less focused on these features and more attracted to learning about mental health and well-being [44]. In essence, opportunities to embed therapeutic processes within game elements are plentiful when researchers and game developers collaborate and are creative.

The protocol and implementation of this study were completed during the COVID-19–lined social distancing practices, which resulted in frequent changes to the restrictions on the size of inside gatherings and how educational facilities operated. Apart from the implications of the pandemic, 3 participants in the study could not download Match Emoji. This was because their phones lacked the storage capacity required to download the latest software and the game. Future research could use methods to compress digital mental health apps such as Match Emoji. In this way, the size of the app may better align with the capacity of users' technology. In addition, some participants struggled to understand several questions on the RCADS. These questions were discussed in more detail with each participant to ensure they understood the meaning of each one. Despite these challenges, no significant issues occurred with conducting the study in a primary and intermediate school context.

Limitations

There are limitations to this study; these include departures from the protocol due to COVID-19 impacts, which resulted in a small exploratory open trial only. There were also limited resources to conduct the study; this meant that the first author (RP) introduced the game to participants, led the recruitment process, supported the completion of assessments before and after playing the game, and facilitated the discussions about the game. Thus, the interpretation of the students' feedback could be overly positive, and participants' opinions and thoughts could be influenced by social pressures, including normative and informative conformity.

Further, the self-assessment outcome measures relied on the comprehension skills of young participants. While some participants raised their hands when unsure of a question, others may have merely guessed. Nevertheless, the 4 assessments appeared to be easy to implement in a short amount of time. Third, when recruiting participants at the secondary school, only 6 (14%) out of the 42 participants who signed the assent form returned their parental consent form for reasons unknown, suggesting that a different process is needed to recruit older adolescents for future trials. Lastly, students were not recruited based on their level of mental distress. Therefore, the results may have been affected by floor effects, whereby their mental health and well-being scores were already good or optimal and thus unlikely to improve any further. Despite these challenges and preliminary results, these findings are of interest as this is the first study to assess the feasibility of a co-designed CVG with psychological well-being concepts.

Conclusion

Findings from this feasibility study suggest that Match Emoji, the purpose-built CVG with brief mental health messages, is promising as an acceptable and feasible approach for young adolescents. Future research should test clinical impacts through a randomized controlled trial. More broadly, the research also highlights the possibility of CVGs as a novel mechanism of delivering simple mental health and well-being messages.

Acknowledgments

We would like to acknowledge the young people who participated in the study and the teachers who helped with the recruitment process. This study was funded by Cure Kids (innovation seed fund: 3918).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

RP and TF were involved in the conceptualization of the game and study. LTM and JM were involved in the methodology and formal analysis. All authors were involved in writing the original draft and reviewing and editing the paper. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

TF is a codeveloper of SPARX, a computerized cognitive behavioral therapy program for adolescent depression. The intellectual property for SPARX is owned by Uniservices at the University of Auckland, and codevelopers can benefit financially from the licensing of SPARX outside of New Zealand.

References

1. Clark TC, Le Grice J, Moselen E, Fleming T, Crengle S, Tiatia-Seath J, et al. Health and wellbeing of Māori secondary school students in New Zealand: trends between 2001, 2007 and 2012. Aust N Z J Public Health 2018;42(6):553-561 [FREE Full text] [doi: 10.1111/1753-6405.12839] [Medline: 30370961]
2. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. JAMA Pediatr 2021;175(11):1142-1150 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.2482] [Medline: 34369987]
3. Fleming T, Ball J, Bavin L, Rivera-Rodriguez C, Peiris-John R, Crengle S, et al. Mixed progress in adolescent health and wellbeing in Aotearoa New Zealand 2001–2019: a population overview from the Youth2000 survey series. J R Soc N Z 2022;52(4):426-449 [FREE Full text] [doi: 10.1080/03036758.2022.2072349]
4. Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. Soc Psychiatry Psychiatr Epidemiol 2019;54(8):987-996 [FREE Full text] [doi: 10.1007/s00127-019-01697-8] [Medline: 30929042]
5. Sutcliffe K, Ball J, Clark TC, Archer D, Peiris-John R, Crengle S, et al. Rapid and unequal decline in adolescent mental health and well-being 2012-2019: findings from New Zealand cross-sectional surveys. Aust N Z J Psychiatry 2023;57(2):264-282. [doi: 10.1177/00048674221138503] [Medline: 36453262]
6. Fortuna LR, Brown IC, Woods GGL, Porche MV. The impact of COVID-19 on anxiety disorders in youth: coping with stress, worry, and recovering from a pandemic. Child Adolesc Psychiatr Clin N Am 2023;32(3):531-542 [FREE Full text] [doi: 10.1016/j.chc.2023.02.002] [Medline: 37201965]
7. Cleare A, Pariante CM, Young AH, Anderson IM, Christmas D, Cowen PJ, et al. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2008 British Association for Psychopharmacology guidelines. J Psychopharmacol 2015;29(5):459-525. [doi: 10.1177/0269881115581093] [Medline: 25969470]
8. Skehan B, Davis M. Aligning mental health treatments with the developmental stage and needs of late adolescents and young adults. Child Adolesc Psychiatr Clin N Am 2017;26(2):177-190. [doi: 10.1016/j.chc.2016.12.003] [Medline: 28314449]
9. Neufeld SAS, Jones PB, Goodyer IM. Child and adolescent mental health services: longitudinal data sheds light on current policy for psychological interventions in the community. J Public Ment Health 2017;16(3):96-99 [FREE Full text] [doi: 10.1108/JPMH-03-2017-0013] [Medline: 29721033]
10. Mojtabai R, Olfson M, Sampson NA, Jin R, Druss B, Wang PS, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. Psychol Med 2011;41(8):1751-1761 [FREE Full text] [doi: 10.1017/S0033291710002291] [Medline: 21134315]
11. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
12. Werntz A, Amado S, Jasman M, Ervin A, Rhodes JE. Providing human support for the use of digital mental health interventions: systematic meta-review. J Med Internet Res 2023;25:e42864 [FREE Full text] [doi: 10.2196/42864] [Medline: 36745497]
13. Garrido S, Millington C, Cheers D, Boydell K, Schubert E, Meade T, et al. What works and what doesn't work? a systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019;10:759 [FREE Full text] [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
14. Riboldi I, Cavaleri D, Calabrese A, Capogrosso CA, Piacenti S, Bartoli F, et al. Digital mental health interventions for anxiety and depressive symptoms in university students during the COVID-19 pandemic: a systematic review of randomized controlled trials. Rev Psiquiatr Salud Ment 2023;16:47-58 [FREE Full text] [doi: 10.1016/j.rpsm.2022.04.005] [Medline: 35755491]
15. Velleman S, Stallard P, Richardson T. A review and meta-analysis of computerized cognitive behaviour therapy for the treatment of pain in children and adolescents. Child Care Health Dev 2010;36(4):465-472. [doi: 10.1111/j.1365-2214.2010.01088.x] [Medline: 20412143]
16. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. J Med Internet Res 2018;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
17. Li J. Digital technologies for mental health improvements in the COVID-19 pandemic: a scoping review. BMC Public Health 2023;23(1):413 [FREE Full text] [doi: 10.1186/s12889-023-15302-w] [Medline: 36859184]

18. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021;20(3):318-335 [FREE Full text] [doi: 10.1002/wps.20883] [Medline: 34505369]
19. Baumel A, Fleming T, Schueller SM. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. J Med Internet Res 2020;22(10):e20631 [FREE Full text] [doi: 10.2196/20631] [Medline: 33118946]
20. Fleming T, Dewhirst M, Haenga-O'Brien A, Chinn V, Ormerod F, Kafatolu D, et al. Digital Tools for Mental Health and Wellbeing: Opportunities & Impact. Findings from the Literature and Community Research. Wellington: Te Hiringa Hauora | Health Promotion Agency; 2021.
21. McCashin D, Coyle D, O'Reilly G. Pesky gNATs for children experiencing low mood and anxiety—a pragmatic randomised controlled trial of technology-assisted CBT in primary care. Internet Interv 2022;27:100489 [FREE Full text] [doi: 10.1016/j.invent.2021.100489] [Medline: 35024346]
22. Fleming TM, Stasiak K, Moselen E, Hermansson-Webb E, Shepherd M, Lucassen M, et al. Revising computerized therapy for wider appeal among adolescents: youth perspectives on a revised version of SPARX. Front Psychiatry 2019;10:802 [FREE Full text] [doi: 10.3389/fpsyt.2019.00802] [Medline: 31824344]
23. Fleming TM, de Beurs D, Khazaal Y, Gaggioli A, Riva G, Botella C, et al. Maximizing the impact of e-therapy and serious gaming: time for a paradigm shift. Front Psychiatry 2016;7:65 [FREE Full text] [doi: 10.3389/fpsyt.2016.00065] [Medline: 27148094]
24. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
25. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
26. Fleming T, Poppelaars M, Thabrew H. The role of gamification in digital mental health. World Psychiatry 2023;22(1):46-47 [FREE Full text] [doi: 10.1002/wps.21041] [Medline: 36640378]
27. What are casual games in the video industry? Consumer Acquisition. 2020. URL: https://www.consumeracquisition.com/faq/what-are-casual-games [accessed 2024-01-19]
28. Desai V, Gupta A, Andersen L, Ronnestrand B, Wong M. Stress-reducing effects of playing a casual video game among undergraduate students. Trends Psychol 2021;29(3):563-579 [FREE Full text] [doi: 10.1007/s43076-021-00062-6]
29. Brand JE, Jervis J, Huggins P, Wilson T. Digital New Zealand 2020: the power of games. Interactive Games and Entertainment Association (IGEA). 2019. URL: https://igea.net/wp-content/uploads/2019/09/DNZ20-Final-Report-2019.pdf [accessed 2014-01-19]
30. Brian M. Angry birds gamers spend 200 million minutes playing each day. The Next Web. 2011. URL: https://thenextweb.com/news/angry-birds-gamers-spend-200-million-minutes-playing-each-day [accessed 2024-01-19]
31. Pine R, Fleming T, McCallum S, Sutcliffe K. The effects of casual videogames on anxiety, depression, stress, and low mood: a systematic review. Games Health J 2020;9(4):255-264 [FREE Full text] [doi: 10.1089/g4h.2019.0132] [Medline: 32053021]
32. Pine R, Sutcliffe K, McCallum S, Fleming T. Young adolescents' interest in a mental health casual video game. Digit Health 2020;6:1-7 [FREE Full text] [doi: 10.1177/2055207620949391] [Medline: 32944270]
33. Pine R, Te Morenga L, Olson M, Fleming T. Development of a casual video game (Match Emoji) with psychological well-being concepts for young adolescents. Digit Health 2021;7:1-9 [FREE Full text] [doi: 10.1177/20552076211047802] [Medline: 34691753]
34. Wolcott MD, Lobczowski NG. Using cognitive interviews and think-aloud protocols to understand thought processes. Curr Pharm Teach Learn 2021;13(2):181-188. [doi: 10.1016/j.cptl.2020.09.005] [Medline: 33454077]
35. Pine R, Mbinta J, Te Morenga L, Fleming T. A casual video game with psychological well-being concepts for young adolescents: protocol for an acceptability and feasibility study. JMIR Res Protoc 2021;10(8):e31588 [FREE Full text] [doi: 10.2196/31588] [Medline: 34387558]
36. Farrier A, Dooris M, Froggett L. Five ways to wellbeing: holistic narratives of public health programme participants. Glob Health Promot 2019;26(3):71-79. [doi: 10.1177/1757975917732352] [Medline: 29125391]
37. Pathways and transitions. Ministry of Education. 2017. URL: https://www.education.govt.nz/communities-of-learning/transitions/ [accessed 2024-01-19]
38. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016;27(2):237-246. [doi: 10.1177/1098214005283748]
39. Schardon L. What is Unity? a guide for one of the top game engines. GameDev Academy. 2022. URL: https://gamedevacademy.org/what-is-unity/ [accessed 2024-01-29]
40. Lorenz-Spreen P, Mønsted BM, Hövel P, Lehmann S. Accelerating dynamics of collective attention. Nat Commun 2019;10(1):1759 [FREE Full text] [doi: 10.1038/s41467-019-09311-w] [Medline: 30988286]

41. Yang Z, Wu Y, Lu C, Tu Y. Effects of paid search advertising on product sales: a Chinese semantic perspective. J Mark Manag 2020;36(15-16):1481-1504. [doi: 10.1080/0267257x.2020.1765001]
42. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
43. Ribanszki R, Fonseca JAS, Barnby JM, Jano K, Osmani F, Almasi S, et al. Preferences for digital smartphone mental health apps among adolescents: qualitative interview study. JMIR Form Res 2021;5(8):e14004 [FREE Full text] [doi: 10.2196/14004] [Medline: 34128814]
44. Fleming T, Merry S, Stasiak K, Hopkins S, Patolo T, Ruru S, et al. The importance of user segmentation for designing digital therapy for adolescent mental health: findings from scoping processes. JMIR Ment Health 2019;6(5):e12656 [FREE Full text] [doi: 10.2196/12656] [Medline: 31066705]

Abbreviations

CAMM: Child and Adolescent Mindfulness Measure
CBT: cognitive behavioral therapy
CVG: casual video game
DMHI: digital mental health intervention
FS: Flourishing Scale
GHSQ: General Health-Seeking Questionnaire
RCADS: Revised Children’s Anxiety and Depression Scale

Edited by T Leung; submitted 21.02.23; peer-reviewed by N Ludin, S Nilsson; comments to author 05.05.23; revised version received 15.08.23; accepted 21.08.23; published 12.02.24.

Please cite as:
Pine R, Mbinta J, Te Morenga L, Fleming T
A Novel Casual Video Game With Simple Mental Health and Well-Being Concepts (Match Emoji): Mixed Methods Feasibility Study
JMIR Serious Games 2024;12:e46697
URL: https://games.jmir.org/2024/1/e46697
doi:10.2196/46697
PMID:38345848

Original Paper

A Smartphone-Gamified Virtual Reality Exposure Therapy Augmented With Biofeedback for Ailurophobia: Development and Evaluation Study

Ali Khaleghi[1], PhD; Abbas Narimani[2*], MSc; Zahra Aghaei[3*], PhD; Anahita Khorrami Banaraki[4,5], MD, PhD; Peyman Hassani-Abharian[4,5], MD, PhD

[1]Iranian Light Source Facility, Institute for Research in Fundamental Sciences, Tehran, Iran

[2]Department of Computer Engineering, Imam Khomeini International University, Qazvin, Iran

[3]Department of Computer Engineering, Bu-Ali Sina University, Hamedan, Iran

[4]Brain and Cognition Clinic, Institute for Cognitive Science Studies, Tehran, Iran

[5]Department of Cognitive Psychology and Rehabilitation, Institute for Cognitive Science Studies, Tehran, Iran

[*]these authors contributed equally

Corresponding Author:
Ali Khaleghi, PhD
Iranian Light Source Facility
Institute for Research in Fundamental Sciences
Opposite the Araj, Artesh Highway, Aghdassieh
Tehran, 19395-5746
Iran
Phone: 98 9121003006
Email: ali.khaleghi.ir@gmail.com

Abstract

Background: To the best of our knowledge, no specialized research has been conducted to address ailurophobia (fear of cats) in Iran or globally. This has driven our project, along with the prevalence of ailurophobia and the absence of a gamified virtual reality exposure therapy (VRET) that incorporates affordable and easily accessible biofeedback (BF) tools. We hypothesize that a gamified VRET augmented with BF will yield more positive effects than a similar device lacking BF.

Objective: This study primarily focuses on the development and preliminary evaluation of a smartphone-gamified VRET integrated with BF, targeting animal phobia, with a specific case study on ailurophobia. The secondary objectives are using affordable and readily available BF found in devices such as smart bands and smartwatches and creating a mobile virtual reality gamified app to improve patients' adherence to treatments while simultaneously enhancing the app's accessibility, scalability, and outreach.

Methods: Evaluations encompassed 3 methods. First, we identified the tool's potential positive effects on phobia interventions, exploring 4 effects: intrinsic motivation, simulation of fearful situations, management of stressful circumstances without therapists' presence and mitigation of catastrophic thoughts, and preliminary effects on ailurophobia treatment. Participants were divided into BF and non-BF groups. Second, we gathered user preferences and opinions about the treatment. Third, we conducted heuristic evaluations using 44 heuristics from existing system usability scales assessing user interfaces, virtual reality platforms, and video games' playability. To interpret the data, mean scores; ANOVA, single factor; and ANOVA, 2-factor with replication were used. A total of 29 individuals were identified, of which 10 met the eligibility criteria or were accessible.

Results: The smartphone-gamified VRET augmented with BF exhibited better results on the identified effects compared with the non-BF version and contributed to normalizing encounters with cats. Moreover, 41 of the 44 heuristics achieved a percentage above 62%, indicating its potential as a therapeutic product and its ability to enhance patient adherence to treatments. Patient preferences on the treatment and its strengths and weaknesses were provided for further improvement.

Conclusions: The tool has the potential to evolve into a comprehensive solution by incorporating various types of cats and their behaviors, simulating environments in which they are commonly found, and enhancing its appeal through an increased sense of adventure without inducing unrealistic fears. By adapting fear elements, the game can be tailored to treat various animal phobias.

Phobia-focused games should avoid action and combat scenarios to prevent reinforcement of fear responses. After rigorous evaluation, further exploration is required to provide remote use beyond clinical settings.

(JMIR Serious Games 2024;12:e34535) doi:10.2196/34535

KEYWORDS

animal phobia; specific phobia; ailurophobia; cat phobia; biofeedback; smartphones; virtual reality; gamification; mobile phone

Introduction

Specific Phobia and Available Therapies

Specific phobia is the most common anxiety disorder, with a lifetime prevalence of 12.5% [1]. It is characterized by an extreme and persistent fear of a specific object or situation [2], leading to substantial disruptions in daily life and heightened anxiety. Many individuals restructure their lives to evade their fears over extended periods [3,4]. Prolonged phobia detrimentally affects academic, social, and family aspects, compromising overall quality of life [5]. Situational (eg, fear of enclosed spaces and flying), natural environment (eg, fear of heights and storms), animal (eg, fears of snakes, spiders, and cats), and blood or injection or injury (eg, fears of medical procedures and seeing blood) fears are subtypes of specific phobias, with animal and natural environment phobias being more prevalent [3].

Phobia interventions are categorized into exposure therapies (eg, direct in vivo exposure, systematic desensitization, imaginal exposure, and virtual reality [VR]) and nonexposure approaches (eg, cognitive therapy and progressive muscle relaxation). There has been a trend toward adopting brief, intensive, or concentrated treatments to manage anxiety [5]. Among the available treatments, exposure therapies are the most commonly used approach for specific phobias [6]. However, although specific phobias are highly treatable, only 31% of patients seek treatment and, among those, only 43.4% seek mental health services [3]. Moreover, some patients might be unable to complete the treatment because of severe reactions, resulting in an attrition rate of 45% [3,7]. In total, 3 main factors contribute to this percentage [7]: (1) perceiving treatments as highly aversive and frightening; (2) the need to visit clinics throughout the treatment, causing relationship and ethical issues, such as perceived cruelty when therapists intentionally evoke fear; and (3) the lack of appealing treatments.

Gamified VR Exposure Therapies Augmented With Biofeedback

To overcome the limitations of exposure therapy methods, incorporating new technologies becomes imperative. Gamification, VR, and biofeedback (BF) are promising options. However, our research indicates that few studies have simultaneously used these technologies for specific phobias. Virtual reality exposure therapy (VRET) uses 360° computer-generated simulations [8,9] similar to traditional exposure therapies [2]. Meta-analyses have shown that VRETs are effective and their performance can rival standard exposure therapies [2]. VR's application in cognitive impairment, anxiety disorders, pain management, phobias, posttraumatic stress disorder, rehabilitation, and eating disorders, among others, has surged because of its immersive realism [8,10]. To treat phobias, VR is a safer, less embarrassing, and cost-effective solution by simulating fear-inducing situations in a controlled environment [8,9,11]. However, VR alone may not address all exposure therapy disadvantages, and enhancing the attractiveness of VRETs is crucial for treatment success. Researchers have explored the potential of gamified VRETs in treating phobias [2,12,13]. Gamification, a strategy derived from video game–based approaches, has proven successful in achieving serious objectives across various fields, including the workplace [14], education [15], marketing [16], mental health [17-19], learning disabilities [20,21], and lazy eye treatment [22]. The primary inherent feature of digital games is their high-level motivational potential [23]. Video games' appeal, engagement, and effectiveness encourage players and frequent use [18]. Attractiveness is beneficial for overcoming people's reluctance to seek treatment, broadening the reach of gamified interventions [18]. The engaging nature of gamification enhances users' experiences, as players are driven to win, explore stories, and ultimately reduce attrition rates [12,18,24]. The effectiveness aspect offers opportunities for achieving serious objectives such as behavior changes [18]. In a gamified product, elements such as scores, badges, and levels are integrated from games into nongame contexts, while not necessarily offering a complete gaming experience [18,25].

Human emotion recognition sensors or BF is another technology that can enhance gamified interventions. This technology serves 2 crucial purposes. First, it boosts their level of attractiveness by leveraging a strategy commonly used in video games to increase engagement [26]. Second, it addresses some of the limitations of traditional methods by potentially reducing or eliminating the need for therapists' constant presence. These sensors work by measuring various body parameters or electrical impulses in the nervous system to identify different emotions and track their changes [27]. Common techniques include electroencephalography, skin resistance measurements, blood pressure, heart rate (HR), eye activity, and motion analysis. With advancements in chipset manufacturing, BF has become more accessible, portable, efficient, and affordable. Users can easily access their data, thereby enabling self-regulation and monitoring. These technologies are incorporated into smart wristbands and watches to help individuals regulate anxiety in their daily lives. BF therapies have shown positive effects in treating conditions such as migraines [28] and attention-deficit/hyperactivity disorder in children [29].

Objectives

The primary aim of this study was to develop and conduct a preliminary assessment of smartphone-gamified VRET augmented with BF for the treatment of cat phobia (ailurophobia). We hypothesize that this tool will outperform gamified VRET without BF in various aspects. Limited evidence exists on animal phobia in Iran, particularly ailurophobia. Observations at the Cognitive and Brain Clinic in Tehran revealed a substantial prevalence of this phobia, as reported by the fourth author, who is a cognitive expert and psychologist attending to cat phobia patients daily. Owing to the abundance of cats in most Iranian cities, encounters are inevitable, resulting in daily challenges for patients walking on the streets and alleys. The secondary objectives were as follows:

1. Using affordable and accessible BF tools in devices such as smart bands and smartwatches to serve as both BF and a game mechanic, enhancing engagement and efficacy.
2. Developing a mobile VR–gamified app to enhance patients' adherence to phobia treatment and expand the app's accessibility, scale, and reach.

To evaluate the effectiveness of the tool, its potential positive effects on phobia interventions were examined. The tool's impact on the effects was examined by dividing the participants into BF and non-BF groups. In addition, we considered the playability and usability aspects of the tool, along with patients' preferences, to optimize its performance and enhance usability for future improvements.

Methods

Design and Development

Our primary objective was to present fear elements indirectly to the player, ensuring that interacting or not interacting with them would not affect the game's progress. The secondary objective was to create a general game design model that could be easily customized for specific phobias, particularly animal phobias. During the initial game development meeting, 2 game design experts (a game designer and a gamification expert) collaborated with 2 cognitive science experts (one of whom also specialized in cognitive games). They engaged in a discussion regarding the essential components required to simulate stress. Size, color, and behavior of the stimuli were introduced as fundamental elements for replicating the desired scenarios. The game team then devised the game stages using a maze design. In the second meeting, cognitive experts suggested simplifying the design to accommodate players of all ages. As it involved memory and problem-solving, it was rejected, leading to a more straightforward game plan that focused on finding lost objects in a park. In the third session, this plan received approval and was tested on a woman aged 40 years with cat phobia, who was selected based on her self-reported fear of cats. She stated, "I experienced a lot of fear during playing." In the fourth session, minor adjustments were made to the game. In total, 2 psychologists from Tehran University found the initial voice of the guide annoying, thereby hindering patient motivation. The overall view of the sessions is presented in Figure 1.

Our game's storyline was inspired by the "Hot and Cold" game. One group hides an object, and the other group should find it using verbal clues such as "colder" as they move away and "hotter" as they get closer. The experience is similar to that of a park with diverse paths. Players are on a quest to discover diamonds concealed within treasure boxes, all while walking along these pathways. Each game session comprises 4 short yet consecutive levels. At each level, players must determine their distance from each box by perceiving changes in the sound consistently played. Moreover, a hint ribbon shows the player's distance to the box for increased engagement. After locating the box, players must stay in front of it for a specific duration to open it, with the time increasing at later levels. The players must open the previous level's box to unlock the next challenge.

Regarding authors' concerns about spreading the game to individuals with phobias, smartphones were chosen as the primary platform. Using smartphones as a VR tool requires affordable mobile VR glasses, which are significantly cheaper than other options such as Oculus or HTC VR. The primary challenge in mobile VR is the user interaction limitations. The game uses Gaze, a pointer on the screen that allows users to interact through head movements, thus providing a mouselike experience. In addition, the game incorporates joysticks connected to the phone, thus offering more interactive possibilities.

The intensity of the fear elements must be balanced based on the game's progress and levels, as in previous studies [7,13]. The escalation of fear stimuli is determined by the following features, each with its own difficulty level. Moreover, these elements can be further amplified in tandem with player's advancement.

1. Visual elements: the fear-triggering elements include cat photos, fantasy cat models, low-poly cat models with minimal details, and high-poly cat models that closely resemble real cats. According to experts, individuals who fear something may also react to objects and shapes that resemble it. For instance, people who are afraid of cats might experience fear when encountering a cat picture or a furry object. This phenomenon is directly related to the degree and intensity of the individual's fear [30]. Figure 2 illustrates the game environment.

2. Fear elements' sound: the scary elements vary from silent to those with terrifying sounds. In intense situations, cats produce specific sounds that could heighten anxiety. The timing of when the sound is played also adds to the diversity. For example, when players are near a cat, the sounds it emits could intensify their fear.

4. The quality of fear elements' behaviors: studying the behavior quality of a stimulus is under investigation [31,32]. A cat jumping from one point to another evokes more fear than a cat simply standing still. Various animations were designed for 3D model cats. The fantastical cat playfully turns its head and randomly spins around. The low-poly cat remains stationary, solely turning its head. In the final level, the high-poly cat features 3 different animations. The first 2 animations portrayed the cat at rest, either shaking its tail and head or cleaning its paws. The third animation involves the cats' walking behavior.

5. Interactable elements: fear elements that respond to the player's presence add a sense of authenticity to the game, elevating immersion and allowing for anxiety manipulation. Cats may react by turning, approaching, or fleeing when a player gets closer. Both low-poly and high-poly cats respond to the

player's presence. The manner in which the elements react was also classified. Although the fantasy cat remains unresponsive, the low-poly cat acknowledges players by turning their heads and looking at them when they enter the zone. At the last level, the realistic cat not only faces the players but also follows them until they exit from its zone.

6. Fear elements' size: element size could also amplify fear. In the final level, some cats are larger, preparing players to confront more intimidating situations.

7. Fear elements' numbers**:** seeing numerous cats creates a feeling of being surrounded, indirectly encouraging players to confront their fear. As players approach the boxes, fear intensifies, peaking around those areas.

The quantity and type of elements can be customized based on players' preferences and conditions (Figure 3), which is beneficial when they need to concentrate on a specific scary element. Furthermore, a player who does not fear an element can eliminate it from the game.

A VR Android game was developed using the Unity game engine, incorporating the Amazfit Bip smartwatch. In anxiety treatments, HR variability is a common BF technique for stress management [33]. However, because of limitations in receiving these signals through conventional smartwatches and wrist bands, HR was used instead of HR variability. HR data are accessible in smartwatches through Bluetooth low energy technology [33]. A plug-in for the Unity3D game engine was implemented to integrate smartwatch data into the game. The player's HR was incorporated into the experience as a game mechanic. The HR was displayed on the corner of the screen. A total of three BF techniques were used in this study: (1) displaying changes in players' bodies to inform and manage anxiety [33]; (2) keeping HR within specific limits allows players to earn the game's prize, a diamond, promoting relaxation skills for stressful situations [34]; and (3) maintaining a low HR for a period allowed players to open boxes and collect more diamonds [33,35].

Figure 1. The overall view of sessions between the game designers and other related experts.

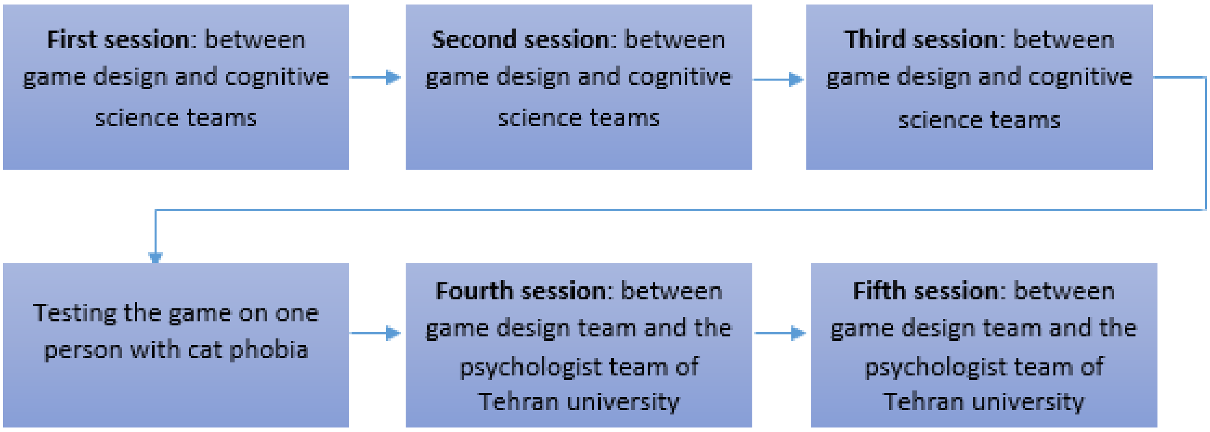


Figure 2. Screenshots of the game's environment with all its cat types.

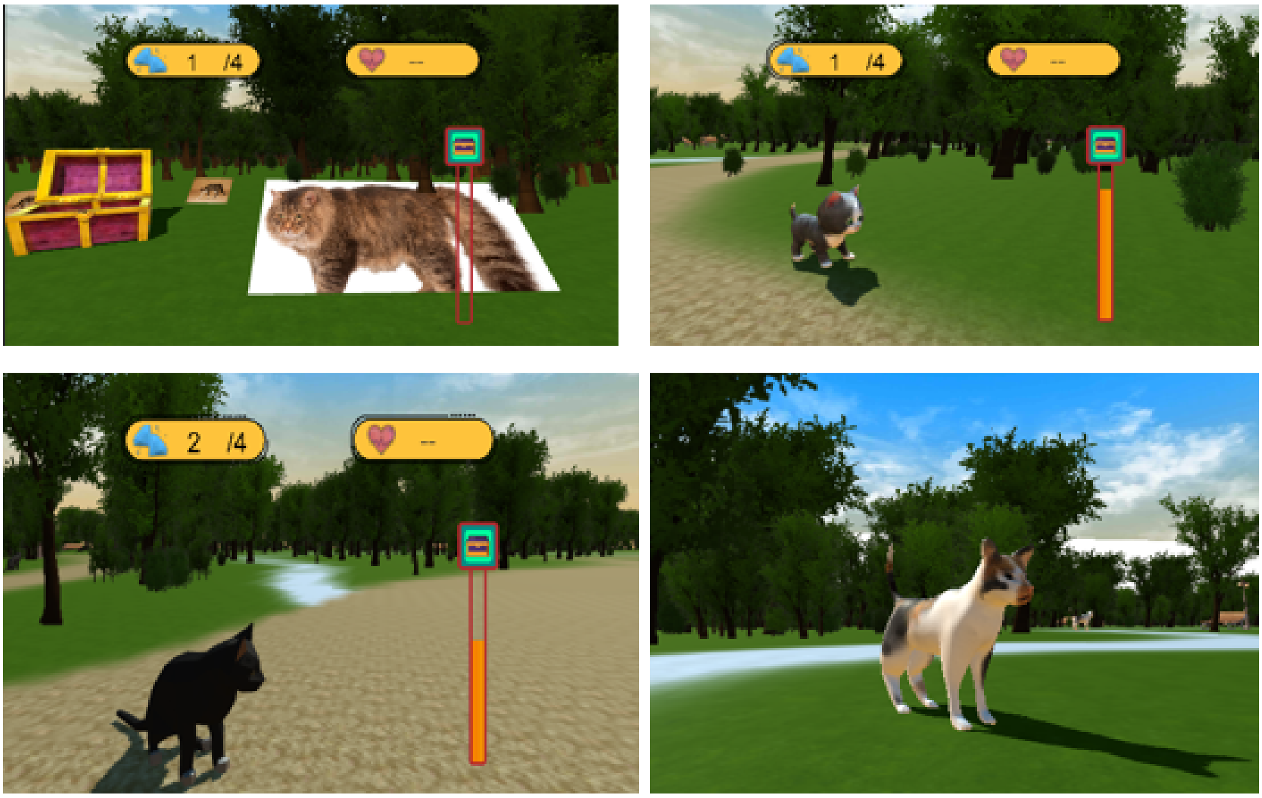


Figure 3. Images of the level customization menu.

Trial Design, Participants, and Procedure

Overview

In total, three methods were used to evaluate the tool: (1) identifying its potential positive effects it could have in phobia interventions. To assess the game's impact on these effects, participants were divided into BF and non-BF groups, with the only differences being the use of smartwatches; (2) gathering user preferences about the treatment; and (3) considering the tool's playability and usability aspects for subsequent optimization and improved usability.

Ethical Considerations

This study was approved by the Research Ethics Committees of the Institute for Cognitive Science Studies (IR.UT. IRICSS.REC.1401.047). Informed consent was obtained from participants. They had the freedom to withdraw from the study at any time. The participants' data were anonymized. To compensate for time, participants were informed that a smartphone-compatible version of the game would be provided free after its finalization.

Participants

The snowball method was used for recruitments. One attractive advertisement was prepared in Farsi and shared within various working, educational, and family groups on Instagram and

WhatsApp. Receivers were asked to help by sharing the advertisement with their own groups. Recruitment took place from September 8 to October 14, 2022, in 2 provinces in Iran: Lorestan and Tehran. Each test session lasted up to 3 hours, and the participants had the flexibility to choose the test location. Random assignment was used to allocate the participants to the study arms.

Inclusion criteria were (1) providing informed written consent, (2) understanding and reading Persian, and (3) scoring ≥55 on the Fear of Cats Questionnaire (FCQ). Exclusion criteria were (1) currently receiving psychological treatment for ailurophobia; (2) having another severe mental disorder (alcohol or substance abuse, psychotic disorder, dementia, or bipolar disorder); (3) diagnosed with a severe personality disorder; (4) experiencing depressive symptoms or suicidal ideation; (5) heart disease; (6) vision or balance problems affecting the VR experience; (7) pregnancies exceeding 3 months; and (7) fear of cats only in a few and exceptional cases. An image of participants is presented in Figure 4.

Figure 4. Depiction of participants.

Identifying the Positive Effects of the Gamified VRET Augmented With BF

Effect 1: Intrinsic Motivation

One primary positive effect that the app could have on phobia interventions is its ability to enhance intrinsic motivation. By incorporating gamification, VR, and BF, the app effectively motivated patients to actively engage in their treatment. We hypothesize that combining gamified VRET with BF will significantly increase motivation compared with a similar tool without BF.

To assess their impact on intrinsic motivation, we used subjective and objective measures. After each level, participants completed a 10-item questionnaire that was previously used to evaluate subjective engagement [36-39].

Participants played a minimum of 5 levels and completed the intrinsic motivation questionnaire after each level, except for the first. First-level data were excluded because of participants' unfamiliarity with the experience. For the training step, the picture level of the game with 8 cat pictures was predetermined. The other settings regarding the type and quantity of stimuli for mandatory games are as follows:

- Game 1: fantasy model with 13 cats
- Game 2: low-poly model with 19 cats
- Games 3 and 4: high-poly model with 23 cats

We deliberately chose the last 2 steps in the same manner to examine the impact of repetitive tasks on the participants.

After the mandatory games, participants had the option to play the game for up to 4 additional times. During the voluntary sessions, participants were allowed to choose the type and number of cats, but the number of cats had to be selected in ascending order. In these sessions, we used a shorter version of the intrinsic motivation questionnaire with only 5 items, as used in Lumsden et al [36].

Effect 2: Simulating Fearful Situations

For phobia treatments to be effective, the game should evoke fear among individuals. To evaluate this, both groups were asked to rate their anxiety levels on a scale from 1 ("no anxiety") to 10 ("extreme anxiety") after any mandatory and voluntary sessions (except the first level).

Effect 3: Controlling Stressful Circumstances, Eliminating Therapists' Presence, and Mitigating Catastrophic Thoughts

The game enables participants to implicitly learn relaxation techniques while confronting their fears. The box-opening mechanism involves standing in front of the box for gradually increasing durations. This combined approach, along with BF, has the potential to reduce the need for therapists' presence.

After the experiments, the participants were asked two questions: (1) How well do you think you could manage your stress when dealing with a real cat after using the gamified app? (2) To what extent can our game eliminate the need for operators? The app's attractive and fantasy environment was expected to alleviate catastrophic thoughts. Participants were also encouraged to share any positive signs of reducing their frequency of thinking about their fears.

Effect 4: Preliminary Effects on Ailurophobia Treatment

The study used before and after assessments with the State-Trait Anxiety Inventory (STAI) and FCQ to measure the game's impact on phobia symptom changes. The STAI questionnaire comprises 40 questions, measuring state (S-scale) and trait (T-scale) anxiety using a 4-point Likert scale. Only the S-scale was used in this study. The evaluation of state anxiety can be used for any situation with a time interval determined by a researcher or a clinical specialist. Mahram developed the Persian version of the STAI, and its internal consistency was confirmed for the S- and T-scales (Cronbach α of .91 and .90, respectively) [40]. Another Iranian study also reported high reliability for the S- and T-scales with Cronbach α values of .93 and .90, respectively [41]. The FCQ questionnaire was derived from the Fear of Spiders Questionnaire (FSQ) to assess cat phobia, with all instances of the word "spider" replaced by "cat." Furthermore, the question format was adjusted to suit the assessment of the cat phobia. The FSQ is an 18-item tool scored on a 7-point Likert-type scale to measure the level of spider phobia, yielding a total score ranging from 18 to 126. The FSQ demonstrates excellent internal consistency with Cronbach α ranging from .88 to .97 [42,43] and good test-retest reliability [42]. The FSQ has been used in previous studies for various phobias such as cockroaches [7,44,45], rats [46], and snakes [47].

After the games, each group of participants was instructed to play 2 levels of the game as the opposite group did. They were then asked to answer the following questions: (1) Which experiences do you prefer? (2) Which experiences had more novelty and were more attractive to you? (3) Which experience was more effective for improving your problem?

Patients' Preferences About the Designed Treatment

To gather patients' opinions on the implemented treatment, an adapted preference questionnaire [48] was used. This 6-item questionnaire focused on patients' preferences regarding the types of cat models, their behaviors, sounds, and sizes. For example, the questions related to cat models are as follows: (1) If you could choose among the cat models, which one would you prefer? (2) Which cat model do you think would be more effective in helping you overcome your problem? (3) Which cat model do you find more logical for aiding in your progress? (4) Which of these cat models do you perceive as more aversive? (5) Which cat model would you recommend to a friend facing the same problem? (6) Are there any cat models missing in the game?

Heuristics Evaluations

The playability and usability aspects of the tool were examined through heuristic evaluations designed as semistructured interviews to optimize its performance and enhance usability. Participants completed a 5-Likert questionnaire covering user interfaces, VR experiences, BF, and game playability. Participants had the opportunity to provide additional comments. The evaluations incorporated 44 heuristics from studies [49-52]. We used Nielsen heuristics [49,50] to assess the interfaces, along with modified Nielsen principles for VR platforms [51]. In terms of game playability, a comprehensive evaluation was necessary to assess additional features, including gameplay, story, and mechanics, which went beyond simple interface usability evaluation [52]. We used the heuristic principles of playability introduced in [52], which carefully examine the various components of a game in terms of playability and enjoyment for the player, encompassing gameplay, mechanics, usability, and game story. In this study, we used the first 3 heuristics from this set.

Statistical Analysis

We evaluated the differences in subjective ratings of intrinsic motivation and levels of anxiety using ANOVA: 2-factor with replication of the total score, with session number as the time factor and task variant (the tool with and without BF) as the between-subjects factor. In addition, we used 1-way ANOVA with task variant as the between-subjects factor to investigate the effects of the tool on mitigating phobia symptoms. For analyzing the semistructured interviews, mean and SD scores were used.

Results

Participants

Of the 17 participants, 7 were excluded for (1) heart disease (n=1); (2) vision or balance problems (n=1; participants with VR-induced dizziness and severe nausea); (3) pregnancy (n=1); (4) personality disorders (n=1); and (5) fear of cats in specific situations (n=3; one was afraid of direct eye contact with cats, whereas 2 others were scared of black cats). Among the 10 included participants (Table 1), 1 individual had 2 other phobias: fear of public toilets (paruresis) and birds (ornithophobia), especially their beaks and legs. Another participant displayed general phobia of animals; even touching chicks elicited an electric shock response. In addition, the sight of cats, dogs (cynophobia) especially when they bark, and foxes caused annoyance and discomfort for her. Interestingly, she was more afraid of kittens than fully grown cats. Another patient had cynophobia and ailurophobia. Finally, 1 participant had a phobia of space and galaxies (to the extent of avoiding space-themed movies) as well as chicks phobia and ornithophobia stating, "I am even afraid of a bird in a cage that might come out and harm me." This participant also avoided going to the park because of the fear of the animals. Given the prevalence of individuals experiencing multiple phobias, particularly fears related to various animals (zoophobia), such as cats, spiders, snakes, and dogs, it is crucial to explore the possibility of modifying the game to effectively address multiple types of phobias. The park environment appears to be conducive to addressing various animal phobias and specific phobias such as paruresis. Accessing 10 participants was hindered by the temporary filtering of Instagram in our country. In addition, 2 individuals declined to participate, expressing shyness and concerns about others noticing their phobia. Our observations suggest that men with ailurophobia conceal their fear more frequently. Notably, ailurophobia predominantly affected women, as 90% (9/10) of our participants were women (Table 1). Ailurophobia began in 70% (7/10) of the participants during childhood and 30% (3/10) during adolescence. The minimum and maximum ages of onset of phobia in the samples were 5 and 18 years, respectively. Regrettably, animal phobias in our country, particularly cat phobia, have been largely overlooked, leading individuals to live for many years in a completely curable condition without seeking treatment. Innovative and early interventions, for example, our tool, could treat patients from childhood when anxiety starts and reduce the negative impact of untreated phobias. There is a pressing need for screening and diagnostic games as a primary step, followed by therapeutic games. The main cause of participants' phobia stemmed from an unexpected childhood encounter with a cat.

Table 1. Participants' characteristics.

Characteristics	Non-BF[a]	BF
Age (years), mean (SD)	24 (7.31)	33.5 (7.16)
Female, n (%)	5 (100)	4 (80)
Video game playing hours per week, mean (SD)	6 (8.52)	1.5 (3.08)
Median level of education	Diploma degree	Master's degree
Years living with cat phobia, mean (SD)	15.6 (5.68)	21.6 (12.01)
Married, n (%)	2 (40)	2 (40)
The onset of phobia, n (%)		
Childhood	4 (80)	3 (60)
Adolescence	1 (20)	2 (40)
Youth	0 (0)	0 (0)
Adulthood	0 (0)	0 (0)
Age of onset of phobia (years), range	9-13	5-18

[a]BF: biofeedback.

Possible Positive Effects of the Gamified VRET Augmented With BF

Effect 1: Intrinsic Motivation

The average intrinsic motivation of the groups indicated better results for the BF group across all 4 mandatory games with 49 scores (the sum of motivation scores for BF vs non-BF in the first to fourth sessions were: 182 vs 174, 178 vs 169, 191 vs 160, and 182 vs 181). However, the results (*P* value [groups]=.15>.05=∝ and $F_{1,3}$=2.165) indicate no statistically significant difference. The analysis used a 2-factor ANOVA with replication.

On the basis of the results (*P* value [sessions]=.91>.05=∝; $F_{3,3}$=0.171), we can conclude that there were no significant differences in the effectiveness of the groups across the different sessions.

There were no significant differences in the interaction between groups and sessions (*P* value [interactions]=.61>.05=∝ and $F_{3,3}$=0.609).

Of the 5 participants in the non-BF group, 4 played 2 levels using BF. Two of them chose each game version, whereas the other 2 preferred the BF version exclusively.

Overall, BF had a greater effect on motivating patients. With greater efforts to leverage its potential within the game, the positive impact on motivation can be substantially enhanced. Nevertheless, it is essential to note that the non-BF version fosters motivation by incorporating 2 vital motivational elements: gamification and VR.

As participants enter new and especially challenging stages, their internal motivation to play tends to decrease, whereas their anxiety increases. However, with repeated attempts at this stage, motivation gradually increased, and anxiety levels tended to decrease.

Effect 2: Simulating Fearful Situations

The non-BF group had, on average, 40 points higher anxiety scores across all 4 rounds of the forced games compared with the BF group (the sum of anxiety scores for BF vs non-BF in the first to fourth sessions were: 11 vs 33, 22 vs 34, 29 vs 34, and 27 vs 28). There was a statistically significant difference between the 2 groups (*P*=.009<.05=∝ and $F_{1,3}$=7.805). The total anxiety score for the non-BF group was 129, whereas that for the other group was 89, indicating the beneficial role of BF in anxiety control. This finding also suggests that using BF could potentially reduce the need for a therapist's presence. Caution is advised when interpreting these data, as it may be influenced by individuals with severe phobias. The crucial point is that both game variants can evoke anxiety, as they simulate fearful situations. During the games, 5 participants (4 without BF and 1 with BF) experienced extreme stress, necessitating temporary pauses to help them calm down. One participant even reported an increase in blinking frequency when feeling nervous while playing the game.

The *P* value (sessions)=.32>.05=∝ and $F_{3,3}$=1.204, indicating no significant differences in the effectiveness of the groups across different sessions. Many participants experienced anxiety even before the games began, which significantly impacted their anxiety levels during training (picture step). One participant even mistook pictures of cats in the training as real cats because of high tension. In addition, 6 participants (4 without BF and 2 with BF) responded to the cat pictures. On the basis of the data and participant feedback, the order of increasing anxiety levels followed the sequence of stages, starting from the trial game and progressing through the forced games in the following order: fantasy, low-poly, and high-poly cats. Similarly, the normalization of cats occurred in the following order: fantasy cats, pictures of cats, low-poly cats, and high-poly cats. For instance, anxiety levels increased as the number of cats increased. No significant differences in interaction between groups and sessions were observed (*P* value [interactions]=.20>.05=∝ and $F_{3,3}$=1.652).

Effect 3: Controlling Stressful Circumstances, Eliminating Therapists' Presence, and Mitigating Catastrophic Thoughts

Most participants about the positive signs of reducing their catastrophic thoughts expressed that encountering cats had started to feel somewhat normal. They noted that with continued play, their irrational fears could be replaced with more rational ones, and these positive changes could extend beyond the game to real-life environments. One participant shared, "Before playing the game, I couldn't even look at cats' stickers or images, and I used to throw my toy cat out of my room window into the street." Another participant expressed, "Encountering fantasy cats in small numbers has become normal for me, and I believe that over time, my fear of other types of cats will decrease." Follow-up data are required to verify the lasting impact of these positive changes.

A total of six noteworthy comments on the elimination of therapists using BF were suggested: (1) after a few sessions, the game can be played independently without therapists; (2) the treatment process can be shortened; (3) patients with milder phobias can benefit from playing without therapists. Otherwise, therapists' support is necessary during the initial sessions; (4) the game is more beneficial for therapists, offering a controlled environment free of danger; (5) combining virtual and face-to-face treatments is recommended, starting with the game to prepare patients for real-life cat encounters; and (6) BF cannot provide the psychological support therapists offer. One participant, Fatemeh, repeatedly reassured herself during gameplay, saying, "Fatemeh, it's just a cat, it's nothing, keep calm." The necessity of a virtual therapist to provide reassurance and guidance during moments of severe anxiety was evident. Participants either managed to calm themselves or received assistance from us. At times, we had to explain the unlocked stage scenarios to convince the participants to proceed with the remaining games.

To enhance the effectiveness, some participants suggested that the game should display their effort by showing the minimum and maximum HR and the time taken to complete a level. In addition, 2 positive comments regarding HR were as follows: "I noticed that my fear is higher before encountering cats, but my heart rate decreases when I face them" and "Before playing, I believed my fear of cats was overwhelming, but the game helped me realize it wasn't as intense as I thought."

Effect 4: Preliminary Effects on Ailurophobia Treatment

Using ANOVA single factor, we could not detect a difference between the groups ($F_{1,8}$=0.073, and *P* value=.79>.05=∝). The S-scale scores worsened by 50 and 33 points in the non-BF and BF groups, respectively (Table 2). Both variants induced anxiety, but the BF group showed lower anxiety levels, suggesting that BF was more effective in reducing stress.

No significant difference between the groups was detected (*P* value=.63>.05=∝, and $F_{1,8}$=0.256). The non-BF group improved by 67 points in the FCQ scores, whereas the BF group worsened by 42 points (Table 3).

The significant difference in scores can be attributed to one participant in the non-BF group who initially experienced high anxiety before and during the game. However, as she played more games, her scores on the S-scale (64-28) and FCQ (119-13) decreased dramatically. She mentioned that she used to be greatly bothered by cats being near her or hearing their voices, but after playing the game, she felt less anxious. The constant presence of cats in the game and being able to hear their voices helped her overcome her fears. It is noteworthy that this participant played the game more than all other players, completing 10 levels, including the training stage. In the last 3 stages, the participants specified an anxiety level of 1 out of 10. Initially, we considered this participant's data as an outlier, but because of the high number of games played, we retained her data. This observation clearly indicates that playing the game more frequently helps to normalize interactions with cats. Her anxiety scores (of 10) for playing 9 levels of the game were (the data related to training was excluded for all participants): 10, 10, 8, 3, 3, 2, 1, 1, 1. By replacing her score with a typical number, we obtained more reasonable scores. The non-BF and BF scores worsened by 9 and 42, respectively. Both game versions induced similar anxiety levels in participants. Some of the participants experienced symptom improvement. To assess the initial positive signs of phobia treatment using the FCQ, we should wait until the completion of 10 game stages on average. All participants completed this questionnaire shortly after the games (within a maximum of 10 minutes), and the effects of anxiety caused by fear were still evident. We had to reassure them that the game was not very scary and that the unpredictable event they feared would not happen in the next level, as 4 participants experienced extreme anxiety. These participants took longer breaks between the phases or temporarily stopped playing the game. This anxiety could adversely affect their grades. In addition, approximately 80% (8/10) of the participants mentioned that playing the game more often helped them become accustomed to seeing cats.

All participants expressed a preference for the gamified VRET with BF, stating that the experience was more novel and perceived as more effective in reducing fear.

Table 2. Pretest and posttest scores of S-scales.

Groups	Pretest	Posttest	Difference
Non-BF[a] (control)	39	61	−22
Non-BF	64	28	36
Non-BF	38	67	−29
Non-BF	46	67	−21
Non-BF	33	47	−14
BF (experimental)	39	45	−6
BF	29	28	1
BF	42	65	−23
BF	39	37	2
BF	34	41	−7

[a]BF: biofeedback.

Table 3. Pretest and posttest scores of Fear of Cats Questionnaire.

Groups	Pretest	Posttest	Difference
Non-BF[a] (control)	71	100	−29
Non-BF	119	13	106
Non-BF	82	94	−12
Non-BF	112	92	20
Non-BF	101	119	−18
BF (experimental)	84	97	−13
BF	70	82	−12
BF	76	104	−28
BF	89	85	4
BF	89	82	7

[a]BF: biofeedback.

Patients' Preferences About the Designed Treatment

Most participants expressed that the game had a positive therapeutic impact and was capable of normalizing their interactions with cats. In total, 2 participants played the game 7 and 10 times and reported significant changes in their perception of cat-related fears. They shared that their perceptions of cat fear transformed, and encountering cats felt normal. Moreover, they believed that this effect could be extended to real-life situations. One participant shared, "I used to feel uneasy when cats were nearby, and the sound of cats was distressing for me. But now, as cats are consistently present in the game, and the sound of cats is played during the gameplay, being close to cats and hearing their sounds has become completely normal for me." In the last 2 stages, their anxiety levels were reduced to a rating of 1 of 10. Before playing, most participants anticipated that cats would appear in the forest and perch on the tree branches. They expected the paths where cats were located to have denser and more crowded areas, featuring an abundance of trees, wooden huts, and gazebos with cats nearby. One commented, "The space provided is too vast, and it could be made more intense to induce fear. It would be beneficial to create some narrower paths leading to a door where cats are positioned. This could instill more fear. Generally, the game's paths are not challenging situations, and a darker environment could make the cat's eyes more prominent." These comments contradict most participants, who appreciated the game's positive aspect of indirectly implementing treatment and displaying everyday interactions people have with cats. Incorporating various environments and cat behaviors could further normalize the interaction with cats from all angles. However, these changes must be introduced with caution to avoid reinforcing the perception that cats are scary. In addition, the suggested locations to be included in the game range from the park environment to urban settings, such as apartments, streets, alleys, markets, cafés, dark scenes, kitchens, and garbage cans.

Some participants preferred the fantasy cats, believing that they alone have the ability to normalize interactions with cats because they highlight the positive aspects of cats such as their beautiful eyes and portray them as attractive, safe, and less harmful. Designing different fantasy cats appears to be a reasonable way to encourage individuals. One participant said, "It bothers me that the cats' heads are small and their tails are long. In contrast, fantasy models had big heads and short tails. In different game levels, placing fantasy cats next to other cats conveys the feeling that all cats are harmless. Starting with images of rough and fat cats and gradually increasing the number of cats, and transitioning them to real models, helped me realize that the initial stage's image was merely in my mind and unreal. As the cats' numbers increased, I discovered that they did not pose any harm." These eye-opening opinions shed light on an overlooked aspect—the psychological impact of the game's difficulty levels and cat types.

Preferring fantasy cats indirectly revealed that low- and high-poly cats mostly evoke fear. Most participants found these cats to be more rational. Increased aversion and avoidance were observed in places with more cat voices and presence. Longer sounds also intensified fear.

On the basis of these findings, it is suggested to gradually introduce sounds, starting from cats with no sounds to short and pleasant sounds and then to real single and multiple sounds. The maximum fear was near the boxes where the number and noise of the cats were higher. Although this arrangement was found to be effective and logical in normalizing interactions with cats, high fear levels may have led some participants to avoid playing altogether. One participant preferred orderly and grouped cats for a calmer experience, whereas disorderly placement near the box increased fear. These reasons highlight the significance of using fantasy cats. Most participants found the size of the cats were found to be suitable. However, larger cat sizes, such as pictures of striped cats and low-poly cats, increased anxiety. The picture level, considered the easiest, induced anxiety and fear in most participants (6 of 10). Concerning cats' behavior, most preferred nonreactive cats, such as fantasy cats that simply look at the sky in a cartoony manner; cats sitting and grooming themselves; or cats moving along the path without any reaction. Most participants disliked black cats waving their hands or white cream cats turning and staring at the player.

Most participants expressed the need for the game's cat designs to closely resemble real-world cats. The following cats were not used based on their comments:

- Spotted (mainly black and white) and gray-striped cats, which are abundant in Iran.
- Kittens: Participants made three points: (1) kittens may not have a significant therapeutic effect, but they enhance the game's appeal and create a more lifelike environment; (2) the treasure finder can be replaced with a fantasy kitten, allowing for a more captivating display of less favored features of cats, such as their nails, tail, and head. Moreover, their beautiful eyes can be showcased as larger; and (3) the option of raising a kitten in the game.
- Fierce-looking cats with grabbing capabilities: adding them requires expert opinions. Although statistics on cat grabbing are limited, the actual occurrence is likely to be minimal. People's intense fears may exaggerate this concern.
- A mother cat breastfeeding her babies for a heartwarming and motherly touch.
- Sphynx cats: despite being rare in Iran, could enhance realism and normalize fear of diverse cat breeds.
- Fat or fluffy cats resembling a doll-like appearance.
- Placing cats amidst the greens and bushes along the paths.
- Injured (eg, cats with one eye or leg) or lifeless cats.
- Sudden movements of cats (eg, cats leaping out of trash cans): mentioned by most participants.
- Feeding cats: some participants did not agree with implementing this feature.

In conclusion, the game layout and models were considered logical by most of the participants. They stressed that fighting with cats in the game could worsen their fear, making a clear distinction between a therapeutic game and one designed solely for entertainment purposes. This opinion is in agreement with the clinical expert (the fourth author) who emphasized that the games for treating animal phobias should avoid action and fighting scenarios. For example, reducing the fear of cockroaches using scenarios where they stomp on or kill them may adversely affect.

Heuristics Evaluations

As presented in Tables 4-7, of the 44 heuristics adapted from the Nielsen user interface, VR, and playability, an impressive 41 principles obtained scores of 62% or higher, underscoring the tool's potential as a therapeutic product. Moreover, it enhances patient adherence to the treatment process.

Overall, 90% (9/10) of the participants found learning to play the game remarkably easy, particularly with the convenience of using just one button under VR glasses, which proved beneficial for those with mobility disabilities. Two suggestions emerged concerning in-game movement: (1) incorporating a back button and (2) movement through walking, potentially achieved with motion-sensing devices. However, careful consideration is necessary to ensure that they positively impact the player experience. Some individuals may prefer a less cumbersome setup. To enhance experience, it is crucial to incorporate a tutorial in a video or audio format for first-time users by introducing relaxation techniques to manage panic situations. Many participants required clarification that frightful situations would not occur at the subsequent levels. Providing detailed descriptions of new levels, including information about cats' types and behaviors, prevents players from creating self-made stories about cat attacks. Moreover, to improve clarity, players needed clearer instructions after opening each box, signaling that they should open 4 boxes per session. Although a ribbon in the corner of the screen displays the number of opened boxes, it does not adequately alert the players to this requirement. Among the VR principles, the navigation and orientation support principle excelled at 82%, with patients being well-informed about their in-game position. Notably, approximately 80% (8/10) of patients experienced no dizziness during extended gameplay. To increase the level of engagement and therapeutic impact, introducing a punishment mechanism, such as reducing players' points, could be beneficial. It might be worth reconsidering the features of allowing players to win the game without encountering cats. Game sounds and music received a relatively low score (51%), causing tension and unease, instead of promoting peace and happiness. The addition of soothing natural sounds was also suggested. In addition, consider a sound to indicate proximity to the box, reducing the need to check the bar constantly and improving the focus on gameplay. The game could benefit from a save and resume feature, especially during panic situations, allowing patients to take a moment to calm down. Some also raised concerns about the suitability of graphics for older adult audiences.

The principle of variety in the game's paths and challenges stands out as one of the main gameplay principles. Although it obtained a relatively good score (68%), most participants said that after a few stages, the game became monotonous. Players quickly realized that cats only appear in certain sections of the roads and near treasure boxes. Certain adjustments were recommended to enhance the game's appeal. Increasing the spacing between trees and raising their height can create a more immersive environment. Adding colorful elements such as flowers, toys, water views, and a gazebo in the park will infuse vibrancy into the game. In general, elevating the game's attractiveness can be achieved by introducing a greater sense of adventure without relying on unrealistic fears. One participant suggested that instead of having the treasure box as the game's goal, it could be placed in various locations within the forest, each rewarding the player with different prizes, such as food. Another suggestion was to replace the guide bar, which received positive feedback from the participants, with a map that indicated the approximate distance to the target. In addition, the introduction of a captivating and fantastical cat character instead of the current bar was recommended. In total, 2 participants pointed out that displaying HR in the corner might be somewhat distracting. It was suggested to show HR only when it was high or to remind players to reduce stress using a heartbeat's sound.

Table 4. Results of the questionnaire designed based on [49-51] for evaluating user interfaces and virtual reality apps, respectively (Tables S1-S4 of Multimedia Appendix 1 provides the noncompressed version of Tables 4-7 containing the list of questions).

Usability heuristic and question	Question, mean (SD)	Heuristic, mean (SD)	Heuristic overall percent, %
1. Natural engagement			
Q1	3.5 (1.08)	3.7 (0.28)	74
Q2	3.9 (0.88)	N/A[a]	N/A
2. Compatibility with the user's task and domain			
Q3	3.7 (0.95)	3.7 (0.3)	74
Q4	4 (1.25)	N/A	N/A
Q5	3.4 (1.17)	N/A	N/A
3. Natural expression of action			
Q6	3.5 (1.27)	3.2 (0.42)	64
Q7	2.9 (1.20)	N/A	N/A
4. Close coordination of action and representation			
Q8	3.5 (1.27)	3.6 (0.1)	72
Q9	3.6 (1.17)	N/A	N/A
Q10	3.7 (1.25)	N/A	N/A
5. Realistic feedback			
Q11	3.7 (1.4)	3.7 (1.4)	74
6. Faithful viewpoints			
Q12	3.6 (1.2)	3.6 (1.2)	72
7. Navigation and orientation support			
Q13	4.1 (1.2)	4.1 (1.2)	82
8. Visibility of system status			
Q14	3.7 (1.25)	3.63 (0.75)	72.5
Q15	3.4 (1.17)	N/A	N/A
Q16	3.7 (1.49)	N/A	N/A
Q17	3.7 (1.16)	N/A	N/A
9. Consistency and standards			
Q18	3.7 (1.3)	3.7 (1.3)	74
10. Error prevention			
Q19	3.7 (1.34)	3.5 (0.28)	70
Q20	3.3 (2)	N/A	N/A
11. Recognition rather than recall			
Q21	3.6 (1.2)	3.6 (1.2)	72
12. Flexibility and efficiency of use			
Q22	3.8 (1.0)	3.8 (1.0)	76
14. Help and documentation			
Q23	2.9 (1.4)	2.9 (1.4)	58

[a]N/A: not applicable.

Table 5. Results of the questionnaire designed based on gameplay part of the playability heuristics [52].

Question	Question, mean (SD)	Heuristic, mean (SD)	Heuristic overall percent
1. Player's fatigue is minimized by varying activities and pacing during game play.			
Q1	3.4 (1.6)	3.4 (1.6)	68
2. Provide consistency between the game elements and the overarching setting and story to suspend disbelief.			
Q2	3.4 (1.3)	3.4 (1.3)	68
3. Provide clear goals, present overriding goal early as well as short-term goals throughout play.			
Q3	4.3 (1.1)	4.3 (1.1)	86
4. There is an interesting and absorbing tutorial that mimics game play.			
Q4	4.5 (1.0)	4.1 (0.57)	90
Q5	3.7 (0.9)	N/A[a]	N/A
5. The game is enjoyable to replay.			
Q6	3.5 (0.7)	3.5 (0.7)	70
6. Game play should be balanced with multiple ways to win.			
Q7	3.8 (1.0)	3.8 (1.0)	76
7. Player is taught skills early that you expect the players to use later, or right before the new skill is needed.			
Q8	3.4 (1.5)	3.4 (1.5)	68
8. Players discover the story as part of game play.			
Q9	4 (0.8)	4 (0.8)	80
9. The game is fun for the Player first, the designer second and the computer third. That is, if the nonexpert player's experience is not put first, excellent game mechanics and graphics programming triumphs are meaningless.			
Q10	3.9 (1.2)	3.9 (1.2)	78
10. Player should not experience being penalized repetitively for the same failure.			
Q11	4.3 (0.7)	4.3 (0.7)	86
11. Player's should perceive a sense of control and impact onto the game world. The game world reacts to the player and remembers their passage through it. Changes the player makes in the game world are persistent and noticeable if they back-track to where they have been before.			
Q12	4.1 (1.0)	3.9 (0.28)	82
Q13	3.7 (1.3)	N/A	N/A
12. The game should give rewards that immerse the player more deeply in the game by increasing their capabilities (power-up), and expanding their ability to customize.			
Q14	3.5 (1.3)	3.7 (0.28)	70
Q15	3.9 (1.0)	N/A	N/A
13. Pace the game to apply pressure but not frustrate the player. Vary the difficulty level so that the player has greater challenge as they develop mastery. Easy to learn, hard to master.			
Q16	3.9 (1.2)	3.75 (0.21)	78
Q17	3.6 (1.2)	N/A	N/A
14. Challenges are positive game experiences, rather than a negative experience (results in their wanting to play more, rather than quitting).			
Q18	3.8 (1.1)	3.8 (1.1)	76

[a]N/A: not applicable.

Table 6. Results of the questionnaire designed based on mechanic part of the playability heuristics [52].

Question	Question, mean (SD)	Heuristic, mean (SD)	Heuristic overall percent
1. Game should react in a consistent, challenging, and exciting way to the player's actions (eg, appropriate music with the action).			
Q1	2.8 (1.6)	2.8 (1.6)	51
2. Make effects of the AI[a] clearly visible to the player by ensuring they are consistent with the player's reasonable expectations of the AI actor.			
Q2	3.1 (0.9)	3.1 (0.9)	62.3
3. A player should always be able to identify their score/status and goal in the game.			
Q3	4.4 (0.5)	4.3 (0.14)	86
Q4	4.2 (0.9)	N/A[b]	N/A
4. Mechanics/controller actions have consistently mapped and learnable responses.			
Q5	4.4 (1.1)	4.15 (0.35)	83
Q6	3.9 (1.6)	N/A	N/A
5. Shorten the learning curve by following the trends set by the gaming industry to meet user's expectations.			
Q7	4.3 (1.3)	3.95 (0.49)	79
Q8	3.6 (1.6)	N/A	N/A
6. Controls should be intuitive, and mapped in a natural way; they should be customizable and default to industry standard settings.			
Q9	4.5 (1.0)	4.35 (0.21)	87
Q10	4.2 (0.9)	N/A	N/A
7. Player should be given controls that are basic enough to learn quickly yet expandable for advanced options.			
Q11	3.3 (1.3)	3.53 (0.32)	70.67
Q12	3.4 (1.7)	N/A	N/A
Q13	3.9 (1.4)	N/A	N/A

[a]AI: artificial intelligence.

[b]N/A: not applicable.

Table 7. Results of the questionnaire designed based on usability part of the playability heuristics [52].

Question	Question, mean (SD)	Heuristic, mean (SD)	Heuristic overall percent
1. Provide immediate feedback for user actions.			
Q1	4.1 (1.5)	4.1 (1.5)	82
2. The player can easily turn the game off and on, and be able to save games in different states.			
Q2	2.3 (1.3)	2.3 (1.3)	46
3. The player experiences the user interface as consistent (in control, color, typography, and dialog design) but the gameplay is varied.			
Q3	3.7 (1.3)	3.35 (0.49)	67
Q4	3 (1.2)	N/A[a]	N/A
4. The player should experience the menu as a part of the game.			
Q5	3.4 (1.0)	3.65(0.35)	68
Q6	3.9 (1.0)	N/A	N/A
5. Sounds from the game provide meaningful feedback or stir a particular emotion.			
Q7	3.5 (1.2)	3.35 (0.35)	67
Q8	3.2 (1.1)	N/A	N/A
Q9	2.8 (1.6)	N/A	N/A
6. Players do not need to use a manual to play the game.			
Q10	4 (0.9)	4 (0.9)	80
7. Make the menu layers well organized and minimalist to the extent the menu options are intuitive.			
Q11	3.9 (1.6)	3.9 (1.6)	78
8. Get the player involved quickly and easily with tutorials and/or progressive or adjustable difficulty levels.			
Q12	4.1 (1.3)	3.95 (0.21)	79
Q13	3.8 (1.4)	N/A	N/A
9. Art should be recognizable to the player, and speak to its function.			
Q14	3.7 (1.2)	3.65 (0.07)	73
Q15	3.6 (0.7)	N/A	N/A

[a]N/A: not applicable.

Discussion

Principal Findings

We developed a gamified VRET augmented with BF to address ailurophobia. To our knowledge, no specialized research on ailurophobia treatment exists, either in Iran or internationally. Motivated by the high prevalence of ailurophobia and the lack of accessible gamified VR environments with BF, our main goal was to create and assess a smartphone-based VRET augmented with BF for animal phobia (cat phobia). We hypothesized that this tool would better motivate patients, manage stress, simulate fearful situations, treat phobia, and reduce therapists' involvement compared with a gamified VRET. The tool was designed based on expert sessions in video games, gamification, cognitive, and psychology. The results indicate its positive impact on specified features. Of the 44 heuristics, 41 scored above 62%, showing the potential for phobia interventions and motivating patients for treatment. Although tested on only 10 participants for a short duration (up to 3 hours without follow-up sessions), the results were reliable. Extensive data and feedback collection have been used to evaluate various aspects of the tool. On average, after 10 sessions, initial signs of improvement were observed, with slight variations depending on individuals' phobia levels. One intriguing finding was that most participants were content with the game's indirect approach to normalize interaction with cats and its nonviolent nature. They emphasized that action or combat scenarios would reinforce unrealistic fears and validate their phobia. Another significant finding was the progression of normalization in dealing with cats, tolerating their behavior, and hearing their voices, which gradually became more challenging. Although the current game normalizes communication with cats and holds good appeal, most participants suggested improvements, such as adding a variety of cats that closely resemble real-world characteristics, including voices and behaviors, to further enhance the normalization process. In addition, most participants expressed satisfaction with the game's easy movements and minimal learning curve. To enhance the experience, adding diversity and adventure while minimizing unrealistic violence was recommended. Moreover, during the evaluations, the participants strongly felt the need for a virtual therapist to provide calming guidance and support during moments of severe anxiety.

Comparison With Prior Work

To our knowledge, no study has simultaneously used BF, VR, and gamification for the treatment of animal phobia. However, various studies have used VR and game concepts to address specific animal phobias, for example, spider phobia [12,32] and snake phobia [47]. Similar to these studies, our tool successfully induced anxiety and led to a reduction in fear levels, avoidance behaviors, and catastrophic thoughts related to phobias. In addition, it positively boosted their motivation for treatment adherence. Unlike previous studies, our unique feature was the initial evaluation, showing that participants preferred a gamified VRET with BF. It has proven to be more effective in reducing symptoms and increasing internal motivation. These findings align with recent reviews highlighting the significant anxiety-reducing benefits of combining VR and BF, along with the advantages in motivation, user experience, involvement, and attentional focus [53,54]. In contrast to our study, where more participants preferred interacting with safe stimuli such as fantasy cats, studies such as those by Dibbets and Schruers [55] and Pittig et al [56] reported that selecting riskier options led to a stronger decrease in self-reported spider fear and disgust, whereas safe choices increased these emotions. The differing outcomes could be attributed to the use of VR and 3D images. VRETs are widely recognized as an appealing treatment modality because of their perceived naturalness in the automated format. However, Albakri et al [57] suggested that augmented reality exposure therapies offer a better experience and increased realism by seamlessly integrating digital information into the real world rather than creating a completely new virtual environment. We plan to explore the implementation of our designed tool with augmented reality and compare the outcomes in future studies. Dibbets and Schruers [55] found that the number of spiders encountered did not correlate with declines in aversive feelings and avoidance behaviors. However, our study concluded that a higher number of stimuli were more effective in normalizing interactions with cats. In addition, we observed that the action and combat scenarios were not beneficial for individuals with phobias. Interestingly, snake phobia treatment in a nearly action genre format [47] lacks a rationale for its selection. Further research is required to determine and devise appropriate scenarios for individuals with phobias. Throughout this study, the need to conduct similar research was highlighted. It was not feasible to make precise comparisons with prior studies in every detail.

Limitations

The initial study on treating ailurophobia using VRETs with gamification and BF had limitations, primarily a small number of respondents. A total of 10 potential participants were inaccessible after Instagram's temporary filtering in our country. The sample was skewed toward educated participants in their twenties and thirties, indicating the need to include diverse educational backgrounds, children, adolescents, and older adults. Owing to time constraints, we did not use any statistical method to calculate the required sample size. The study by Mor et al [48] recommended a minimum of 20 participants in each study arm for feasibility pilot studies on treating flying phobia using 360° images. Certainly, a larger number of patients is needed in each arm for the primary assessments. One future work is to replicate the study quantitatively and more rigorously while also introducing another arm that uses standard and clinical exposure therapies, enabling us to evaluate the tool and showcase more applications. In addition, the small sample sizes prohibited us from examining dropout rates. The results are exploratory, and long-term effects remain unknown due to the lack of follow-up data. Only one self-rating scale, the FCQ, has been used to diagnose individuals with ailurophobia. However, it is advisable to supplement such questionnaires with a telephone or face-to-face diagnostic interview conducted by an expert clinician, typically lasting approximately 30 minutes [2,12,32]. These interviews not only boost diagnostic reliability but also enable descriptive analysis [2]. It is worth mentioning that the patients were initially asked to explain the origin and signs of their ailurophobia. Participants were randomly divided into groups; however, the equality of their stress levels was not considered. It appears that by preserving randomness, the stress levels of individuals in the study groups should be nearly equal. For example, if one group has 2 extreme cases, the other groups should also have 2 similar cases to ensure transparency and enhance the reliability of the results. Creating a real-world game proved challenging owing to the limitations of the smartphone platform. Although playability and system usability questionnaires were not rigorously assessed, they were designed based on popular usability scales, including Nielsen [49,50], VR [51], and playability [52] heuristic evaluations. Changes in the individual's physiological status, particularly HR, influence their experience. Unfortunately, this feature could not be assessed in the BF arm owing to the small sample size. Understanding its effectiveness in high-tension situations and its role in reducing anxiety remains a top priority.

Conclusions and Future Work

The gamified VRET incorporating BF for treating cat phobia could be effective and has the potential to evolve into a comprehensive tool. One way to enhance its utility is by expanding the variety of cat types and behaviors, simulating different environments where cats are commonly found, and boosting its appeal through increased adventure while avoiding the use of unrealistic fears. After modifying the tool and using more robust study designs with ample sample sizes, further investigation can explore how this tool can be used in treatments without the presence of a therapist or combined (virtual and real simulation of fear), both in clinics and remotely. The park environment has the potential to effectively treat various animal phobias and other specific phobias. Implementing a gradual progression of sound stimuli could improve the therapeutic process. Starting with serene and pleasant sounds and gradually advancing to more challenging and potentially distressing voices, like cats squealing (inspired by a participant's recollection of hearing a cat giving birth) or their aggressive vocalizations during fights. The final suggestion is to add the possibility of interacting with cats during more challenging stages, thereby bridging the game environment with the real world.

Acknowledgments

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The data presented in this study are available from the corresponding author upon request.

Authors' Contributions

AK contributed to the conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing the original draft, and reviewing and editing. AN contributed to conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing the original draft, and reviewing and editing. ZA contributed to the conceptualization, data curation, formal analysis, methodology, project administration, resources, validation, visualization, writing the original draft, and reviewing and editing. AKB and PHA contributed to conceptualization, formal analysis, funding acquisition, investigation, resources, software, validation, and reviewing and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaires were designed based on user interfaces, virtual reality, and playability heuristics, along with their results.

[DOC File , 136 KB - games_v12i1e34535_app1.doc]

References

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun 01;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
2. Lindner P, Miloff A, Bergman C, Andersson G, Hamilton W, Carlbring P. Gamified, automated virtual reality exposure therapy for fear of spiders: a single-subject trial under simulated real-world conditions. Front Psychiatry 2020 Mar 3;11:116 [FREE Full text] [doi: 10.3389/fpsyt.2020.00116] [Medline: 32210850]
3. Wolitzky-Taylor KB, Horowitz JD, Powers MB, Telch MJ. Psychological approaches in the treatment of specific phobias: a meta-analysis. Clin Psychol Rev 2008 Jul;28(6):1021-1037. [doi: 10.1016/j.cpr.2008.02.007] [Medline: 18410984]
4. Garcia R. Neurobiology of fear and specific phobias. Learn Mem 2017 Sep 16;24(9):462-471 [FREE Full text] [doi: 10.1101/lm.044115.116] [Medline: 28814472]
5. Öst LG, Ollendick TH. Brief, intensive and concentrated cognitive behavioral treatments for anxiety disorders in children: a systematic review and meta-analysis. Behav Res Ther 2017 Oct;97:134-145. [doi: 10.1016/j.brat.2017.07.008] [Medline: 28772195]
6. Eaton WW, Bienvenu OJ, Miloyan B. Specific phobias. Lancet Psychiatry 2018 Aug;5(8):678-686. [doi: 10.1016/s2215-0366(18)30169-x]
7. Botella C, Breton-López J, Quero S, Baños RM, García-Palacios A, Zaragoza I, et al. Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: a single case study. Comput Hum Behav 2011 Jan;27(1):217-227. [doi: 10.1016/j.chb.2010.07.043]
8. Zhang T, Booth R, Jean-Louis R, Chan R, Yeung A, Gratzer D, et al. A primer on usability assessment approaches for health-related applications of virtual reality. JMIR Serious Games 2020 Oct 28;8(4):e18153 [FREE Full text] [doi: 10.2196/18153] [Medline: 33112243]
9. Garrett B, Taverner T, Gromala D, Tao G, Cordingley E, Sun C. Virtual reality clinical research: promises and challenges. JMIR Serious Games 2018 Oct 17;6(4):e10839 [FREE Full text] [doi: 10.2196/10839] [Medline: 30333096]
10. Park MJ, Kim DJ, Lee U, Na EJ, Jeon HJ. A literature overview of virtual reality (VR) in treatment of psychiatric disorders: recent advances and limitations. Front Psychiatry 2019 Jul 19;10:505 [FREE Full text] [doi: 10.3389/fpsyt.2019.00505] [Medline: 31379623]
11. Strickland D, Hodges L, North M, Weghorst S. Overcoming phobias by virtual exposure. Commun ACM 1997 Aug;40(8):34-39. [doi: 10.1145/257874.257881]
12. Lindner P, Rozental A, Jurell A, Reuterskiöld L, Andersson G, Hamilton W, et al. Experiences of gamified and automated virtual reality exposure therapy for spider phobia: qualitative study. JMIR Serious Games 2020 Apr 29;8(2):e17807 [FREE Full text] [doi: 10.2196/17807] [Medline: 32347803]

13. Khaliq I, Fowles J, Moore C. Theorizing gamified virtual reality approach to overcome fear of height. In: Proceedings of the Smart Objects and Technologies for Social Good. 2018 Presented at: GOODTECHS 2017; November 29-30, 2017; Pisa, Italy. [doi: 10.1007/978-3-319-76111-4_18]
14. Kumar JM, Herger M. Gamification at Work: Designing Engaging Business Software. Aarhus, Denmark: Interaction Design Foundation; 2013.
15. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev 2019 Aug 15;32(1):77-112. [doi: 10.1007/s10648-019-09498-w]
16. Werbach K, Hunter D. For the Win: How Game Thinking Can Revolutionize Your Business. Philadelphia, PA: Wharton Digital Press; 2012.
17. Khaleghi A, Aghaei Z, Mahdavi MA. A gamification framework for cognitive assessment and cognitive training: qualitative study. JMIR Serious Games 2021 May 18;9(2):e21900 [FREE Full text] [doi: 10.2196/21900] [Medline: 33819164]
18. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. Front Psychiatry 2016;7:215 [FREE Full text] [doi: 10.3389/fpsyt.2016.00215] [Medline: 28119636]
19. Schroeder PA, Lohmann J, Ninaus M. Preserved inhibitory control deficits of overweight participants in a gamified stop-signal task: experimental study of validity. JMIR Serious Games 2021 Mar 12;9(1):e25063 [FREE Full text] [doi: 10.2196/25063] [Medline: 33709936]
20. Khaleghi A, Aghaei Z, Behnamghader M. Developing two game-based interventions for dyslexia therapeutic interventions using gamification and serious games approaches entertainment computing journal. Entertain Comput 2022 May;42:100482. [doi: 10.1016/j.entcom.2022.100482]
21. Behnamghader M, Khaleghi A, Izadpanah P, Rahmani F. Using gamification based on mobile platform in therapeutic interventions for children with dyslexia. In: Proceedings of the 13th International Conference on Interactive Mobile Communication Technologies and Learning. 2021 Presented at: 13th International Conference on Interactive Mobile Communication Technologies and Learning; October 31-November 01, 2019; Thessaloniki, Greece. [doi: 10.1007/978-3-030-49932-7_76]
22. Khaleghi A, Aghaei Z, Hosseinnia F. Toward using effective elements in adults' amblyopia treatment in a virtual reality-based gamified binocular application. Entertain Comput 2022 Aug;43:100504. [doi: 10.1016/j.entcom.2022.100504]
23. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DM. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. JMIR Serious Games 2020 Aug 10;8(3):e18644 [FREE Full text] [doi: 10.2196/18644] [Medline: 32773374]
24. Sailer M, Hense JU, Mayr SK, Mandl H. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput Hum Behav 2017 Apr;69:371-380. [doi: 10.1016/j.chb.2016.12.033]
25. Landers RN, Tondello GF, Kappen DL, Collmus AB, Mekler ED, Nacke LE. Defining gameful experience as a psychological state caused by gameplay: replacing the term 'gamefulness' with three distinct constructs. Int J Hum Comput Stud 2019 Jul;127:81-94. [doi: 10.1016/j.ijhcs.2018.08.003]
26. Flowers BA, Tatoian R, Witkowski A, Katenka N, Herve JY. Using multi-sensor voting for resilience in VR biofeedback games. In: Proceedings of the IEEE Games, Entertainment, Media Conference (GEM). 2018 Presented at: IEEE Games, Entertainment, Media Conference (GEM); August 15-17, 2018; Galway, Ireland. [doi: 10.1109/gem.2018.8516522]
27. Dzedzickis A, Kaklauskas A, Bucinskas V. Human emotion recognition: review of sensors and methods. Sensors (Basel) 2020 Jan 21;20(3):592 [FREE Full text] [doi: 10.3390/s20030592] [Medline: 31973140]
28. Giggins OM, Persson UM, Caulfield B. Biofeedback in rehabilitation. J Neuroeng Rehabil 2013 Jun 18;10:60 [FREE Full text] [doi: 10.1186/1743-0003-10-60] [Medline: 23777436]
29. Sonne T, Jensen MM. ChillFish: a respiration game for children with ADHD. In: Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction. 2016 Presented at: TEI '16; February 14-17, 2016; Eindhoven, Netherlands. [doi: 10.1145/2839462.2839480]
30. Michałowski JM, Droździel D, Matuszewski J, Koziejowski W, Jednoróg K, Marchewka A. The set of fear inducing pictures (SFIP): development and validation in fearful and nonfearful individuals. Behav Res Methods 2017 Aug;49(4):1407-1419 [FREE Full text] [doi: 10.3758/s13428-016-0797-y] [Medline: 27613018]
31. Johnstone KA, Page AC. Attention to phobic stimuli during exposure: the effect of distraction on anxiety reduction, self-efficacy and perceived control. Behav Res Ther 2004 Mar;42(3):249-275. [doi: 10.1016/S0005-7967(03)00137-2] [Medline: 14975769]
32. Miloff A, Lindner P, Hamilton W, Reuterskiöld L, Andersson G, Carlbring P. Single-session gamified virtual reality exposure therapy for spider phobia vs. traditional exposure therapy: study protocol for a randomized controlled non-inferiority trial. Trials 2016 Feb 02;17:60 [FREE Full text] [doi: 10.1186/s13063-016-1171-1] [Medline: 26833396]
33. Yu B, Funk M, Hu J, Wang Q, Feijs L. Biofeedback for everyday stress management: a systematic review. Front ICT 2018 Sep 7;5. [doi: 10.3389/fict.2018.00023]
34. Chittaro L, Sioni R. Affective computing vs. affective placebo: study of a biofeedback-controlled game for relaxation training. Int J Hum Comput Stud 2014 Aug;72(8-9):663-673. [doi: 10.1016/j.ijhcs.2014.01.007]

35. Houzangbe S, Christmann O, Gorisse G, Richir S. Integrability and reliability of smart wearables in virtual reality experiences: a subjective review. In: Proceedings of the Virtual Reality International Conference - Laval Virtual. 2018 Presented at: VRIC '18; April 4-6, 2018; Laval, France. [doi: 10.1145/3234253.3234305]
36. Lumsden J, Skinner A, Coyle D, Lawrence N, Munafo M. Attrition from web-based cognitive testing: a repeated measures comparison of gamification techniques. J Med Internet Res 2017 Nov 22;19(11):e395 [FREE Full text] [doi: 10.2196/jmir.8473] [Medline: 29167090]
37. Lumsden J, Skinner A, Woods AT, Lawrence NS, Munafò M. The effects of gamelike features and test location on cognitive test performance and participant enjoyment. PeerJ 2016 Jul 6;4:e2184 [FREE Full text] [doi: 10.7717/peerj.2184] [Medline: 27441120]
38. Hawkins GE, Rae B, Nesbitt KV, Brown SD. Gamelike features might not improve data. Behav Res Methods 2013 Jun 9;45(2):301-318. [doi: 10.3758/s13428-012-0264-3] [Medline: 23055169]
39. Miranda AT, Palmer EM. Intrinsic motivation and attentional capture from gamelike features in a visual search task. Behav Res Methods 2014 Mar 22;46(1):159-172. [doi: 10.3758/s13428-013-0357-7] [Medline: 23835649]
40. Mahram B. Standardization of spielberger anxiety test in Mashhad. Allameh Tabataba'i University. 1994. URL: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=ICB0EFwAAAAJ&citation_for_view=ICB0EFwAAAAJ:Tiz5es2fbqcC [accessed 2024-01-01]
41. Christoforou M, Sáez Fonseca JA, Tsakanikos E. Two novel cognitive behavioral therapy-based mobile apps for agoraphobia: randomized controlled trial. J Med Internet Res 2017 Nov 24;19(11):e398 [FREE Full text] [doi: 10.2196/jmir.7747] [Medline: 29175809]
42. Muris P, Merckelbach H. A comparison of two spider fear questionnaires. J Behav Ther Exp Psychiatry 1996 Sep;27(3):241-244. [doi: 10.1016/s0005-7916(96)00022-5] [Medline: 8959425]
43. Szymanski J, O'Donohue W. Fear of spiders questionnaire. J Behav Ther Exp Psychiatry 1995 Mar;26(1):31-34. [doi: 10.1016/0005-7916(94)00072-t] [Medline: 7642758]
44. Botella C, Bretón-López J, Quero S, Baños R, García-Palacios A. Treating cockroach phobia with augmented reality. Behav Ther 2010 Sep;41(3):401-413. [doi: 10.1016/j.beth.2009.07.002] [Medline: 20569788]
45. Botella C, Quero S, Banos RM, Garcia-Palacios A, Breton-Lopez J, Alcaniz M, et al. Telepsychology and self-help: the treatment of phobias using the internet. Cyberpsychol Behav 2008 Dec;11(6):659-664. [doi: 10.1089/cpb.2008.0012] [Medline: 18991528]
46. Hemyari C, Zomorodian K, Shojaee M, Sahraian A, Dolatshahi B. The effect of personality traits on cognitive behavioral therapy outcomes in student pharmacists with rat phobia: a randomized clinical trial. Iran J Med Sci 2021 Jan;46(1):23-31 [FREE Full text] [doi: 10.30476/ijms.2019.82341.1045] [Medline: 33487789]
47. Toma E, Balan O, Lambru C, Moldoveanu A, Moldoveanu F. Ophiophobia 3D – a game for treating fear of snakes. In: Proceedings of the IEEE 10th International Conference on Intelligent Systems (IS). 2020 Presented at: IEEE 10th International Conference on Intelligent Systems (IS); August 28-30, 2020; Varna, Bulgaria. [doi: 10.1109/is48319.2020.9200189]
48. Mor S, Botella C, Campos D, Carlbring P, Tur C, Quero S. An internet-based treatment for flying phobia using 360° images: a feasibility pilot study. Internet Interv 2022 Apr;28:100510 [FREE Full text] [doi: 10.1016/j.invent.2022.100510] [Medline: 35242593]
49. Nielsen J. Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994 Presented at: CHI '94; April 24-28, 1994; Boston, MA. [doi: 10.1145/191666.191729]
50. 10 usability heuristics for user interface design. Nielsen Norman Group. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2024-01-01]
51. Sutcliffe A, Gault B. Heuristic evaluation of virtual reality applications. Interact Comput 2004 Aug;16(4):831-849. [doi: 10.1016/j.intcom.2004.05.001]
52. Desurvire H, Caplan M, Toth JA. Using heuristics to evaluate the playability of games. In: Proceedings of the CHI '04 Extended Abstracts on Human Factors in Computing Systems. 2004 Presented at: CHI EA '04; April 24-29, 2004; Vienna, Austria. [doi: 10.1145/985921.986102]
53. Lüddecke R, Felnhofer A. Virtual reality biofeedback in health: a scoping review. Appl Psychophysiol Biofeedback 2022 Mar;47(1):1-15 [FREE Full text] [doi: 10.1007/s10484-021-09529-9] [Medline: 34860290]
54. Kothgassner OD, Goreis A, Bauda I, Ziegenaus A, Glenk LM, Felnhofer A. Virtual reality biofeedback interventions for treating anxiety: a systematic review, meta-analysis and future perspective. Wien Klin Wochenschr 2022 Jan;134(Suppl 1):49-59 [FREE Full text] [doi: 10.1007/s00508-021-01991-z] [Medline: 34989862]
55. Dibbets P, Schruers K. An online spider game: overcome your fear, exposure is near. Comput Hum Behav Rep 2022 May;6:100201. [doi: 10.1016/j.chbr.2022.100201]
56. Pittig A, Treanor M, LeBeau RT, Craske MG. The role of associative fear and avoidance learning in anxiety disorders: gaps and directions for future research. Neurosci Biobehav Rev 2018 May;88:117-140. [doi: 10.1016/j.neubiorev.2018.03.015] [Medline: 29550209]
57. Albakri G, Bouaziz R, Alharthi W, Kammoun S, Al-Sarem M, Saeed F, et al. Phobia exposure therapy using virtual and augmented reality: a systematic review. Appl Sci 2022 Feb 05;12(3):1672. [doi: 10.3390/app12031672]

Abbreviations

BF: biofeedback
FCQ: Fear of Cats Questionnaire
FSQ: Fear of Spiders Questionnaire
HR: heart rate
STAI: State-Trait Anxiety Inventory
VR: virtual reality
VRET: virtual reality exposure therapy

Edited by T Leung; submitted 28.10.21; peer-reviewed by T Gargot, L Suppan, A Teles; comments to author 27.05.22; revised version received 17.10.22; accepted 10.07.23; published 06.03.24.

Please cite as:
Khaleghi A, Narimani A, Aghaei Z, Khorrami Banaraki A, Hassani-Abharian P
A Smartphone-Gamified Virtual Reality Exposure Therapy Augmented With Biofeedback for Ailurophobia: Development and Evaluation Study
JMIR Serious Games 2024;12:e34535
URL: https://games.jmir.org/2024/1/e34535
doi:10.2196/34535
PMID:38446522

Original Paper

Using a Virtual Reality Tool to Provide Primary Prevention Training in the Construction Field Following a Periodic Medical Visit: Cross-Sectional Study

Sylvain Chamot[1,2], MD, MSc; Isabelle Mahieu[1], MD; Marion Delzard[3], MD; Léa Leroy[1], MSc; Gwen Marhic[1,2]; Maxime Gignon[4,5,6], MD, PhD

[1]Regional Center for Occupational and Environmental Pathologies of Hauts-de-France, Amiens University Hospital, Amiens, France
[2]Péritox (UMR_I 01), UPJV/INERIS, Amiens, France
[3]Department of General Medicine, Amiens University Hospital, Amiens, France
[4]CRP-CPO, UR UPJV 7273, Université de Picardie Jules Verne, Amiens, France
[5]Education and Health Practices Laboratory UR3412, Sorbonne Paris Cité, University of Sorbonne Paris Nord, Bobigny, France
[6]Department of Prevention, Risks, Medical Information and Epidemiology, Amiens-Picardie University Hospital, Amiens, France

Corresponding Author:
Sylvain Chamot, MD, MSc
Regional Center for Occupational and Environmental Pathologies of Hauts-de-France
Amiens University Hospital
1 rond point du Pr Christian Cabrol
Amiens, 80000
France
Phone: 33 322087760
Email: chamot.sylvain@chu-amiens.fr

Abstract

Background: The construction field is highly concerned with the risk of work-related accidents, and training employees is difficult due to their small numbers in most companies.

Objective: This study aimed to study the impact of a virtual reality (VR) training tool following a periodic occupational health medical visit on the feeling of personal effectiveness in preventing occupational risks related to co-activity on a construction site.

Methods: We conducted a cross-sectional study with employees who had a periodic medical visit between April 1, 2022, and October 13, 2022, in a French occupational health service specializing in the construction field (Services Médicaux Interentreprises Bâtiment Travaux Publics [SMIBTP]). The employees were divided into 2 groups according to the training received: a medical visit alone or coupled with a session with a VR tool. We compared the scores for a “feeling of self-efficacy in occupational risk prevention” using the Fisher exact test.

Results: Of the 588 employees included, 210 had a medical visit alone, and 378 had a medical visit coupled with VR training. Training with the VR tool was associated with an increased “feeling of self-efficacy in occupational risk prevention.” The employees who benefited from the training reported a willingness to apply the advice given on prevention to a greater extent than those who did not, and they believed that risks on the worksite could be reduced using this tool.

Conclusions: Using VR training as a complement to periodic medical visits in an occupational health service improves the feeling of personal effectiveness in occupational risk prevention at the end of the training. If this trend is confirmed over a longer period of time, it could be an easily accessible prevention lever for employees in the future.

(JMIR Serious Games 2024;12:e49218) doi:10.2196/49218

KEYWORDS

virtual reality; virtual training tool; prevention; occupational medicine; construction

Introduction

Overview

The building and construction field is one of the most hazardous occupations in France. The main risks identified are the risk of road accidents, chemical risks, and musculoskeletal disorders, as well as risks related to the work environment and working with equipment. There are numerous work-related accidents, both fatal and nonfatal, with a major medicosocial impact on the individual and the community (long and costly medical care, prolonged absence from work negatively affecting the employer) [1]. In France, it is the leading field in terms of the frequency of work-related accidents, and the prevention of occupational risks remains difficult to achieve [2].

These difficulties are specifically related to collective prevention, which needs to be applied to a wide variety of tasks, sometimes in varying conditions and subject to change. Co-activities may be performed with employees from other companies in locations regularly situated very far from the company's head office. Small and medium-sized construction companies are the most affected by the lack of accessible risk prevention [3]. They also represent 99.8% of the companies and 45.7% of the jobs in France [4].

Health Monitoring of Employees in France

In France, all employees must undergo a periodic medical visit (at least every 5 years) in an occupational health services center. The main mission is to avoid any alteration in the employees' health due to their work. Thus, during these visits, employees receive advice to prevent occupational risks [5].

The Services Médicaux Interentreprises Bâtiment Travaux Publics (SMIBTP) is an occupational health service in charge of medical visits in the field of building and construction activity. In 2018, the SMIBTP monitored 2001 companies and 15,176 employees, mostly working in small and medium-sized companies. These companies work on a large number of sites throughout northern France. Since February 2022, the SMIBTP has been experimenting with a virtual reality (VR) training tool to train employees in the primary prevention of occupational risks related to co-activities on construction sites at one of its 2 consultation centers. The VR training is always preceded by a medical visit.

Virtual Reality

VR is a computer technology that involves real-time simulation and interaction through visual and auditory sensorial channels [6]. Computer-based 3D environments provide sensory information in a form similar to that received from the real world. VR allows individuals to experience and interact with or within environments with enhanced feedback [7-9]. To do this, users are required to be equipped with a VR headset that uses the principle of a stereoscopic 3D display connected to a computer interface to enable reproduction of the sensation of interaction with the artificial environment. The SMIBTP is the first occupational health service in France to have used a VR tool to provide additional prevention training for employees undergoing periodic medical visits. The goal of the SMIBTP is to provide employees with additional training in occupational risk prevention, with the aim of reducing the risk of accidents on site.

Objectives

No study has yet been conducted on periodic medical visits in an occupational health service coupled with a VR educational tool. A study in Finland compared VR with lecture-based safety training and found that the feeling of personal effectiveness in occupational risk prevention was increased by VR at the end of the training [10]. On a more general note, a review of the literature was carried out in 2023 on VR training and its impact on prevention, focusing in particular on the construction sector and its risks, highlighting that, although there appeared to be a positive impact, there was a lack of experimental studies in this field [11]. This was also highlighted in a meta-analysis published in 2023 [12]. However, it's important to keep in mind that these reviews pool together studies with different methods. Some studies are based on immersive technologies such as head-mounted displays, which rely on a computer connection [13], and mobile VR, which relies on the use of a smartphone [14]. Others have used less immersive techniques such as the Cave Automatic Virtual Environment, which involves virtual reality spaces where the walls, floor, and ceiling act as huge projection surfaces [15]. For the same method, the tools may vary (eg, headset brand), and above all, the context of the serious game may be very different (eg, risk prevention specific to certain trades vs risk prevention linked to co-activity on construction sites here).

The main objective of our study was to determine whether VR training had an impact on the feeling of self-efficacy in occupational risk prevention compared with a medical visit alone.

As a secondary objective, we wanted to know how the employees rated this additional VR training compared with the medical visit alone.

Methods

Design

This cross-sectional study included employees coming for a periodic medical visit to the SMIBTP who presented between April 1, 2022, and October 13, 2022, at one of the 2 centers.

The employees received 2 types of prevention training depending on the center in which they were examined. The employees in the first group had a medical visit coupled with VR (MV+VR group) training at the end of their periodic medical visit (Site A). The employees in the second group (Site B) had a medical visit alone (MV group).

Only employees performing manual work on construction sites were included in the study (engineers or secretaries were not included). In addition, in the VR group, only employees who completed the entire training (eg, no interruption due to motion sickness) were included.

The only exclusion criteria were an employee's past or present refusal of personal data collection and an insufficient knowledge of French.

Periodic Medical Visit

The role of occupational health services in France is to prevent any damage to workers' health caused by their work. The periodic medical visit is a preventive training tool used to this end. During these visits, workers can receive individual advice adapted to the workstation they occupy within their company. This involves oral advice, for example, on wearing personal protective equipment or using collective protective equipment. It may also, for example, involve a physical examination to assess the way employees bend over to pick up items from the floor. Finally, it may involve the delivery of paper documentation specific to the risks and workstations concerned.

Virtual Reality Tool

The VR training tool used by the SMIBTP is a serious game entitled SRC-Bâti VR (ViRtual Création), which aims to improve the occupational risk prevention skills of construction workers using VR digital simulation. SRC-Bâti VR emphasizes the co-activity aspect of construction sites and therefore the interaction between employees with very different workstations. Relative to a typical medical visit, it is less theoretical and more closely approaches real work, which is expected to have a positive impact in terms of prevention [16]. ViRtual Création was created in 2018 to develop software as an educational tool to improve worker safety.

The training sessions lasted between 7 minutes and 10 minutes, and a technician was present to equip the employee and explain how the device works. The training took place in a dedicated area of more than 10 m². The technician did not interfere during the training, except, for example, to prevent the employee from colliding with the equipment in the room.

During the training, the employees moved freely on a construction site. Workstations were clearly identified by markers. When the employee went to a workstation, a multiple choice question appeared about an accident risk at the workstation. If the employee did not answer correctly, the accident occurred, and a correction was provided. When an accident occurred, the employee's senses were stimulated to raise awareness of the risk. Workstations at which there was a risk of falling made a strong impression, as the impression of falling was real, as were situations in which there was a risk of being crushed. SRC-Bâti VR therefore offered a realistic simulation that served to teach skills in the prevention of occupational hazards linked to on-site co-activity. This realistic aspect gave a dimension of play to the VR simulation, with employees positively reacting to these virtual accidents, sometimes providing them with a simulation of what would happen (employees were never evaluated on their performance in the questions, which served only as an introductive teaching aid).

Depending on the employee's profession, 2 types of VR training were possible: one focused more on road work, and the other focused on building construction. Of the 20 possible workstations, 7 were randomly presented during the training, and 1 had to be present (possible workstations are shown in Table 1). No other customization was implemented in addition to the basic tool. Employees moved around the site by teleporting from one workstation to another over short distances, rather than gliding along, using joysticks. Although the training is short, the involvement of participants and interactivity and immersion offered by VR distinguish it from a simple paper questionnaire with the same questions (certainly greater involvement). The risks addressed were representative of the major risks on a construction site. Figure 1 illustrates how an employee is notified of a workstation, and an example of a workstation is shown in Figure 2. Figure 2 shows, on the left, the initial risky situation in which a truck backs up toward the employee in training and, on the right, the correction involves the employee moving away from the truck (green proposal). The red proposal indicates that the employee made the wrong choice before the correction and was run over by the truck. Demos can be viewed online [17]. The headset was a VIVE Focus 3 because, at the time the training was set up in 2021, it was the headset recommended by ViRtual Création and ViRtual Création was, at the time, the only French company identified by the SMIBTP that offered ready-made training material for building construction and road works. Since then, another solution dedicated to on-site risk prevention has appeared in France: VIRTUAL CONSTRUCT (Mimbus).

Table 1. Possible workstations depending on the training scenario in the Bâti VR serious game.

Scenario	Workstations common to both scenarios	Scenario-specific workstations
Road work	• Putting on personal protective equipment upon arrival on site[a] • Handling an unstable catwalk • Putting safety caps on the ends of iron bars • Waiting for trench walls to be reinforced to avoid being buried and limit machine traffic nearby • How to limit the risks associated with the vibrations from a jackhammer • How to deal with a truck backing up on a worksite • In front of an area cluttered with equipment, clearing a passageway without the possibility of falling objects before carrying out work in this space • In front of a colleague passing close to a load-lifting machine, informing the driver of the presence of the colleague to avoid any accidents • Using safety barriers when passing near holes in the ground	• Not driving past construction machinery but going around it by following the markings on studs • Using appropriate personal protective equipment when operating a circular saw • Using available handling equipment to carry loads instead of carrying them yourself • Replacing defective site signage • When laying asphalt on the road, wearing gloves, long sleeves, and pants for protection • Not working on a running construction machine engine • Using antipollution kits in the event of an accidental chemical spill on site • Bypassing work areas and following safe paths when moving around the site • When a construction machine reaches a buried network, stopping the machine and continuing work by hand • Warning a truck driver if he is going to hit a power cable when reversing • When climbing into a construction machine, always maintaining 3 points of support
Building construction	• Putting on personal protective equipment upon arrival on site[a] • Handling an unstable catwalk • Putting safety caps on the ends of iron bars • Waiting for trench walls to be reinforced to avoid being buried and limit machine traffic nearby • How to limit the risks associated with the vibrations from a jackhammer • How to deal with a truck backing up on a worksite • In front of an area cluttered with equipment, clearing a passageway without the possibility of falling objects before carrying out work in this space • In front of a colleague passing close to a load-lifting machine, informing the driver of the presence of the colleague to avoid any accidents • Using safety barriers when passing near holes in the ground	• How to avoid injury when carrying a heavy load • On scaffolding, limiting the risk of accidents by avoiding the presence of people working on several levels • Ventilating and vacuuming when using a sander • Before working on a pressurized water pipe, turning off the water supply completely • When using electrically-powered machines, never repairing the machine or its connections yourself • Using appropriate personal protective equipment when working near a colleague using a grinder • Disposing of rags soaked in chemical products after use • Reducing noise exposure by enclosing the compressor in dedicated rooms • Wearing appropriate gloves when welding • Alerting the colleague in charge of any abnormalities in load-bearing equipment • Using rolling scaffolding for occasional work at height

[a]Workstation mandatory for all training sessions.

Figure 1. Representation of a workstation accessible by the employee in the Bâti VR serious game.

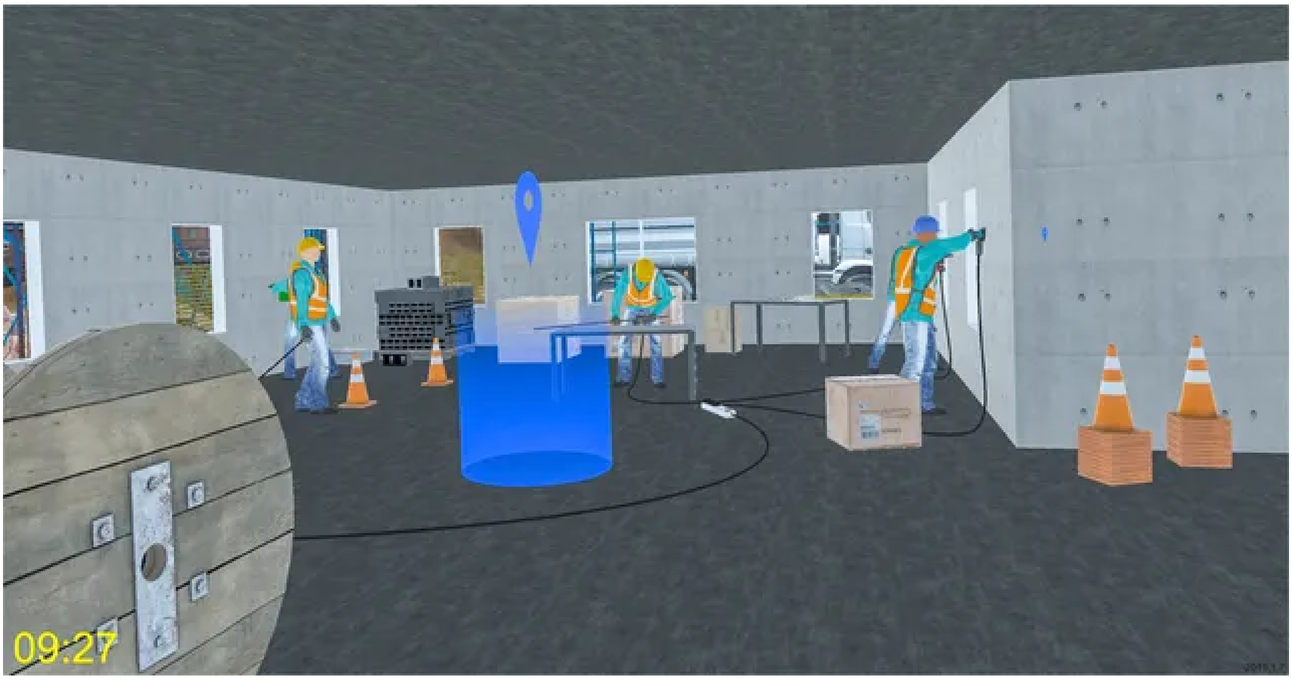
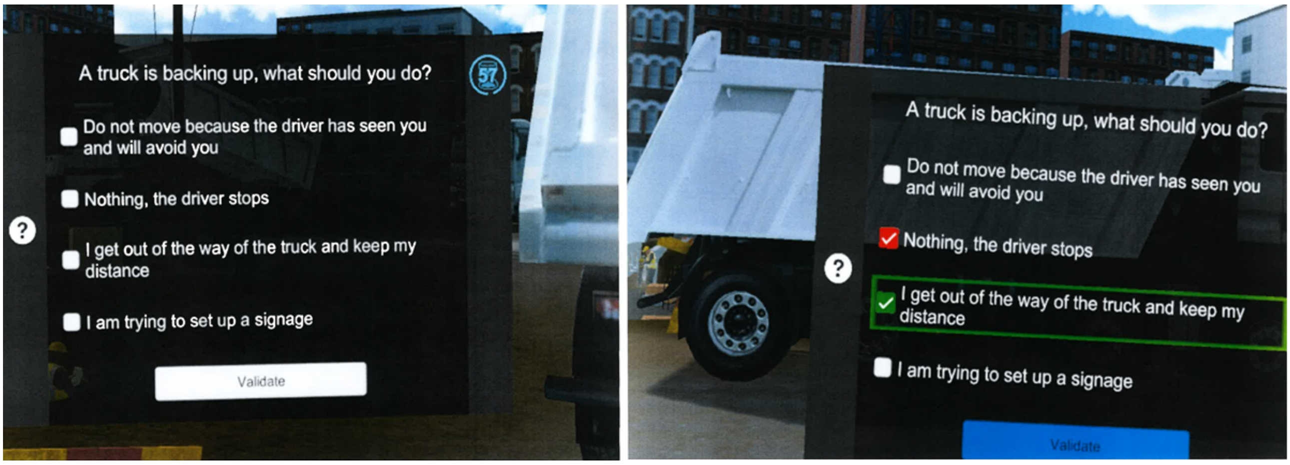


Figure 2. Example of a workstation in the Bâti VR serious game in which an employee is at risk of being run over by a truck.



Ethical Considerations

As this was a cross-sectional study evaluating current practice in the use of virtual reality, it did not require review by an institutional review board. Virtual reality was used independently of the study, with only an anonymous virtual reality evaluation questionnaire added by our teams. The study was carried out in compliance with good data protection practice, with the agreement of the data protection officer of the Université de Picardie Jules Verne. Our study was not funded by ViRtual Création and we did not collaborate with the company in the conduct of the study.

Data Collection and Variables

We collected data using a questionnaire built using the LIMESUVEY tool provided by the University of Picardie Jules Verne. The questionnaire was completed directly following the intervention (site A) or after the medical visit (site B). The data collected were demographics (age, gender, size of the company in which the employee worked), type of medical visit (with or without VR), and questions related to the feeling of self-efficacy and their rating of the training using 5-point Likert scales. These questions have not been validated and were defined by the authors. The questionnaire we used, based on the LIMESUVEY tool, also did not undergo a prior validation study. It was, however, partially based on the model for self-efficacy questions by Kirkpatrick and Kirkpatrick [18], which is a training evaluation method based on 4 levels: reaction, learning, behavior, and results. It enables assessment of the effectiveness of a training program at different levels, from participant reactions to concrete results for the employer. This model is widely used in training and human resources development to measure the impact of training programs. We only studied reactions, as our study design did not allow for employees to be contacted at a later date. The immediate reaction was assessed by the statement “I feel more effective in prevention.” We wished to address the question of what employees felt they could apply in practice just after their training, in particular regarding on-site co-activity, using the following 2 statements: “I am ready to apply these prevention rules” and “I think that these prevention rules can reduce the risks with regard to other colleagues on the site." The other 2 questions were aimed at evaluating the training received in itself: "My visit to the SMIBTP was worth it" and "I learned about prevention."

To explore gender, we asked employees to indicate whether they defined themselves as male or female.

Statistical Analysis

Employees were divided into 2 groups based on the 2 types of prevention training (MV *vs* MV + VR). The primary endpoint was a difference (as a percentage) between the responses of the 2 groups for each item (on our Likert scale) on questions relating to "feelings of self-efficacy in the prevention of occupational risks." The secondary endpoint was the difference (as a percentage) between the responses of the 2 groups to questions relating to the rating of the training. Responses measured on the Likert scales were not transformed into a quantitative variable, to not distort the nature of this mode of questioning.

Baseline demographics and clinical characteristics are expressed as means (SDs) or medians (IQRs) for numerical variables and frequencies (percentages) for categorical variables. Between-group comparisons were performed using the Mann-Whitney *U* test (for age) and Fisher exact test for categorical variables (size of the company and gender). The chi-squared test could not be used because the number of participants for certain response modalities was <5. The Fisher exact test was used to assess the association between the type of prevention training and the primary and secondary endpoints. A *P* value of .05 was considered significant for all tests.

All statistical tests were performed using R software (version 4.0.0, R Core Team, R Foundation for Statistical Computing). The data and R scripts are available on MENDELEY [19].

Results

During the study period (April 1, 2022, to October 13, 2022), 588 employees were recruited.

The baseline participant characteristics by type of prevention training are summarized in Table 2. The study population was predominantly male (571/588, 97.1%). The mean age was 33.15 (SD 12.1) years. By comparison, in 2019, men represented 87.89% of employees in the construction industry in France, and the mean age was 42 years [20,21]. There was not a statistically significant difference between the 2 groups in terms of gender, but there were statistically significant differences for age and company size. Of the 588 employees, there were 210 employees (35.7%) who had the medical visit alone (MV group) and 378 employees (64.3%) who had the medical visit coupled with VR training (MV+VR group). There were no missing data.

The results for the "feeling of self-efficacy in occupational risk prevention" are shown in Table 3. The MV+VR group had a greater feeling of self-efficacy in prevention than the MV group. For each question, there was a statistically significant difference at the 5% risk level, indicating that the MV+VR group felt more effective in prevention in general and, more specifically, in co-activity on worksites and would be more inclined to apply the prevention rules learned during their visit to the occupational health service.

The results of the ratings of the interventions received by the 2 groups are shown in Table 4. Employees in the MV+VR group found the intervention to be more useful and to provide more knowledge in terms of prevention than those in the MV group.

Table 2. Study population characteristics by type of prevention training at the Services Médicaux Interentreprises Bâtiment Travaux Publics (SMIBTP).

Characteristic	Overall (n=588)	MV[a] (n=210)	MV+VR[b] (n=378)	*P* value
Age (years), median (IQR)	32 (23-42)	38 (28-47.75)	29 (21-37)	<.001
Male, n (%)	571 (97.1)	201 (95.7)	370 (97.8)	.19
Size of the company				.01
1-10 employees	214 (36.4)	60 (28.6)	154 (40.7)	
11-49 employees	224 (38.1)	91 (43.3)	133 (35.2)	
50-299 employees	129 (21.9)	53 (25.2)	76 (20.1)	
≥300 employees	21 (3.6)	6 (2.9)	15 (4)	

[a]MV: medical visit.

[b]MV+VR: medical visit coupled with virtual reality training.

Table 3. Distribution of answers relating to the "feeling of self-efficacy" statements.

Questions	Responses, n (%)					*P* value
	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	
I feel more effective in prevention.						.002
MV[a] (n=210)	77 (36.7)	87 (41.4)	32 (15.2)	3 (1.4)	11 (5.2)	
MV+VR[b] (n=378)	139 (36.8)	197 (52.1)	36 (9.5)	2 (0.5)	4 (1.1)	
I am ready to apply these prevention rules.						<.001
MV (n=210)	118 (56.2)	70 (33.3)	16 (7.6)	2 (1)	4 (1.9)	
MV+VR (n=378)	277 (73.3)	96 (25.4)	5 (1.3)	0	0	
I think that these prevention rules can reduce the risks with regard to other colleagues on the site.						<.001
MV (n=210)	117 (55.7)	66 (31.4)	20 (9.5)	2 (1)	5 (2.4)	
MV+VR (n=378)	282 (74.6)	90 (23.8)	5 (1.3)	0	1 (0.3)	

[a]MV: medical visit.
[b]MV+VR: medical visit coupled with virtual reality training.

Table 4. Distribution of answers relating to the evaluation of the training.

Statements	Responses, n (%)					*P* value
	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	
My visit to SMIBTP[a] was worth it.						.002
MV[b] (n=210)	129 (61.4)	65 (31)	13 (6.2)	1 (0.5)	2 (1)	
MV+VR[c] (n=378)	268 (70.9)	104 (27.5)	6 (1.6)	0	0	
I learned about prevention.						<.001
MV (n=210)	75 (35.7)	93 (44.3)	19 (9)	4 (2)	19 (9)	
MV+VR (n=378)	186 (49.2)	158 (41.8)	25 (6.6)	3 (0.8)	6 (1.6)	

[a]SMIBTP: Services Médicaux Interentreprises Bâtiment Travaux Publics.
[b]MV: medical visit.
[c]MV+VR: medical visit coupled with virtual reality training.

Discussion

Principal Findings

The results of our study show that the use of a VR training tool at the end of periodic occupational medical visits had an impact on the feeling of self-efficacy in terms of occupational risk prevention in the construction field. This is an important finding, suggesting that the use of VR could have a significant impact on the occupational risk prevention practices of construction site employees. This is a useful finding, given that all employees in France systematically and regularly have such medical visits. Our results highlight a potentially important lever for the prevention of occupational risks in the construction field in the future through the improvement of employee competence.

Other Uses of Virtual Reality in the Health Field

Eiris et al [22] sought to validate safety training using 360-degree augmented reality panoramas. Their study showed the interest in the use of this method in the identification and recognition of hazards on construction sites. However, the rate of hazard identification was quite low, as only 30% of the hazards were identified by the participants. They explained this by the fact that their population was composed of students specializing in construction management (n=30) and were not building and construction professionals. They also emphasized the constructive comments concerning the ease of use of the platform, feedback that we also had in our study using VR. In our study, we did not analyze the responses to the questions asked during the VR training, as this did not correspond to our research question.

Nykänen et al [10] evaluated both the effectiveness of an immersive VR-based safety training program and a participatory human factors safety training program. The study was conducted with 119 employees working on 8 construction sites in Finland. The employees evaluated the training with questionnaires at the start, immediately after the intervention, and at a 1-month follow-up. They considered VR to be a serious tool for improving prevention skills and found that it motivated them to apply prevention rules more than after safety training based on passive learning methods. This study was conducted only with employees of medium-sized and large companies.

Simeonov et al [23] investigated the value of reducing mechanical vibration of support structures used as walking or working surfaces when performing construction tasks at height (falls from height account for one-third of fatal accidents in construction). Employees (n=12) used instrument-carrying gel insoles connected to a VR system to test sensory perception of the feet. The study did not show any effectiveness for this technology in 2008, but given the evolution of VR technologies, it is possible that the results would be different today.

We also found studies that assessed the use of VR as a prevention and training tool in fields other than construction.

The mining industry is a field in which the risk of serious accidents and fatalities is very high. Filigenzi et al [24] highlighted the value of using VR to train surface and underground mine employees and rescue personnel in hazard recognition and evacuation routes and procedures. This study, carried out in 2000, was innovative, demonstrated possibilities, and generated interest in extending such an approach to other fields of high-risk activity, such as construction, agriculture, and the oil industry.

In the logistics field, the use of handling equipment is responsible for a large number of occupational accidents, in particular to third parties. Choi et al [25] focused on forklift drivers, conducting a study with 20 students at Hong Kong Polytechnic University specializing in construction engineering. Their goal was to investigate how a forklift driver's situational awareness of others around him can be influenced by the type of subtasks he performs. A VR environment was used as the experimental environment in which participants performed a series of subtasks, such as driving, turning, reversing, loading, and unloading: the more concentration that was required for the tasks, the higher the risk of an accident. The authors concluded that it would be beneficial to not only use additional safety devices (such as person detection devices) but also have more detailed safety training, making VR meaningful.

In the area of electrical risk training, in 2015, Zhao and Lucas [26] reported that human error was responsible for approximately 50% of all electrical-related fatalities in multiple industries in the United States. They hypothesized that effective employee safety training programs, including VR, would be the most direct approach to mitigate such errors. Their study showed the success of using VR, highlighting training that effectively visualizes invisible risks without endangering employees. Such training increases awareness of the risk and trains employees to use the necessary protective equipment.

In the health care field, VR interventions appear to be an effective tool to boost the intention to be vaccinated [27-29].

The results of our study, as well as those of others in various fields, show that VR training tools hold great potential and should be further developed to improve the prevention of occupational risks, particularly in the construction field.

Strengths and Limitations

One of the strengths of our study is that it was conducted with a large population and 2 groups who were similar in terms of gender. In addition, the completion rate was 100% due to the use of a short and acceptable questionnaire.

However, the study population was mainly composed of men, which did not allow us to obtain data on the female population in the construction field. Women are not as well represented as men in the national population of construction employees. In addition, this intervention was intended only for certain construction jobs, mainly on construction sites, where women are much less present. The female population is mainly present in the administrative field of construction and public works companies and is therefore not subject to the same occupational risks.

Employees in the MV+VR group were younger than those in the MV group, which is similar to the overall population of construction employees in France. This result was expected, given the appetite of the younger generation for new technologies, such as VR. This age difference suggests that, if this tool is deployed on a larger scale, the older portion of the construction employee population might not benefit from it, as they may not want to use it.

The employees in the MV+VR group were also more often from small companies, which can be explained by the fact that they were the target population for the occupational health service. It is possible that this influenced our results, as larger companies have more resources for prevention. The employees of larger companies might therefore find this training less useful, but we believe that this does not affect the interpretation of our results.

It should also be noted that the use of VR is already a common practice in occupational health services and that our study did not change these practices, apart from the addition of the questionnaire. We therefore believe that our intervention did not bias our results.

On the other hand, we excluded individuals with the least mastery of the French language from our study. Individuals in this group are among those most at risk of having an accident at work due to the language barrier, in particular because of difficulties in understanding safety instructions. This does not call into question the validity of our results but highlights this group’s limited access to prevention through this tool. A translated version could be envisaged.

All the employees who participated were only seen once by the SMIBTP. It was therefore not possible to evaluate the impact of repeating these VR training sessions. Similarly, the design of the study did not allow an evaluation of the impact of this training at a later date. This was a major limitation of our study. Although these results are encouraging, other studies are needed to evaluate the long-term impact of VR training on the knowledge and perception of personal effectiveness in preventing occupational hazards. Longer-term studies are also needed to study the tool’s impact in terms of reducing the occurrence of occupational accidents.

In the context of our study, no data were collected that could be used to identify employees. Our objective was to reproduce, as closely as possible, the real-life conditions of using the tool, and we knew that collecting identification data could have significantly reduced participation in our study. If we were to carry out an evaluation at a later time point, it would logically be conducted under the normal conditions of periodic medical visits in occupational health services and therefore completed within 2 years to 5 years following our study. We would ask the employees coming for a visit whether they had already received training via VR. If so, we would request that they complete a questionnaire.

Further studies will be needed to assess the acceptability of VR. Indeed, one of the classic side effects of VR is motion sickness, and some VR sessions had to be interrupted because of symptoms such as nausea [30]. VR can also alter sensorimotor and perceptual abilities, with effects that can last several hours after exposure, and cause visual fatigue and headaches [31].

The routine use of VR during medical visits by occupational health services could have an impact on occupational risk prevention in the construction field. It could be a tool of major importance, given its accessibility, but its long-term impact and accessibility need to be assessed.

Acknowledgments

The authors would like to thank Ms Angélique Verzele, Dr Gérard Araszkiewirz, Dr Sébastien Amasse and the entire SMIBTP team for developing this virtual reality tool in their department.

Conflicts of Interest

None declared.

References

1. Pinto A, Nunes IL, Ribeiro RA. Occupational risk assessment in construction industry – Overview and reflection. Safety Science 2011 Jun;49(5):616-624. [doi: 10.1016/j.ssci.2011.01.003]
2. Rapport annuel 2020 de l'Assurance Maladie - Risques professionnels. Caisse nationale de l'Assurance Maladie. 2021 Dec 06. URL: https://assurance-maladie.ameli.fr/etudes-et-donnees/2020-rapport-annuel-assurance-maladie-risques-professionnels [accessed 2024-02-18]
3. Lecocq C, Dupuis B, Forest H, Lanouziere H. Santé au travail : vers un système simplifié pour une prévention renforcée. Republique Francaise. 2018 Aug 28. URL: https://www.vie-publique.fr/rapport/27019-sante-au-travail-vers-un-systeme-simplifie-pour-une-prevention-renforcee [accessed 2024-02-18]
4. Les entreprises en France. Institut national de la statistique et des études économiques. 2018 Nov 08. URL: https://www.insee.fr/fr/statistiques/3639594 [accessed 2024-02-18]
5. Caroly S, Landry A, Cholez C, Davezies P, Bellemare M, Poussin N. Innovation in the occupational health physician profession requires the development of a work collective to improve the efficiency of MSD prevention. Work 2012;41 Suppl 1:5-13. [doi: 10.3233/WOR-2012-0128-5] [Medline: 22316692]
6. Rutkowski S, Rutkowska A, Jastrzębski D, Racheniuk H, Pawełczyk W, Szczegielniak J. Effect of virtual reality-based rehabilitation on physical fitness in patients with chronic obstructive pulmonary disease. J Hum Kinet 2019 Oct;69(1):149-157 [FREE Full text] [doi: 10.2478/hukin-2019-0022] [Medline: 31666897]
7. Rutkowski S, Rutkowska A, Kiper P, Jastrzebski D, Racheniuk H, Turolla A, et al. Virtual reality rehabilitation in patients with chronic obstructive pulmonary disease: a randomized controlled trial. COPD 2020 Jan;Volume 15:117-124. [doi: 10.2147/copd.s223592]
8. Namkoong K, Chen J, Leach J, Song Y, Vincent S, Byrd A, et al. Virtual reality for public health: a study on a VR intervention to enhance occupational injury prevention. J Public Health (Oxf) 2023 Mar 14;45(1):136-144. [doi: 10.1093/pubmed/fdab407] [Medline: 35051993]
9. Rutkowski S, Kiper P, Cacciante L, Cieślik B, Mazurek J, Turolla A, et al. Use of virtual reality-based training in different fields of rehabilitation: A systematic review and meta-analysis. J Rehabil Med 2020 Nov 19;52(11):jrm00121 [FREE Full text] [doi: 10.2340/16501977-2755] [Medline: 33073855]
10. Nykänen M, Puro V, Tiikkaja M, Kannisto H, Lantto E, Simpura F, et al. Implementing and evaluating novel safety training methods for construction sector workers: Results of a randomized controlled trial. J Safety Res 2020 Dec;75:205-221. [doi: 10.1016/j.jsr.2020.09.015] [Medline: 33334479]
11. Stefan H, Mortimer M, Horan B. Evaluating the effectiveness of virtual reality for safety-relevant training: a systematic review. Virtual Reality 2023 Aug 19;27(4):2839-2869. [doi: 10.1007/s10055-023-00843-7]
12. Man SS, Wen H, So BCL. Are virtual reality applications effective for construction safety training and education? A systematic review and meta-analysis. Journal of Safety Research 2023 Dec:1. [doi: 10.1016/j.jsr.2023.11.011]
13. Abotaleb I, Hosny O, Nassar K, Bader S, Elrifaee M, Ibrahim S, et al. An interactive virtual reality model for enhancing safety training in construction education. Comp Applic In Engineering 2022 Nov 23;31(2):324-345. [doi: 10.1002/cae.22585]
14. Bhagwat K, Kumar P, Delhi VSK. Usability of visualization platform–based safety training and assessment modules for engineering students and construction professionals. J. Civ. Eng. Educ 2021 Apr;147(2):1. [doi: 10.1061/(asce)ei.2643-9115.0000034]

15. Sacks R, Perlman A, Barak R. Construction safety training using immersive virtual reality. Construction Management and Economics 2013 Sep;31(9):1005-1017. [doi: 10.1080/01446193.2013.828844]
16. Clot Y. Le travail à coeur. Paris, France: La Découverte; Jun 28, 2010.
17. Virtual Création. YouTube. 2021 Mar 02. URL: https://www.youtube.com/watch?v=BZgBInXTiWc [accessed 2023-12-12]
18. Kirkpatrick DL, Kirkpatrick JD. Implementing the Four Levels: A Practical Guide for Effective Evaluation of Training Programs. San Francisco, CA: Berrett-Koehler Publishers; 2007.
19. Chamot S. VR SMIBTP. Mendeley Data. 2023 Jan 05. URL: https://data.mendeley.com/datasets/m323fk6fhg/1 [accessed 2024-02-18]
20. Les chiffres clés du BTP: Les salariés du BTP. Observatoire des métiers du BTP. 2019. URL: https://dataviz.metiers-btp.fr/salaries/age [accessed 2022-10-13]
21. Les salariés du BTP. CIBTP. 2019. URL: https://www.cibtp.fr/documentation/rapports-statistiques/les-salaries-du-btp [accessed 2022-10-13]
22. Eiris R, Gheisari M, Esmaeili B. PARS: using augmented 360-degree panoramas of reality for construction safety training. Int J Environ Res Public Health 2018 Nov 03;15(11):1 [FREE Full text] [doi: 10.3390/ijerph15112452] [Medline: 30400318]
23. Simeonov P, Hsiao H, Powers J, Ammons D, Kau T, Amendola A. Postural stability effects of random vibration at the feet of construction workers in simulated elevation. Appl Ergon 2011 Jul;42(5):672-681. [doi: 10.1016/j.apergo.2010.10.002] [Medline: 21071015]
24. Filigenzi MT, Orr TJ, Ruff TM. Virtual reality for mine safety training. Appl Occup Environ Hyg 2000 Jun 30;15(6):465-469. [doi: 10.1080/104732200301232] [Medline: 10853286]
25. Choi M, Ahn S, Seo J. VR-Based investigation of forklift operator situation awareness for preventing collision accidents. Accid Anal Prev 2020 Mar;136:105404. [doi: 10.1016/j.aap.2019.105404] [Medline: 31927452]
26. Zhao D, Lucas J. Virtual reality simulation for construction safety promotion. Int J Inj Contr Saf Promot 2015;22(1):57-67. [doi: 10.1080/17457300.2013.861853] [Medline: 24471426]
27. Vandeweerdt C, Luong T, Atchapero M, Mottelson A, Holz C, Makransky G, et al. Virtual reality reduces COVID-19 vaccine hesitancy in the wild: a randomized trial. Sci Rep 2022 Mar 17;12(1):4593 [FREE Full text] [doi: 10.1038/s41598-022-08120-4] [Medline: 35301359]
28. Mottelson A, Vandeweerdt C, Atchapero M, Luong T, Holz C, Böhm R, et al. A self-administered virtual reality intervention increases COVID-19 vaccination intention. Vaccine 2021 Nov 05;39(46):6746-6753 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.004] [Medline: 34654579]
29. Plechatá A, Vandeweerdt C, Atchapero M, Luong T, Holz C, Betsch C, et al. Experiencing herd immunity in virtual reality increases COVID-19 vaccination intention: Evidence from a large-scale field intervention study. Comput Human Behav 2023 Feb;139:107533 [FREE Full text] [doi: 10.1016/j.chb.2022.107533] [Medline: 36277032]
30. Leung AK, Hon KL. Motion sickness: an overview. Drugs Context 2019;8:1 [FREE Full text] [doi: 10.7573/dic.2019-9-4] [Medline: 32158479]
31. Anses. Expositions aux technologies de réalité virtuelle et/ou augmentée. Anses. 2021 Jun. URL: https://www.anses.fr/fr/system/files/AP2017SA0076Ra.pdf [accessed 2024-02-18]

Abbreviations

MV: medical visit
SMIBTP: Services Médicaux Interentreprises Bâtiment Travaux Publics
VR: virtual reality

Edited by T Leung, R Baranyi; submitted 23.05.23; peer-reviewed by C Yang, Y Chen; comments to author 30.10.23; revised version received 23.11.23; accepted 15.01.24; published 15.03.24.

Please cite as:
Chamot S, Mahieu I, Delzard M, Leroy L, Marhic G, Gignon M
Using a Virtual Reality Tool to Provide Primary Prevention Training in the Construction Field Following a Periodic Medical Visit: Cross-Sectional Study
JMIR Serious Games 2024;12:e49218
URL: https://games.jmir.org/2024/1/e49218
doi:10.2196/49218
PMID:38488851

Original Paper

Impact of Virtual Reality–Based Group Activities on Activity Level and Well-Being Among Older Adults in Nursing Homes: Longitudinal Exploratory Study

Yijun Li[1], MSc; Carlotta Wilke[1], MSc; Irina Shiyanov[2], BSc; Beate Muschalla[1], Prof Dr

[1]Department of Psychotherapy and Diagnostics, Technische Universität Braunschweig, Institute of Psychology, Braunschweig, Germany

[2]VirtuaLounge GmbH, Braunschweig, Germany

Corresponding Author:
Yijun Li, MSc
Department of Psychotherapy and Diagnostics
Technische Universität Braunschweig
Institute of Psychology
Humboldtstraße 33
Braunschweig, 38106
Germany
Phone: 49 0531 391 3603
Email: yijun.li@tu-braunschweig.de

Abstract

Background: In addition to illness, inactivity is a risk factor for high mortality in nursing homes. Using innovative technology, such as virtual reality (VR), for meaningful group activities could provide new opportunities for solving this problem. VR interventions have already been approved as a promising method for enhancing the health of older adults.

Objective: In this study, we examined whether VR-based group activities can have a positive impact on activity level and group interaction among older adults living in nursing homes.

Methods: We conducted a longitudinal study and provided VR interventions as a group activity once a week for 4 consecutive weeks in nursing homes. Participants were recruited based on the experience of the nursing staff members and the natural decisions of the older adults. Within a virtual cottage, designed according to the needs of the target group, older adults were able to perform daily tasks that they were no longer able to do in real life, such as gardening and making pizza. Overall, 2 psychologists measured the psychosocial capacities, activities of daily life, and well-being before and after the interventions using standardized instruments.

Results: The results focus on a total of 84 older adults from 14 nursing homes who completed at least 3 VR interventions. The results indicate that several psychosocial capacities among the older adults improved, including adherence to regulations ($P<.001$; $\eta^2=0.122$), flexibility ($P<.001$; $\eta^2=0.109$), and group integration ($P<.001$; $\eta^2=0.141$). Problems related to competence also showed a slight decrease ($P=.04$; $\eta^2=0.039$). In addition, the VR intervention promoted their proactivity ($P<.001$; $\eta^2=0.104$) and mobility ($P=.04$; $\eta^2=0.039$). During the VR group intervention, older adults' well-being could be maintained at a high level. The results highlight the beneficial effects of VR intervention as a meaningful activity in nursing homes, showcasing the potential of VR applications in this setting.

Conclusions: This study provides a novel and naturalistic perspective, offering new insights into the use of VR in nursing homes. The VR intervention was well accepted and fulfilled the aim of enhancing capacity and well-being. It could be a meaningful group activity in nursing homes to improve social group interaction. To provide stronger evidence, randomized controlled trials are necessary.

(JMIR Serious Games 2024;12:e50796) doi:10.2196/50796

KEYWORDS

virtual reality; group activity; aging care; older adults; meaningful activity; mental health; well-being; social interaction; psychosocial capacities; activity of daily living

Introduction

Background

As a result of demographic changes and the development of physical and mental illnesses in the older adult population, residing in a nursing home for assisted living is a commonly chosen solution in later life. The perspectives of older adults residing in nursing homes are characterized by a prevailing sense of awaiting death and a dearth of activities that foster a sense of purpose and fulfillment [1]. Vossius et al [2] conducted a longitudinal study involving 690 older adults living in nursing homes over a span of 3 years. The median survival time in nursing homes was approximately 2.2 (95% CI 1.9-2.4) years. The annual mortality rate was approximately 30% [2,3].

On the one hand, the high mortality rate can be explained by the baseline health situation and comorbidity of the residents [3,4]. On the other hand, there is a loss of activities of daily living (ADLs). It is possible to reduce the risk of mortality by improving ADLs to maintain physical functioning [2]. ADLs encompass essential daily activities and mobility, such as eating and using the toilet. Ouden et al [5] observed a significant number of inactive older adults in nursing homes. Most of them were observed to be in a lying or sitting position [5]. A considerable proportion (67%) of the older population engaged in sedentary behavior for >8.5 hours per day [6]. This sedentary behavior and lack of communicative activity have critical implications for the prevention of physical, psychological, and social health problems [7]. This phenomenon also indicates social isolation and loneliness among older adults in nursing homes [8,9]. Older adults in nursing homes tend to be lonelier than community-dwelling older adults, even though they are often surrounded by other residents and caregivers [9-11]. Connecting with individuals with varying cognitive fitness levels is challenging in nursing home settings [10].

Several studies have been conducted regarding how loneliness and social isolation in nursing homes negatively affect mental and physical health, well-being, and mortality [12-15]. Studies have consistently shown that both loneliness and social isolation are associated with various mental health issues, including depression, feelings of hopelessness, and cognitive impairment [16,17]. Older adults residing in nursing homes experienced elevated levels of loneliness and anxiety during the COVID-19 pandemic compared to those receiving home care [18]. Overall, 69% of older individuals in nursing homes reported feelings of loneliness and 63% reported anxiety. On the other hand, among those receiving home care, 53% reported loneliness and 47% reported anxiety [18]. In addition, these conditions have been linked to impaired motor function, cardiovascular health problems, disrupted sleep, and increased frailty [13,19]. Zhao et al [20] found that higher engagement in activities was associated with lower levels of loneliness and frailty among older adults in nursing homes. The authors emphasized the importance of developing strategies to increase social and activity engagement in this population. Higher levels of activity engagement and meaningful relationships have been linked to greater satisfaction, well-being, and quality of life [15,21,22]. Therefore, there is a need to develop strategies that focus on improving ADLs, promoting engagement in activities, and enhancing social interactions among nursing home residents. These strategies have the potential to enhance overall well-being and quality of life and potentially reduce loneliness, social isolation, and mortality among older individuals residing in nursing homes [1,20,23].

Enhancing activity and social interaction among older adults can be effectively supported through meaningful daily group activities [24]. Participating in group activities fosters a sense of belonging, promotes social engagement, and contributes to overall well-being [24]. It provides opportunities for increased social interaction with fellow residents and emotional support through participating in games and identifying with teams [24]. It is also important that these activities are “meaningful” to the residents. Research by Tak et al [25] demonstrated that if activities are not relevant or meaningful to the residents, they may prefer to do nothing or passively watch television. Meaningful group activities are described as those that hold significance or provide enjoyment for individuals, aligning with their current and past interests, routines, habits, and roles and improving their mental or physical function [26-29]. It has been shown that meaningful activities enhance social engagement and well-being and reduce loneliness among older adults living in nursing homes [30]. Nevertheless, there are several barriers to providing meaningful group activities in nursing homes. One major challenge is the shortage of personnel. Nursing homes are already facing difficulties in filling nursing home positions due to a shortage of skilled workers [31], and this is expected to persist and worsen in the coming years. Insufficient staffing limits the capacity to organize several meaningful activities [32,33]. In addition, there may be constraints related to limited space and equipment within nursing homes [33]. Therefore, there is a need to develop new, low–resource dependent, easily applicable, meaningful group activities [34,35]. Moreover, the demands of older adults in need of care are evolving, including their expectations regarding the technical equipment in nursing homes. On the basis of a population survey conducted in Germany between 2009 and 2014, only a small percentage of older adults aged >65 years used smartphones, but by 2019, more than half of them were already using these devices. In addition, internet use has also experienced significant growth since 2009, with 74% of the older adult population using the internet in 2019 [35]. Therefore, introducing innovative, technology-based interventions such as virtual reality (VR) [36] for group activities could be a promising solution to improve social connections and activities for older adults in nursing homes [12,37,38].

Fully immersive VR has emerged as a feasible method of intervention in older adults’ rehabilitation and aging care [36,39]. By using head-mounted devices (HMDs) and controllers, VR technology provides users with a fully immersive experience and a sense of presence. The unique characteristics of VR offer a viable solution to the barriers faced in nursing homes [33]. These barriers include the health status of older adults, limited physical space in the environment, and organizational challenges such as staffing shortages and funding constraints. First, VR proves to be accessible and accommodating for older adults with limited mobility. For example, individuals can remain seated in a wheelchair while experiencing the sensation of being on a mountain surrounded by stars. This enables older adults to engage in activities that would otherwise be physically challenging or impossible. Moreover, VR interventions ensure safety as they eliminate environmental risks and can be paused at any moment, which is particularly important for populations considered vulnerable such as older adults [40]. Second, VR interventions are flexible and require minimal physical space, similar to the space requirements for small-group activities [41]. This flexibility empowers the staff members to efficiently organize activities, while also reducing the costs associated with transportation. Moreover, the effectiveness of VR interventions for older adults has been demonstrated. The immersive and presence-inducing nature of VR has shown evidence comparable to that of traditional interventions in mental health [42]. Using VR as a medium to improve ADLs for older adults has already shown promising effects [43], and it was found to be effective in reducing loneliness and social isolation [44-47]. In the field of mental health, VR is considered an effective method for training and therapy for cognitive functions and for enhancing the well-being of older adults [48-50]. In addition, several studies have reported high acceptance of VR experiences among older adults [39,51,52].

However, recent interventions have primarily focused on singular concepts such as training or entertainment, and there is a lack of studies exploring VR interventions as daily group events in nursing homes that provide opportunities for older adults to connect and interact socially. In this study, we developed and evaluated a series of VR interventions, aiming to enhance the overall health and well-being of older adults in nursing homes. The VR interventions combine training activities and entertainment to create meaningful group ADLs in nursing homes. As discussed previously, meaningful activities should align with older adults' current and past interests, routines, habits, and roles [28,29]. Studies have shown that older adults who engage in daily and household activities experience less decline in mobility [53]. For example, gardening has been recognized as a promising activity for reducing loneliness and improving socialization [54,55]. Building upon these findings, our VR group intervention focuses on providing older adults in nursing homes with virtual environments that allow them to engage in daily tasks they may no longer be able to perform, such as baking a pizza, handcrafting, and gardening. Through these simulated activities, older adults have the opportunity to experience the fulfillment of completing familiar tasks, while preserving and enhancing their physical and mental functioning. By using tasks that are familiar to them, we aimed to reduce fear or demotivation, which might occur when being confronted with new technology. The meaningful activities chosen for the VR game offer older adults an enjoyable experience and contribute to their overall well-being. Organizing these daily activities in a virtual environment incurs lower costs in terms of time and equipment compared to real-life implementation. In addition to the VR scenario, we developed an aid system and an automated program that enable staff members to easily facilitate VR group sessions with older adults.

In summary, this exploratory study investigated the effectiveness of implementing VR interventions as meaningful group activities for older adults in nursing homes. The primary focus was on evaluating the older adults' activity and mobility levels, well-being, social interaction, and mental capacities over the course of a 4-week VR group intervention. By supporting well-being and psychosocial capacities, these interventions have the potential to address key challenges faced by older adults in nursing homes.

Research Question and Hypotheses

This observational intervention study examines the following question: Does VR-based group activity have any positive impact on the daily lives of older adults living in nursing homes? The following hypotheses were tested:

- Hypothesis 1: Over the course of a 4-week VR intervention, psychosocial capacities and ADLs of older adults in nursing homes will remain stable or even improve.
- Hypothesis 2: Over the course of a 4-week VR intervention, older adults' well-being will remain stable or even improve.

Methods

Study Design

This longitudinal study using pre-post measures was conducted in naturalistic settings in nursing homes in a city with 250,000 inhabitants in Germany. After contacting all 31 local nursing homes, a total of 15 (48%) nursing homes chose to participate in the project.

Selection of Participants

We contacted all nursing homes by telephone. Subsequently, those nursing homes expressing interest were provided with a comprehensive briefing via email outlining the selection criteria. Participants were selected from the nurses in the nursing homes, who considered both the basic data (eg, medical history) about the older adults from the nursing home information system and their extensive experience in assisting older adults with their daily living, while also assessing the older adults' willingness to participate. The selection criteria were as follows: (1) older adults were aged >60 years; (2) they had at least 1 arm and 1 hand that they were able to use (this ensures their interaction with the virtual environment); (3) they were still able to see and hear, and the use of audio aids or glasses was permitted; (4) they were able to participate in an oral interview with the researchers and did not have severe dementia; and (5) they did not have diseases, such as epilepsy, that are contraindications for VR activities.

We set a control group that underwent the same measurement procedure as the intervention group, except that they did not undergo the VR intervention phase. The control group participants were chosen based on the advice from the nursing staff members. The older adults in the control group had a low willingness to participate in the VR intervention and expressed a preference for interviews. Of the 15 nursing homes, 1 (7%) chose to solely participate in the control group due to low willingness to organize new group events and expressed a preference for participating in the project only through interviews with older adults.

Procedure and Data Collection

The interventions were conducted for approximately 3 months in each participating nursing home. During this period, the intervention group went through 3 phases: premeasurement phase, 4 weeks of VR group intervention, and postmeasurement phase (Table 1). During the premeasurement phase, each older adult was offered an individual initial appointment. During this appointment, a psychologist explained the schedule, privacy agreement, and consent form to the older adult. One week after the initial appointment, the older adult underwent the first pretest interview (T0; baseline) with a psychologist. This baseline assessment covered mental capacity, ADLs, and well-being. In the following week, a brief, second pretest (T1) was conducted to familiarize the older adults with the well-being assessment. This served as a warm-up assessment and was repeated after each subsequent intervention session. The intervention phase began after the premeasurement phase. Group interventions were conducted every week for 4 weeks (T2-T5). These sessions were facilitated by a project psychologist and a technical assistant. Their role was to introduce the older adults to the VR program and provide support throughout the VR sessions. There were always 3 to 5 older adults in a group for an intervention. During each VR intervention session, the older adults were presented with tasks to solve in a virtual environment. After completing the tasks, the older adults participated in a 3-minute virtual tour of a landscape to relax. Following each VR group session, the older adults' individual well-being was assessed. One week after completing all 4 VR group interventions, the same psychologist who conducted the pretest interviews assessed the older adults' mental capacities, ADLs, and well-being in the posttest phase through a postintervention interview (T6). A follow-up interview (T7) with the same content was conducted 3 weeks after the posttest phase to assess the stability of the posttest results.

Table 1. Procedure of the virtual reality (VR) group intervention study in a nursing home (as per the focus of this paper)[a].

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 10
Event	• Preintervention assessment 1 (T0)	• Preintervention assessment 2 (T1)	• VR activity 1 (T2)	• VR activity 2 (T3)	• VR activity 3 (T4)	• VR activity 4 (T5)	• Postintervention assessment 1 (T6)	• Follow-up (T7)
Content	• Baseline demographic information • Mini-ICF-APP[b] • ADL[c] • WHO-5[d]	• Warm-up • WHO-5	• VR group activity • WHO-5	• VR group activity • WHO-5	• VR group activity • WHO-5	• VR group activity • WHO-5	• Postintervention interview • Mini-ICF-APP • ADL • WHO-5	• Follow-up-interview • Mini-ICF-APP • ADL • WHO-5
Implementation	• Psychologist	• Psychologist	• Psychologist and technical assistant	• Psychologist and technical assistant	• Psychologist and technical assistant	• Psychologist and technical assistant	• Psychologist	• Psychologist

[a]Some procedures such as a questionnaire for feedback and obtaining perceptions from the older adults are not presented in this table.

[b]Mini-ICF-APP: Mini-ICF-Rating for Impairment in Psychological Activities and Capacities [56].

[c]ADL: activity of daily living [57].

[d]WHO-5: World Health Organization–Five Well-Being Index [58].

The study was conducted during the COVID-19 pandemic, and the safety of our older adult participants was our top priority. From our initial contact with the nursing homes, we inquired about and strictly adhered to the COVID-19 guidelines governing group events. Our team members underwent COVID-19 testing at public testing institutes within 24 hours before each visit to a nursing home. Even if a team member exhibited symptoms similar to those of COVID-19 but obtained a negative test result, they were not allowed to enter the nursing home as an extra precaution. During individual interviews, all team members wore masks, and regular hand disinfection was practiced throughout their stay in the nursing homes. During the VR interventions, we maintained a safe distance between the older adults, and the VR equipment was thoroughly disinfected after each use to ensure the highest level of safety for all participants.

VR Intervention and Equipment

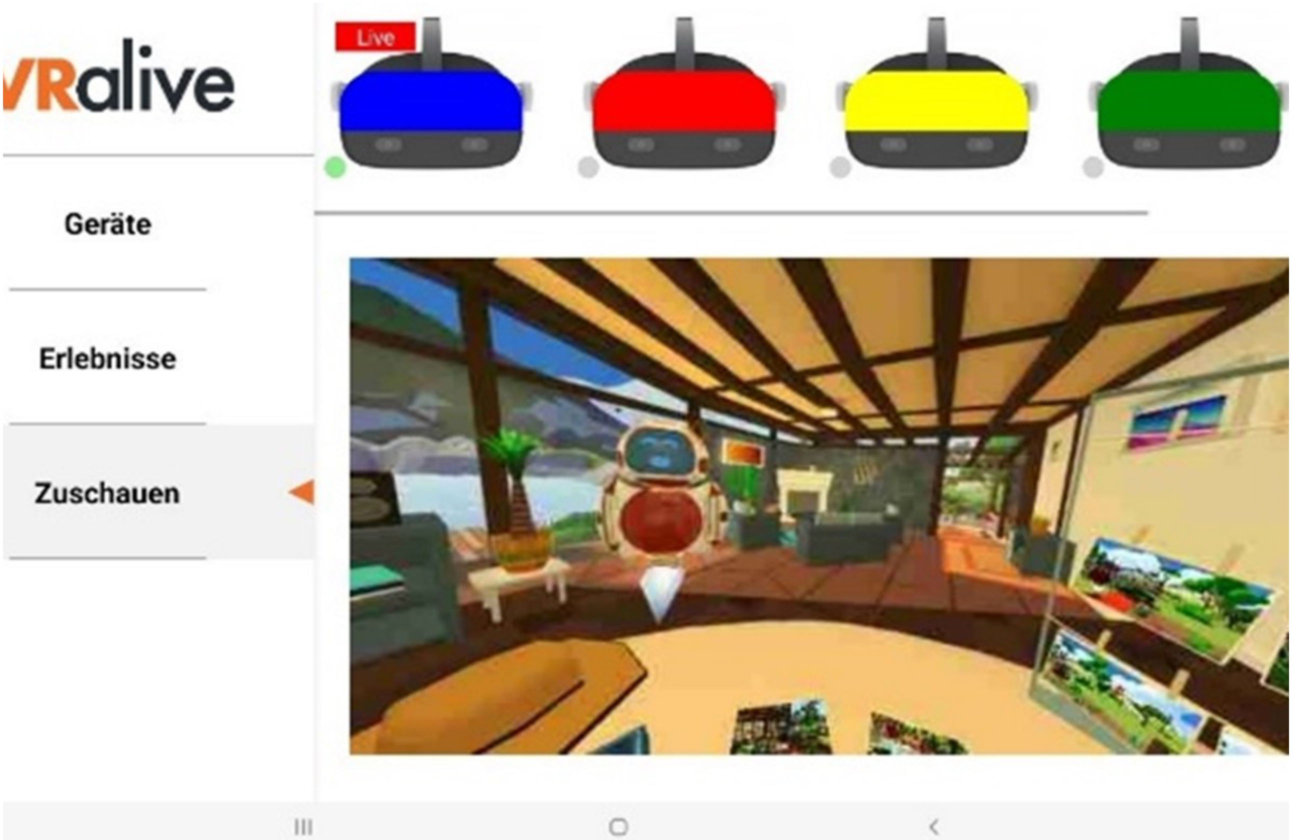
In collaboration with the technical company, VirtuaLounge, we developed a virtual vacation home to facilitate VR interventions for the older adults. We designed the meaningful VR activities in the virtual vacation home with an older adult–centered approach, drawing upon our understanding of older adults' daily routines and incorporating the valuable suggestions received from older adults during the pilot program testing phase while developing the VR intervention. The central aspect was to make the VR intervention as accessible as possible to the older adult population, while minimizing the barriers to use. For example, we ensured that older adults, including those using wheelchairs, could actively participate in the entire intervention by setting up the VR experience in a sitting format. We optimized the interaction with the VR environment to be easily manageable with only 1 finger, and all the tasks were designed to be completed using only 1 hand.

The interventions were conducted over a period of 4 weeks in regional nursing homes, with each session lasting approximately 30 minutes. Within each VR session, participants engaged in 4 or 5 tasks that were no longer possible for them to perform in real life. These tasks were integrated into a cohesive storyline, resulting in an immersive experience for the older adults. Our storyline revolved around 4 classic settings within a vacation home: living room, crafts station, garden, and kitchen (Figures 1 and 2). The tasks encompassed various routine activities, such as building furniture, gardening, and cooking in the kitchen. We devised a tablet control system specifically for those conducting the procedures to address the challenge of assisting older individuals wearing the nontransparent HMD. This system enables remote connectivity between the tablet and HMD. Through the tablet interface, a live view of the older adults' perspective is displayed, allowing the technical support personnel to monitor the progress of individual tasks for all participants. In addition, the tablet allows supporters to adjust the sound settings and remotely initiate or terminate the program on each HMD (Figure 3). This innovative solution enhances the ability to provide real-time assistance and control during the VR interventions.

Figure 1. Example task—gardening.

Figure 2. Example task—baking a pizza.

Figure 3. Screenshot of the tablet system.



The VR setup used in this study involved the use of the stand-alone VR Headset Pico Neo 3 Pro along with the Pico Neo 3 controller (Pico Technology Co, Ltd). The resolution of the VR headset was set at 1832 × 1920 per eye. The headset operated at a refresh rate of 72 Hz and featured 6 dfs inside-out tracking capabilities. In addition to the VR equipment, a Samsung Galaxy Tab S6 Lite (Samsung Electronics Co, Ltd) with an Android 12 operating system served as the remote tablet for the study. The system was programmed using .Net and C# programming languages.

Instruments

Consistent with our hypotheses, *psychosocial capacities* (hypothesis 1) included the mental capacities to adapt to daily life in a nursing home, which was assessed using the Mini-ICF-Rating for Impairment in Psychological Activities and Capacities (Mini-ICF-APP) scale [56]. This scale encompasses 13 psychosocial capacities, including proactivity and mobility. *ADLs* (hypothesis 1) included physical activity and mobility of the participants, which were assessed in more depth using the ADL-Barthel Index [57] that encompasses basic ADLs. *Well-being* (hypothesis 2) was measured using the World Health Organization–Five Well-Being Index (WHO-5) [58]. These measures were administered to both the intervention group and the control group, allowing for a comprehensive evaluation of the outcomes in both groups (Table 1).

The *Mini-ICF-APP capacity rating* [56] is an established instrument for describing a person's psychosocial capacity status. It has been translated internationally into many languages and is recommended in social medicine guidelines [59]. Among other settings, it is commonly used in settings of psychosocial rehabilitation. It is internationally recognized and has proven to be manageable, reliable, and valid in clinical practice [60-63]. In this study, the Mini-ICF-APP rating was designed to assess the psychosocial capacities for living in the nursing home, that is, performing certain basic activities on their own. The scale covers the following capacity dimensions: (1) adherence to regulations, (2) planning and structuring of tasks, (3) flexibility, (4) competence and knowledge application, (5) capacity to make decisions and judgments, (6) proactivity and spontaneous activities, (7) endurance, (8) self-assertiveness, (9) contact with others, (10) group integration, (11) intimate relationships, (12) self-care, and (13) mobility. Each dimension is rated using an 8-point rating scale (0=this is a strength of me to 7=this is impossible for me). The Mini-ICF-APP interview guide and questionnaire were adapted for the older adults by focusing on activities that individuals in need of care could still perform independently while residing in a nursing home. In a pilot study, 8 participants were interviewed by both project psychologists. Of the 2 psychologists, 1 conducted the interview, and both the interviewers completed Mini-ICF-PP rating sheet based on the responses provided by the older adults. On average, the interrater reliability over all 13 capacity dimensions was r=0.857. Psychosocial capacities were measured at 3 measurement time points (T0, T6, and T7).

ADL [57] was measured before the first VR intervention (T0), after the intervention (T6), and at a 3-week follow-up after the posttest phase (T7). It includes 10 dimensions of daily activity: (1) eating, (2) washing and showering, (3) body care, (4) dressing and undressing, (5) stool control, (6) urine control, (7) toilet use, (8) transfer from bed to chair, (9) movement and mobility, as well as (10) climbing stairs. The total score on the Activity of Daily Living-Barthel Index (ADL-BI) ranges from 0 to 100, with higher scores indicating greater independence in performing daily activities. A score of 0 indicates complete dependence on assistance for all activities, whereas a score of 100 indicates complete independence. An ADL score <80 indicates a need for care of >2 hours a day [64]. In the interview with the study participants, the psychologists asked about the activities one after the other and checked the plausibility of the answers of the older adults were. In cases of doubt, supplementary external judgments were obtained from the caregivers to ensure data validity.

The third instrument used in this study was the *WHO-5* [58], which is a concise self-report measure of current mental well-being. The assessment of well-being using the WHO-5 was conducted at baseline, after each VR intervention session, and during the postintervention and follow-up interviews. Numerous studies have demonstrated the validity of the WHO-5 as a screening tool for depressive mood and as a measure of treatment outcomes in clinical trials, and it has also shown good construct validity for assessing well-being in both younger and older populations [65,66]. The WHO-5 has been translated into >30 languages [66]. It consists of five statements that assess the individual's (1) good mood and cheerfulness, (2) relaxation, (3) activity and energy, (4) regenerative capacity through sleep, and (5) enthusiasm. Each statement is rated on a scale ranging from 1 ("at no time") to 5 ("all the time"). For example, 1 statement reads as follows: "Last week, I was happy and in a good mood." These statements are straightforward and nonintrusive in nature [66]. Typically, the questionnaire covers a 14-day period; however, considering the older adult participants and the study's weekly interventions, a 7-day period was deemed appropriate. Therefore, the assessment inquired about the individual's well-being, relaxation, activity level, quality of sleep, and interest in life over the past 7 days. Well-being was measured at each of the 8 measurement time points (T0-T7; Table 1).

Statistical Analyses

The data collected from the study were entered into the statistical software SPSS (IBM Corp) [67] for analysis. A repeated-measures ANOVA was conducted to analyze the data obtained from the 84 older adults in the intervention group who had participated in at least 3 interventions. The analysis focused on psychosocial capacities (Mini-ICF-APP), ADL, and well-being (WHO-5).

Owing to unequal sample sizes, a 2-factor, repeated-measures ANOVA between the intervention group and the control group could not be performed. However, the data from the control group (consisting of 11 participants) are presented descriptively, enabling a comparison with the values obtained from the intervention group.

Ethical Considerations

This study was funded by the German Federal Ministry of Education and Research (BMBF) (project number: 16SV8561 VRalive). This study was approved by the ethics committee of Technische Universität Braunschweig (FV-2020-18). Before the study, informed consent, confidentiality, and informed data protection were obtained from the participants or their life caregivers under the supervision of nursing staff. The older adults were informed of their ability to opt out at any time. The VR activities were provided as daily activities in the nursing home, and there was no compensation provided. All activities carried out in the nursing home were in strict compliance with the current nursing home COVID-19 prevention and treatment policy. For secondary analyses using existing data, we specified that the original consent approval covers secondary analysis without additional consent. The collected data were anonymized and deidentified.

Results

Participants and Demographic Information

A total of 116 older adults aged ≥60 years initially participated in the VR intervention group. Of these 116 participants, 31 (26.7%) discontinued their involvement in the VR group. The primary reasons cited for dropping out were concerns related to data protection and illness or death (Table 2).

Table 2. Reasons for dropping out (n=31).

Reasons	Values, n (%)
No more interest, without reasons	5 (16)
Several events are occurring	4 (13)
Incapable of participating due to illness or death	12 (39)
Virtual reality–related sickness (“cybersickness”)	5 (16)
No interest in interviews	1 (3)
Alternative events are preferred	2 (6)
Leaves retirement home	1 (3)
Cancellation of the group owing to very few participants	1 (3)

Sociodemographic information was collected at baseline and included older adults’ age, sex (male, female, or intersex), and educational and working history. The intervention group (N=116) had an average age of 80.74 (SD 8.49) years. The age range at the time of the intervention varied from 60 to 97 years. There was a higher proportion of female participants (81/116, 69.8%). Most participants (107/116, 92.2%) in the intervention group had no previous experience with VR. Table 3 presents more detailed demographic information.

We decided to consider only the data from older adults who participated in a minimum of 3 VR interventions to ensure that the analysis focused on the impact of the VR intervention. This resulted in a final sample size of 84 older adults to be analyzed statistically. Hence, for the purpose of this paper, statistical evaluation will be conducted on the data obtained from 84 older adults from the VR intervention group.

Furthermore, 12 older adults participated in the control group. One of the participants discontinued due to death (this has already been accounted for in the dropout statistics, as indicated in Table 2). The age range of the control group participants spanned from 61 to 94 years, with an average age of 83.75 (SD 8.97) years, and 10 (83%) of the 12 participants were women. It is important to note that the selection of participants for the control group was based on the perceptions of the nursing staff members and the natural decisions made by the older adults themselves. Consequently, the sample size of the control group in this study is notably small, rendering it insufficient for a robust comparison with the intervention group. Therefore, detailed information about the control group, which is presented alongside the intervention group data, is provided in Multimedia Appendices 1-3. The limitations associated with the small control group sample size are discussed further in the Strengths and Limitations section.

Table 3. Sociodemographic data about the older adults participating in the virtual reality (VR) intervention (N=116).

Features	Values
Age (y), mean (SD; range)	80.74 (8.49; 60-97)
Sex, n (%)	
Female	81 (69.8)
Male	35 (30.2)
Education, n (%)	
None	7 (6)
Special school	2 (1.7)
Primary school	1 (0.9)
Secondary school	17 (14.7)
Primary school or grade 9 or 10	79 (68.1)
Abitur	10 (8.6)
Professional qualification, n (%)	
None	37 (31.9)
Craft profession or skilled work	67 (57.8)
Master	7 (6)
University studies	5 (4.3)
Longest professional activity in working life, n (%)	
Craft, industry, or production	37 (31.9)
Research and development	3 (2.6)
Agriculture	3 (2.6)
Office or management activities	24 (20.7)
Service, gastronomy, or customer service	26 (22.4)
Practical health care (nurse, physician, therapist, or similar)	10 (8.6)
Housewife	12 (10.3)
Missing indication	1 (0.9)
Frequency of visits from trusted people, n (%)	
Several times a week	63 (54.3)
Weekly	27 (23.3)
Every 2-3 weeks	8 (6.9)
Monthly	2 (1.7)
Less frequently than monthly	2 (1.7)
No regular contacts	14 (12.1)
Previous experience with VR, n (%)	
No	107 (92.2)
Yes	9 (7.8)

Outcomes

Tables 4 and 5 present a summary of the results for the Mini-ICF-APP, ADL, and WHO-5 measures at specific measurement time points for both the intervention group and the control group. A macrolevel analysis indicates significant differences in the mean scores of Mini-ICF-APP ($P<.001$; $\eta^2=0.150$) and WHO-5 ($P=.04$; $\eta^2=0.032$) and the sum score of ADL ($P=.02$; $\eta^2=0.050$) within the intervention group.

Table 4. Comparison of the older adults' scores at baseline, at the end of the intervention, and 3 weeks after the postintervention assessment.

	Baseline (T0), mean (SD)	Postintervention assessment (T6), mean (SD)	Follow-up (T7), mean (SD)	rANOVA[a] (n=84)	
				P value	η²
Capacities (Mini-ICF-APP[b])					
Adjustment to rules and routines	2.57 (0.88)	2.15 (1.05)	2.04 (1.08)	<.001[c]	0.122
Planning and structuring tasks	3.13 (1.82)	2.95 (1.92)	3.13 (2.08)	.56	0.007
Flexibility and adaptability	2.38 (0.90)	1.86 (1.19)	1.87 (1.22)	<.001[c]	0.109
Competence and knowledge application	2.25 (1.25)	1.99 (1.55)	1.94 (1.52)	.04[d]	0.039[e]
Capacity to make decisions and judgments	2.57 (1.12)	2.42 (1.40)	2.48 (1.35)	.57	0.007
Proactivity and spontaneous activities	2.39 (1.19)	2.04 (1.25)	1.90 (1.26)	<.001[c]	0.104
Resilience and perseverance	2.54 (1)	2.25 (1.18)	2.29 (1.14)	.07	0.032
Self-assertiveness	2.60 (1.09)	2.43 (1.15)	2.40 (0.96)	.30	0.014
Capacity to talk with and contact third parties	2.39 (1.41)	2.14 (1.35)	2.25 (1.42)	.15	0.023[e]
Group integration	2.71 (1.39)	2.26 (1.36)	2.04 (1.21)	<.001[c]	0.141
Capacity to form close relationships	2.61 (1.58)	2.45 (1.66)	2.49 (1.75)	.57	0.007
Self-care and self-sufficiency	3.29 (1.76)	3.18 (1.84)	3.11 (1.86)	.63	0.006
Mobility and transportability	2.39 (1.46)	2.37 (1.59)	2.40 (1.54)	.95	0.001
Average score	2.60 (0.75)	2.35 (0.85)	2.33 (0.88)	<.001[c]	0.150[e]
ADL[f]					
Food	9.29 (1.76)	9.46 (1.56)	9.64 (1.30)	.11	0.027[e]
Bath	1.85 (2.43)	1.73 (2.39)	2.02 (2.47)	.42	0.010
Washing	4.52 (1.48)	4.58 (1.39)	4.64 (1.30)	.76	0.003
Dressing and undressing	7.20 (3.41)	7.38 (3.68)	7.32 (3.51)	.84	0.002
Stool control	7.92 (3.74)	8.04 (3.72)	7.80 (3.75)	.83	0.002
Urine control	6.31 (4.40)	6.55 (4.25)	6.90 (4.31)	.27	0.016[e]
Using the toilet	8.27 (3.76)	8.63 (3.41)	9.11 (2.71)	.03[d]	0.040
Bed or wheelchair transfer	12.74 (4.93)	13.15 (4.58)	12.98 (4.66)	.22	0.018[e]
Movement or mobility	9.11 (4.40)	9.46 (4.39)	9.64 (4.16)	.04[d]	0.039
Climbing stairs	4.35 (4.40)	4.76 (4.17)	4.88 (4.46)	.28	0.015
Total score	71.55 (23.19)	73.75 (23.32)	74.94 (22.29)	.02[d]	0.050[e]

[a]rANOVA: repeated ANOVA.

[b]Mini-ICF-APP: Mini-ICF-Rating for Impairment in Psychological Activities and Capacities; the scale ranges from 0 ("clearly a strength of mine") to 7 ("I cannot do at all").

[c]*P*<.001.

[d]*P*<.05.

[e]The value was corrected according to Greenhouse Geisser.

[f]ADL: activity of daily living; the scale ranges from 0 to 15.

Table 5. Comparison of older adults' scores regarding their well-being (World Health Organization–Five Well-Being Index [WHO-5][a]) before the intervention (T0 and T1), during the intervention (T2-T5), after the intervention (T6), and 3 weeks after the postintervention assessment (T7).

	Time points, mean (SD)								rANOVA[b]		
	T0	T1	T2	T3	T4	T5	T6	T7	*F* test (*df*)	*P* value	η²
Good mood and cheerfulness	4.08 (0.95)	3.90 (1.04)	3.78 (1.13)	3.94 (1.02)	3.90 (0.99)	3.85 (1.03)	3.83 (0.98)	3.90 (1.04)	1.893 (5.51)	.09	__[c]
Relaxation	4.02 (0.01)	3.89 (1.22)	3.98 (1.09)	3.90 (0.96)	3.94 (1.04)	3.93 (1.02)	4.06 (1.09)	3.89 (1.22)	0.554 (7.00)	.79	0.008
Activity and energy	3.54 (1.27)	3.27 (1.25)	3.32 (1.33)	3.52 (1.21)	3.57 (1.18)	3.23 (1.26)	3.60 (1.11)	3.27 (1.25)	2.050 (7.00)	.048[d]	0.029
Regenerative capacity through sleep	3.49 (1.35)	3.65 (1.36)	3.71 (1.26)	4.03 (1.05)	4.03 (1.21)	3.89 (1.36)	4.01 (1.12)	3.65 (1.36)	3.867 (5.85)	.001[e]	0.054[c]
Enthusiasm	3.64 (1.17)	3.72 (1.28)	3.60 (1.17)	3.73 (1.22)	3.85 (1.22)	3.68 (1.29)	4.12 (1.12)	3.72 (1.28)	3.541 (7.00)	.001[e]	0.049
WHO-5 (total)	3.75 (0.75)	3.69 (0.83)	3.68 (0.83)	3.82 (0.79)	3.86 (0.82)	3.72 (0.84)	3.92 (0.78)	3.91 (0.87)	2.235 (5.85)	.04[d]	0.032[c]

[a]A 5-point Likert scale ranging from 0 ("at no time") to 5 ("all the time").

[b]rANOVA: repeated ANOVA.

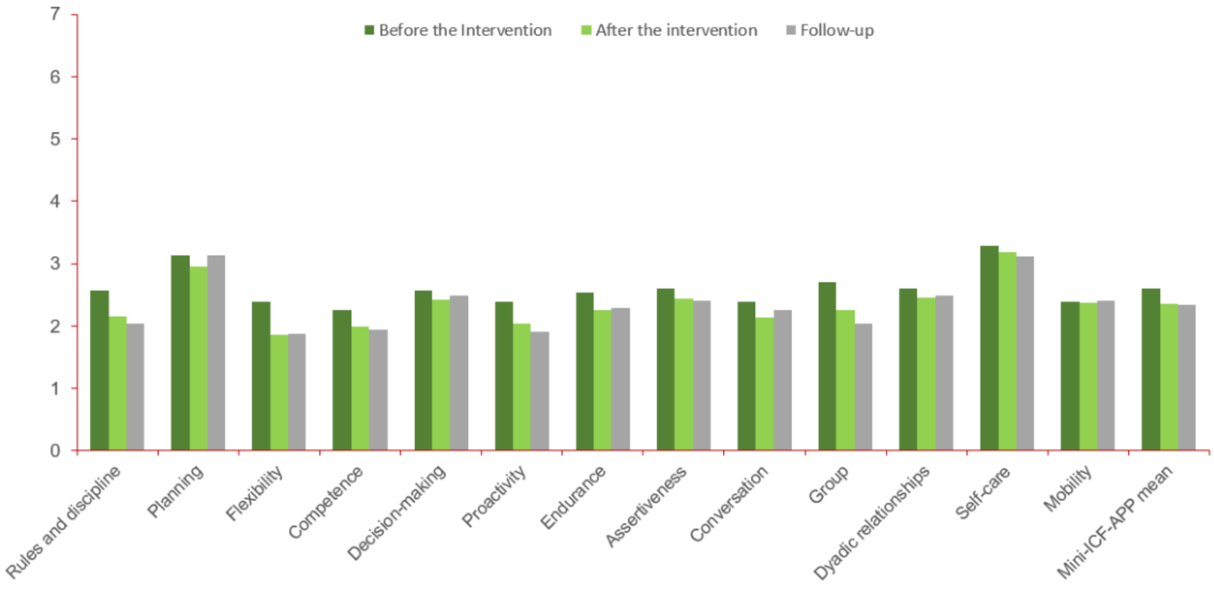
[c]The value was corrected according to Greenhouse Geisser.

[d]*P*<.05.

[e]*P*<.01.

Specifically, the data from the Mini-ICF-APP indicated slight reductions in some psychosocial capacity impairments within the intervention group (Table 4 and Figure 4): adherence to regulations (*P*<.001; η²=0.122), flexibility (*P*<.001; η²=0.109), proactivity (*P*<.001; η²=0.104), and group integration ICF (*P*<.001; η²=0.141). Problems related to competence also showed a slight decrease (*P*=.04; η²=0.039).

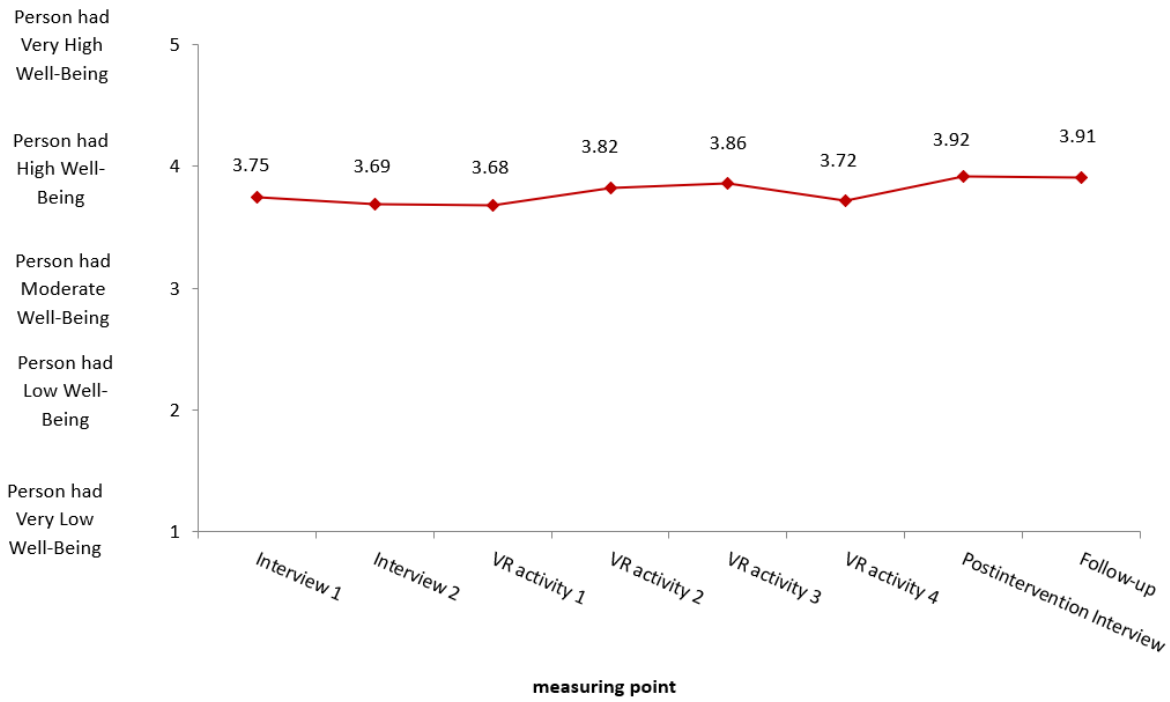
Figure 4. Changes in older adults' scores for the ability dimensions (Mini-ICF-Rating for Impairment in Psychological Activities and Capacities; Mini-ICF-APP) across the measurement time points: before the intervention, after the intervention, and follow-up during the intervention.. 0=This is clearly a strength of person, 1=Person is better than many others, 2=Person can do this well, 3=Person can somehow work with this, 4=Person does not always get this to work, 5=Person has problem with this, 6=Person needs help in this regard, and 7=Person is fully unfit.



Overall, 2 ADLs seemed to improve over the course of the VR intervention (Table 4): "Using the toilet" (*P*=.03; η²=0.040) and "mobility" (*P*=.04; η²=0.039).

In terms of well-being (WHO-5), 3 of 5 items in the WHO-5 showed significant changes (Table 5 and Figure 5): "feeling active" (*P*=.048; η²=0.029), "sleeping well" (*P*<.001; η²=0.054), and "being full of interest for life" (*P*<.001; η²=0.049). This indicates a slight variation in well-being over the course of the 4-week VR intervention period.

Figure 5. Mean of the World Health Organization–Five Well-Being Index for the intervention group at 8 weekly time points. VR: virtual reality.



Side Effects of VR

Throughout the VR intervention, there were no reported instances of older adults falling or accidentally colliding with nearby objects based on their sitting poses. In addition, there was no indication of potential interference with medical devices. While a few cases of cybersickness (5/31, 16%; Table 2) were reported, these were promptly addressed and resolved.

Discussion

Summary

The study aimed to determine whether a specifically designed VR intervention had a positive impact on nursing home inhabitants in terms of psychosocial capacities, activities, and well-being. In general, some dimensions of psychosocial capacities, activity, mobility, and well-being of older adults in the intervention group showed a slight positive improvement over the course of the intervention. Improvements were observed in adherence to regulations, flexibility, proactivity, competence, and group integration. The older adults showed improvement in their ability to use the toilet and physical mobility. The older adults reported feeling active, experiencing better sleep, and displaying greater interest in daily life.

Principal Findings and Comparison to Previous Studies

Psychosocial Capacities

The results of the study demonstrate an improvement substantially in the older adults' capacity to adhere to daily routines in the nursing home. This indicates that they have become more proficient in following schedules and maintaining self-discipline. In addition, the older adults show a greater willingness to adapt to and switch between different tasks or activities, both within the group setting and in public. Studies have shown that VR interventions can enhance the cognitive abilities of older adults, and these improvements are often correlated with their intrinsic motivation for training. This motivation can be triggered either by the engaging nature of the game itself or by the immersive experience facilitated by VR technology [68,69]. In addition, the study conducted by Fan et al [46] demonstrated that using VR as a form of entertainment for older adults can enhance their achievement motive and self-esteem, leading to improved mental health outcomes and reduced isolation, particularly among community-dwelling older adults. It is reasonable to assume that similar benefits could be experienced by older adults residing in nursing homes as well. In our case, the older adults' intrinsic motivation may stem from their own desire to participate in meaningful VR activities and engage with others, whereas extrinsic motivation could arise from the supportive environment provided by the staff members and the positive experiences associated with the VR group activity. The combination of intrinsic and extrinsic motivation appears to have contributed to the observed improvements in discipline and flexibility among the older adults; with these improvements, the older adult could fit better with daily life in the nursing home. However, it should be acknowledged that flexibility may also have been influenced by factors such as the visits from and relationship with the VR project psychologist. Over time, as the psychologist became more familiar with the older adults, they could have corrected any initial bias in the older adults' self-assessment, leading to more accurate evaluations based on their pragmatic experience.

The study findings indicate that the older adults' use and retention of competence and knowledge have improved. This outcome aligns with the study's design, which focused on implementing VR interventions with meaningful activities based on the older adults' familiar daily tasks. The positive impact of this approach is evident in the results obtained. The older adults were able to engage with the new VR activities, thus increasing their competence. Moreover, the result of improved capability of group integration indicated that after the VR intervention, the older adults gradually developed an interest in the group activity and expressed a desire to retry the task in the following week. A previous study by Padilha et al [70] found that VR offers a learning experience from interaction with the virtual environment and enhances knowledge acquisition in nursing education. It is plausible that VR can also be a promising tool for older adults to enhance these mental capabilities. Multiple studies have demonstrated that VR interventions have a positive impact on the memory and information-processing abilities of older adults [68,71,72]. VR video games could enhance the working memory and reasoning abilities of older adults [68]. The authors also suggested that VR interventions may improve problem-solving and planning skills, which are findings that align with those of our own study.

In summary, the VR group activity incorporating daily life tasks has shown to be a promising method for improving psychosocial capacities, including adherence to regulations, flexibility, and competence retention, among older adults in nursing homes.

Activity and Mobility

The results of the study demonstrate significant enhancement in the older adults' proactivity (Mini-ICF-APP). This suggests that the older adults experience less boredom and express a desire to engage in more daily activities that interest them while living in the nursing home. They are better prepared and motivated to initiate activities on their own. Furthermore, the ADL score, particularly in the mobility domain, also showed a significant improvement. This improvement in proactivity can be linked to the enhanced mobility observed in the ADL scale. The nursing home already provides gymnastic courses and physical therapy to help older adults maintain or rebuild their physical functioning. When older adults are more willing to participate in these activities, it can lead to better mobility and daily activity (eg, using the toilet) outcomes, aligning with the theoretical background that meaningful group activities can improve activity and mobility in nursing homes. Previous studies have investigated the impact of VR interventions on ADLs and instrumental ADLs among older adults, yielding different results. One of these studies aligns closely with our own study. Liao et al [47] conducted a VR cognitive training program that involved tasks such as locating stores and acting as a virtual kitchen chef, which is similar to our approach. Their results demonstrated a significant improvement in older adults' activity levels, with the effect size being larger than that observed with traditional cognitive training methods. Moreover, the improvement in the activity of older adults in nursing homes through VR intervention was also demonstrated by Saredakis et al [73]. In contrast, Optale et al [43] conducted a VR memory training program that consisted of repeated memory tasks focused on objects and orientation. Their study did not find a positive impact on ADL. This discrepancy suggests that the content of the VR intervention may play a crucial role in determining its effectiveness in enhancing older adults' daily activities. Overall, these findings highlight the importance of considering the specific content and nature of VR interventions when assessing their potential impact on the daily lives of older adults.

On the basis of the exploratory findings of this study, conducting an experimental study to provide evidence about the impact of VR interventions on proactivity and mobility would be valuable. Such a study can further validate the potential benefits of VR interventions in promoting proactivity and improving mobility among older adults in nursing homes.

Group Integration and Social Interaction

There was significant improvement in the group capacity of the older adults, indicating their increased willingness to participate in group activities and enjoy the benefits of group engagement. Thus, VR intervention could be used as a meaningful group activity that contributes to reducing social isolation. Staff members in some care homes also reported that older adults were pleasantly surprised by the VR technology and were more open to group activities after participating in the VR project. These findings are consistent with the results reported by Fan et al [46], who conducted a VR intervention involving horticultural group activities such as gardening for community-dwelling older adults, aiming to reduce social isolation. In addition, other previous studies involving VR horticultural activities in nursing homes have also demonstrated a reduction in older adults' loneliness and an improvement in their social interaction [44,74]. However, it is important to note that, unlike our study, these previous studies did not specifically focus on VR interventions as group events. Therefore, these studies have not reported about the impact of VR interventions specifically on group capacities of older adults. In addition to the VR horticultural activities implemented in nursing homes, the study by Saredakis et al [73] examined the effectiveness of VR reminiscence therapy in reducing older adults' loneliness but did not observe significant effects. This suggests that VR horticultural activities may hold greater potential in reducing loneliness among older adults. Engagement in meaningful activities, such as virtual gardening or horticultural group activities, might have a more profound impact on addressing the issue of loneliness in this population.

It is important to note that the project spanned periods of the COVID-19 pandemic and the winter and summer seasons, which could potentially act as confounding factors. While there was generally limited availability of group activities during the pandemic, the introduction of VR group sessions may have enhanced older adults' interest in social interaction.

Well-Being

Although there are statistically significant differences in the changes in well-being over the course of the intervention, it is important to note that the VR group activity has not yet demonstrated its full potential in consistently improving the well-being of older adults. The results indicate that there have been very small, incremental improvements in the well-being curve. The statistically significant difference observed may be attributed to the number of measurement points used in the study. We do not have a sufficient, practically relevant effect to confirm an increase in well-being.

The well-being of the older adults was already good at the beginning of the VR intervention study, which makes significant and consistent additional improvements less likely. Furthermore, well-being is influenced by various situational factors, particularly among older adults with health problems or disabilities considered vulnerable. Another factor to consider is the frequency and duration of the VR intervention, which may not have been sufficient to produce further improvement in well-being. Other VR studies that have demonstrated improvements in well-being often involve more frequent and longer VR interventions [44,50,74] or are only measured once after a 1-time intervention [48,75,76]. Furthermore, recent entertainment-oriented VR interventions targeting well-being or quality of life among older adults have predominantly used a passive interaction approach, for instance, virtual travel in Hong Kong [75]. These studies consistently achieved their research goals in terms of enhancing older adults' mood and well-being. In contrast, VR interventions that primarily focus on functional training with hand interaction have generally shown limited improvement in overall well-being, such as the one conducted by Brito et al [77]. The learning process associated with using the hand console can act as a barrier for older adults, potentially hindering their ability to improve their well-being through functional training. Consequently, it is crucial to approach the didactic process of VR devices with care to ensure that older adults are not discouraged at the initial stages of training. Furthermore, when developing VR interventions for older adults, it is important to consider the design and usability of the console or device being used. Older adults may have specific needs and challenges when it comes to interacting with technology. Therefore, the console or device should be tailored to accommodate their physical abilities, cognitive capabilities, and potential sensory impairments [78].

In summary, the well-being of older adults could be maintained at a high level over the course of the VR group intervention. It would be interesting to see if a more frequent intervention could further improve the impact of VR intervention on the well-being of older adults.

Strengths and Limitations

This study explored VR group activities in nursing homes, adopting a naturalistic approach to gain a deeper understanding of technology's role for older adults in the digital age. Our findings revealed the potential of VR as a tool in meaningful activity programs for older adults residing in nursing homes. Notably, this intervention leads to an enhancement in older adults' abilities and engagement in activities, while sustaining a high level of well-being. Our study offers novel insights into the transformative possibilities of VR for enriching the lives of older residents within nursing home settings. Despite the significant findings of this study, it is important to acknowledge several limitations.

First, the impact of the COVID-19 pandemic on the effectiveness of VR interventions cannot be ignored. During the pandemic, there were restricted group activities and increased vulnerability among older adults, which may have magnified the positive impacts of the VR group intervention. It is crucial to consider this unique context when interpreting the results.

Second, there may be a selection bias in the sample of participants. The selection of participants in the nursing home was based on defined and standardized selection criteria, which the nurses applied in the field. This is the most natural, accepted, and standardized way to select participants for psychosocial activities in nursing homes. There can be a slight selection bias due to the various individual interaction processes of the nurses with the participants. The selected sample may not accurately represent the range of responses in the population, but it represents older adults with complex disabilities who are nevertheless able to cope with specially designed VR tools. Moreover, recruitment was also based on older adults' willingness to participate in the study. The older adults who chose to participate in the study may be more open to new experiences compared to those who declined. Furthermore, the findings indicate that the older adults initially reported good well-being and had regular contact with family or friends, suggesting a limited scope for improvement in well-being and social interaction. It is essential to find ways to extend the reach of VR group interventions to a wider range of older adults, particularly those who are more isolated and lonely. Using a VR session to introduce the intervention to these older adults may be a potential solution.

Third, there is a possibility that the older adults may have overestimated or underestimated their own capacities. This could be addressed by staff members closely monitoring the older adults' daily behavior. However, due to limited personnel resources, this was not feasible in this study.

Finally, another weakness of this study is in its study design. This was not a randomized controlled trial. Although a control group was included, the sample size was very small, making it challenging to establish a valid comparison with the intervention group. Participation in the control group was based on the natural decisions of the older adults. Therefore, the results should be interpreted as a point of reference rather than indicative of a causal "effect." Nevertheless, it is important to note that this longitudinal study is naturalistic and externally valid. It offers a novel perspective on the pragmatic application of VR intervention as a group event in nursing homes.

Future Studies

On the basis of the findings of this exploratory study, a randomized controlled experimental trial that specifically focuses on VR group interventions within the daily lives of older adults in nursing homes should be conducted. Without the specific conditions during the COVID-19 pandemic, a more favorable social environment will be available, resulting in fewer hindering factors such as limited group interventions. In addition, in this study, large variation was observed in the basic cognitive functions of the participating older adults, according to age and type of disease. However, it is important for researchers to be mindful about the competencies and skill levels of older adults when introducing VR interventions [79]. Some older adults may feel socially excluded if they lack the necessary skills to participate in these digital activities [80]. Therefore, it is crucial to prioritize accessibility and provide adequate support and training to ensure inclusivity. This can involve tailoring the VR experiences to accommodate older adults with varying cognitive functioning, such as providing different levels of difficulty based on individual capabilities. In our study, we received diverse feedback from older adults regarding VR tasks. The highly independent older adults expressed that the VR tasks were very easy for them, whereas those with cognitive impairments or dementia found the tasks challenging to complete. In the next phase, it would be beneficial to group older adults based on their cognitive capacities and provide tailored VR interventions at different difficulty levels. Exploring an older adult–centered VR design is another intriguing direction for further investigation. This could involve studying the optimal form of interaction that minimizes the learning curve associated with using VR devices, ultimately enhancing the overall user experience for older adults.

Conclusions

In conclusion, the project successfully explored the benefits of a VR-based group intervention in nursing home settings. The results indicate that the VR intervention could be a meaningful group activity in nursing homes to support social group interaction, activity level, and well-being. The 4 sessions of the VR group intervention—with tasks that the older adults were unable to perform in their current environment— led to significant improvements in adherence to rules, flexibility, competence, proactivity, group integration, and mobility. Future research could benefit from conducting a randomized controlled trial to provide stronger evidence.

Acknowledgments

Funding for this project was provided by the German Federal Ministry of Education and Research (project number V5KMU1910044-02 EID 0017003). This study would not have been possible without the technical support of the VirtuaLounge team.

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

BM and IS conceptualized the study and acquired the funding. BM designed the research question. IS and her team developed the virtual reality software. YL and CW conducted and implemented the study with technical support from IS in nursing homes. YL and CW collected and analyzed the data and prepared the tables for this paper. YL wrote the paper. BM supervised the research process and contributed to the writing and revision of the paper.

Conflicts of Interest

IS is the cofounder of the company, VirtuaLounge, which developed the virtual reality program used in this study. The developed program may be used commercially in the future.

Multimedia Appendix 1
Sociodemographic data of the participating older adults in the virtual reality intervention for the control group and the intervention group.
[DOCX File , 35 KB - games_v12i1e50796_app1.docx]

Multimedia Appendix 2
Comparison of the scores of the older adults from the control group and the intervention group at baseline, at the end of the virtual reality intervention, and 3 weeks after the postintervention assessment regarding their psychosocial capacities.
[DOCX File , 20 KB - games_v12i1e50796_app2.docx]

Multimedia Appendix 3
Comparison of the scores of the older adults from the control group and the intervention group regarding their well-being before the intervention (T0 and T1), during the intervention (T2-T5), after the intervention (T6), and 3 weeks after the postintervention assessment (T7).
[DOCX File , 18 KB - games_v12i1e50796_app3.docx]

References

1. Österlind J, Ternestedt BM, Hansebo G, Hellström I. Feeling lonely in an unfamiliar place: older people's experiences of life close to death in a nursing home. Int J Older People Nurs 2017 Mar 14;12(1):e12129. [doi: 10.1111/opn.12129] [Medline: 27624362]

2. Vossius C, Selbæk G, Šaltytė Benth J, Bergh S. Mortality in nursing home residents: a longitudinal study over three years. PLoS One 2018 Sep 18;13(9):e0203480 [FREE Full text] [doi: 10.1371/journal.pone.0203480] [Medline: 30226850]
3. Braggion M, Pellizzari M, Basso C, Girardi P, Zabeo V, Lamattina MR, et al. Overall mortality and causes of death in newly admitted nursing home residents. Aging Clin Exp Res 2020 Feb 01;32(2):275-280. [doi: 10.1007/s40520-019-01441-x] [Medline: 31894560]
4. Goldberg TH, Botero A. Causes of death in elderly nursing home residents. J Am Med Dir Assoc 2008 Oct;9(8):565-567. [doi: 10.1016/j.jamda.2008.04.011] [Medline: 19083290]
5. den Ouden M, Bleijlevens MH, Meijers JM, Zwakhalen SM, Braun SM, Tan FE, et al. Daily (In)activities of nursing home residents in their wards: an observation study. J Am Med Dir Assoc 2015 Nov 01;16(11):963-968. [doi: 10.1016/j.jamda.2015.05.016] [Medline: 26155723]
6. Harvey JA, Chastin SF, Skelton DA. Prevalence of sedentary behavior in older adults: a systematic review. Int J Environ Res Public Health 2013 Dec 02;10(12):6645-6661 [FREE Full text] [doi: 10.3390/ijerph10126645] [Medline: 24317382]
7. Maher JP, Conroy DE. Daily life satisfaction in older adults as a function of (In)activity. J Gerontol B Psychol Sci Soc Sci 2017 Jul 01;72(4):593-602. [doi: 10.1093/geronb/gbv086] [Medline: 26405078]
8. Dahlberg L, McKee KJ, Frank A, Naseer M. A systematic review of longitudinal risk factors for loneliness in older adults. Aging Ment Health 2022 Feb;26(2):225-249. [doi: 10.1080/13607863.2021.1876638] [Medline: 33563024]
9. Pinquart M, Sorensen S. Influences on loneliness in older adults: a meta-analysis. Basic Appl Soc Psych 2010 Jun 07;23(4):245-266. [doi: 10.1207/s15324834basp2304_2]
10. Gardiner C, Laud P, Heaton T, Gott M. What is the prevalence of loneliness amongst older people living in residential and nursing care homes? A systematic review and meta-analysis. Age Ageing 2020 Aug 24;49(5):748-757 [FREE Full text] [doi: 10.1093/ageing/afaa049] [Medline: 32396600]
11. Victor CR. Loneliness in care homes: a neglected area of research? Aging Health 2012 Dec;8(6):637-646. [doi: 10.2217/ahe.12.65]
12. Wu B. Social isolation and loneliness among older adults in the context of COVID-19: a global challenge. Glob Health Res Policy 2020 Jun 5;5(1):27 [FREE Full text] [doi: 10.1186/s41256-020-00154-3] [Medline: 32514427]
13. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
14. Pitkala KH. Loneliness in nursing homes. J Am Med Dir Assoc 2016 Aug 01;17(8):680-681. [doi: 10.1016/j.jamda.2016.04.007] [Medline: 27217096]
15. Brownie S, Horstmanshof L. The management of loneliness in aged care residents: an important therapeutic target for gerontological nursing. Geriatr Nurs 2011 Sep;32(5):318-325. [doi: 10.1016/j.gerinurse.2011.05.003] [Medline: 21831481]
16. Santini ZI, Jose PE, York Cornwell E, Koyanagi A, Nielsen L, Hinrichsen C, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. Lancet Public Health 2020 Jan;5(1):e62-e70 [FREE Full text] [doi: 10.1016/S2468-2667(19)30230-0] [Medline: 31910981]
17. DiNapoli EA, Wu B, Scogin F. Social isolation and cognitive function in Appalachian older adults. Res Aging 2014 Mar 04;36(2):161-179. [doi: 10.1177/0164027512470704] [Medline: 25650688]
18. Johansson-Pajala RM, Alam M, Gusdal A, Heideken Wågert PV, Löwenmark A, Boström AM, et al. Anxiety and loneliness among older people living in residential care facilities or receiving home care services in Sweden during the COVID-19 pandemic: a national cross-sectional study. BMC Geriatr 2022 Dec 01;22(1):927 [FREE Full text] [doi: 10.1186/s12877-022-03544-z] [Medline: 36456904]
19. Mushtaq R, Shoib S, Shah T, Mushtaq S. Relationship between loneliness, psychiatric disorders and physical health ? A review on the psychological aspects of loneliness. J Clin Diagn Res 2014 Sep;8(9):WE01-WE04 [FREE Full text] [doi: 10.7860/JCDR/2014/10077.4828] [Medline: 25386507]
20. Zhao M, Gao J, Li M, Wang K. Relationship between loneliness and frailty among older adults in nursing homes: the mediating role of activity engagement. J Am Med Dir Assoc 2019 Jun;20(6):759-764. [doi: 10.1016/j.jamda.2018.11.007] [Medline: 30630725]
21. Horowitz BP, Vanner E. Relationships among active engagement in life activities and quality of life for assisted-living residents. J Hous Elderly 2010 May 25;24(2):130-150. [doi: 10.1080/02763891003757056]
22. Park NS. The relationship of social engagement to psychological well-being of older adults in assisted living facilities. J Appl Gerontol 2009 Feb 03;28(4):461-481. [doi: 10.1177/0733464808328606]
23. Grenade L, Boldy D. Social isolation and loneliness among older people: issues and future challenges in community and residential settings. Aust Health Rev 2008 Aug;32(3):468-478. [doi: 10.1071/ah080468] [Medline: 18666874]
24. Inoue Y, Wann DL, Lock D, Sato M, Moore C, Funk DC. Enhancing older adults' sense of belonging and subjective well-being through sport game attendance, team identification, and emotional support. J Aging Health 2020 Mar 15;32(7-8):530-542. [doi: 10.1177/0898264319835654] [Medline: 30873892]
25. Tak SH, Kedia S, Tongumpun TM, Hong SH. Activity engagement: perspectives from nursing home residents with dementia. Educ Gerontol 2015 Mar;41(3):182-192 [FREE Full text] [doi: 10.1080/03601277.2014.937217] [Medline: 25489122]

26. Morley JE, Philpot CD, Gill D, Berg-Weger M. Meaningful activities in the nursing home. J Am Med Dir Assoc 2014 Feb;15(2):79-81. [doi: 10.1016/j.jamda.2013.11.022] [Medline: 24461236]
27. Groenendaal M, Loor A, Trouw M, Achterberg WP, Caljouw MA. Perspectives of healthcare professionals on meaningful activities for persons with dementia in transition from home to a nursing home: an explorative study. Healthcare (Basel) 2019 Aug 19;7(3):98 [FREE Full text] [doi: 10.3390/healthcare7030098] [Medline: 31430855]
28. Travers C, Brooks D, Hines S, O'Reilly M, McMaster M, He W, et al. Effectiveness of meaningful occupation interventions for people living with dementia in residential aged care: a systematic review. JBI Database System Rev Implement Rep 2016 Dec;14(12):163-225. [doi: 10.11124/JBISRIR-2016-003230] [Medline: 28009676]
29. Tierney L, Beattie E. Enjoyable, engaging and individualised: a concept analysis of meaningful activity for older adults with dementia. Int J Older People Nurs 2020 Jun 05;15(2):e12306. [doi: 10.1111/opn.12306] [Medline: 32022414]
30. Morley JE, Kusmaul N, Berg-Weger M. Meaningful engagement in the nursing home. J Gerontol Soc Work 2021 Dec 22;64(1):33-42. [doi: 10.1080/01634372.2020.1864543] [Medline: 33353488]
31. Arbeitsmarktsituation im pflegebereich. Bundesagentur für Arbeit. URL: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Statistiken/Themen-im-Fokus/Berufe/Generische-Publikationen/Altenpflege.pdf?__blob=publicationFile [accessed 2023-06-28]
32. Ericson-Lidman E, Strandberg G. Learning to deal constructively with troubled conscience related to care providers' perceptions of not providing sufficient activities for residents. Clin Nurs Res 2015 Apr 09;24(2):211-227. [doi: 10.1177/1054773813500139] [Medline: 24021209]
33. Benjamin K, Edwards N, Ploeg J, Legault F. Barriers to physical activity and restorative care for residents in long-term care: a review of the literature. J Aging Phys Act 2014 Jan;22(1):154-165. [doi: 10.1123/japa.2012-0139] [Medline: 23434919]
34. Cotten SR, Schuster AM, Seifert A. Social media use and well-being among older adults. Curr Opin Psychol 2022 Jun;45:101293. [doi: 10.1016/j.copsyc.2021.12.005] [Medline: 35065352]
35. Seifert A. Digitale transformation in den haushalten älterer menschen. Z Gerontol Geriatr 2022 Jul 28;55(4):305-311 [FREE Full text] [doi: 10.1007/s00391-021-01897-5] [Medline: 33909129]
36. D'Cunha NM, Nguyen D, Naumovski N, McKune AJ, Kellett J, Georgousopoulou EN, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology 2019;65(4):430-440 [FREE Full text] [doi: 10.1159/000500040] [Medline: 31108489]
37. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: systematic umbrella review. JMIR Aging 2022 Oct 24;5(4):e40125 [FREE Full text] [doi: 10.2196/40125] [Medline: 36279155]
38. Juul A, Wilding R, Baldassar L. The best day of the week: new technology enhancing quality of life in a care home. Int J Environ Res Public Health 2019 Mar 19;16(6):1000 [FREE Full text] [doi: 10.3390/ijerph16061000] [Medline: 30893945]
39. Skurla MD, Rahman AT, Salcone S, Mathias L, Shah B, Forester BP, et al. Virtual reality and mental health in older adults: a systematic review. Int Psychogeriatr 2022 Feb;34(2):143-155. [doi: 10.1017/S104161022100017X] [Medline: 33757619]
40. Dascal J, Reid M, IsHak WW, Spiegel B, Recacho J, Rosen B, et al. Virtual reality and medical inpatients: a systematic review of randomized, controlled trials. Innov Clin Neurosci 2017;14(1-2):14-21 [FREE Full text] [Medline: 28386517]
41. Emmelkamp PM, Meyerbröker K. Virtual reality therapy in mental health. Annu Rev Clin Psychol 2021 May 07;17:495-519. [doi: 10.1146/annurev-clinpsy-081219-115923] [Medline: 33606946]
42. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400 [FREE Full text] [doi: 10.1017/S003329171700040X] [Medline: 28325167]
43. Optale G, Urgesi C, Busato V, Marin S, Piron L, Priftis K, et al. Controlling memory impairment in elderly adults using virtual reality memory training: a randomized controlled pilot study. Neurorehabil Neural Repair 2010 May;24(4):348-357. [doi: 10.1177/1545968309353328] [Medline: 19934445]
44. Lin TY, Huang CM, Hsu HP, Liao JY, Cheng VY, Wang SW, et al. Effects of a combination of three-dimensional virtual reality and hands-on horticultural therapy on institutionalized older adults' physical and mental health: quasi-experimental design. J Med Internet Res 2020 Nov 02;22(11):e19002 [FREE Full text] [doi: 10.2196/19002] [Medline: 33135666]
45. Chau PH, Kwok YY, Chan MK, Kwan KY, Wong KL, Tang YH, et al. Feasibility, acceptability, and efficacy of virtual reality training for older adults and people with disabilities: single-arm pre-post study. J Med Internet Res 2021 May 04;23(5):e27640 [FREE Full text] [doi: 10.2196/27640] [Medline: 33944795]
46. Fan CC, Choy CS, Huang CM, Chih PS, Lee CC, Lin FH, et al. The effects of a combination of 3D virtual reality and hands-on horticultural activities on mastery, achievement motives, self-esteem, isolation and depression: a quasi-experimental study. BMC Geriatr 2022 Sep 12;22(1):744 [FREE Full text] [doi: 10.1186/s12877-022-03431-7] [Medline: 36096746]
47. Liao YY, Tseng HY, Lin YJ, Wang CJ, Hsu WC. Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. Eur J Phys Rehabil Med 2020 Feb;56(1):47-57 [FREE Full text] [doi: 10.23736/S1973-9087.19.05899-4] [Medline: 31615196]

48. Kalantari S, Bill Xu T, Mostafavi A, Lee A, Barankevich R, Boot WR, et al. Using a nature-based virtual reality environment for improving mood states and cognitive engagement in older adults: a mixed-method feasibility study. Innov Aging 2022;6(3):igac015 [FREE Full text] [doi: 10.1093/geroni/igac015] [Medline: 35592668]
49. Niki K, Yahara M, Inagaki M, Takahashi N, Watanabe A, Okuda T, et al. Immersive virtual reality reminiscence reduces anxiety in the oldest-old without causing serious side effects: a single-center, pilot, and randomized crossover study. Front Hum Neurosci 2020;14:598161 [FREE Full text] [doi: 10.3389/fnhum.2020.598161] [Medline: 33536887]
50. Fiocco AJ, Millett G, D'Amico D, Krieger L, Sivashankar Y, Lee SH, et al. Virtual tourism for older adults living in residential care: a mixed-methods study. PLoS One 2021 May 20;16(5):e0250761 [FREE Full text] [doi: 10.1371/journal.pone.0250761] [Medline: 34014948]
51. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519 [FREE Full text] [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]
52. Roberts AR, De Schutter B, Franks K, Radina ME. Older adults' experiences with audiovisual virtual reality: perceived usefulness and other factors influencing technology acceptance. Clin Gerontol 2019 Mar 05;42(1):27-33. [doi: 10.1080/07317115.2018.1442380] [Medline: 29505343]
53. Visser M, Pluijm SM, Stel VS, Bosscher RJ, Deeg DJH. Physical activity as a determinant of change in mobility performance: the Longitudinal Aging Study Amsterdam. J Am Geriatr Soc 2002 Nov;50(11):1774-1781. [doi: 10.1046/j.1532-5415.2002.50504.x] [Medline: 12410894]
54. Quan NG, Lohman MC, Resciniti NV, Friedman DB. A systematic review of interventions for loneliness among older adults living in long-term care facilities. Aging Ment Health 2020 Dec 11;24(12):1945-1955. [doi: 10.1080/13607863.2019.1673311] [Medline: 31602993]
55. Wang D, MacMillan T. The benefits of gardening for older adults: a systematic review of the literature. Activ Adapt Aging 2013 Apr;37(2):153-181. [doi: 10.1080/01924788.2013.784942]
56. Linden M, Baron S, Muschalla B. Mini-ICF-Rating Für Psychische Störungen (Mini-ICF-APP): Ein Kurzinstrument Zur Beurteilung Von Fähigkeits- Bzw. Göttingen, Germany: Hans Huber; 2009.
57. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. JAMA 1963 Sep 21;185:914-919. [doi: 10.1001/jama.1963.03060120024016] [Medline: 14044222]
58. Wellbeing measures in primary health care. World Health Organization. URL: https://apps.who.int/iris/handle/10665/349766 [accessed 2024-03-14]
59. Schneider W, Henningsen P, Dohrenbusch R, Freyberger HJ, Irle H, Köllner V, et al. Begutachtung bei Psychischen und Psychosomatischen Erkrankungen: Autorisierte Leitlinien und Kommentare. Göttingen, Germany: Hogrefe Verlag; 2016.
60. Linden M, Baron S, Muschalla B, Ostholt-Corsten M. Fähigkeitsbeeinträchtigungen bei Psychischen Erkrankungen: Diagnostik, Therapie und Sozialmedizinische Beurteilung in Anlehnung an das Mini-ICF-APP. Göttingen, Germany: Hogrefe Verlag; 2015.
61. Muschalla B. Different work capacity impairments in patients with different work-anxieties. Int Arch Occup Environ Health 2016 May 29;89(4):609-619. [doi: 10.1007/s00420-015-1099-x] [Medline: 26515184]
62. Balestrieri M, Isola M, Bonn R, Tam T, Vio A, Linden M, et al. Validation of the Italian version of Mini-ICF-APP, a short instrument for rating activity and participation restrictions in psychiatric disorders. Epidemiol Psychiatr Sci 2012 Sep 19;22(1):81-91. [doi: 10.1017/s2045796012000480]
63. Molodynski A, Linden M, Juckel G, Yeeles K, Anderson C, Vazquez-Montes M, et al. The reliability, validity, and applicability of an English language version of the Mini-ICF-APP. Soc Psychiatry Psychiatr Epidemiol 2013 Aug 19;48(8):1347-1354. [doi: 10.1007/s00127-012-0604-8] [Medline: 23080483]
64. Leischker AH, Kolb GF. Diagnostik- und therapieziele beim alten Patienten. Internist (Berl) 2007 Nov 13;48(11):1195-1196. [doi: 10.1007/s00108-007-1946-2] [Medline: 17932635]
65. Lucas-Carrasco R, Allerup P, Bech P. The validity of the WHO-5 as an early screening for apathy in an elderly population. Curr Gerontol Geriatr Res 2012;2012:171857-171855 [FREE Full text] [doi: 10.1155/2012/171857] [Medline: 22991511]
66. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 well-being index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176 [FREE Full text] [doi: 10.1159/000376585] [Medline: 25831962]
67. IBM SPSS Statistics for Windows. IBM Corp. URL: https://www.ibm.com/products/spss-statistics [accessed 2024-03-14]
68. Li X, Niksirat KS, Chen S, Weng D, Sarcar S, Ren X. The impact of a multitasking-based virtual reality motion video game on the cognitive and physical abilities of older adults. Sustainability 2020 Nov 02;12(21):9106. [doi: 10.3390/su12219106]
69. Huang KT. Exergaming executive functions: an immersive virtual reality-based cognitive training for adults aged 50 and older. Cyberpsychol Behav Soc Netw 2020 Mar;23(3):143-149. [doi: 10.1089/cyber.2019.0269] [Medline: 31794673]
70. Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical virtual simulation in nursing education: randomized controlled trial. J Med Internet Res 2019 Mar 18;21(3):e11529 [FREE Full text] [doi: 10.2196/11529] [Medline: 30882355]
71. Maeng S, Hong JP, Kim WH, Kim H, Cho S, Kang JM, et al. Effects of virtual reality-based cognitive training in the elderly with and without mild cognitive impairment. Psychiatry Investig 2021 Jul;18(7):619-627 [FREE Full text] [doi: 10.30773/pi.2020.0446] [Medline: 34265200]

72. Wais PE, Arioli M, Anguera-Singla R, Gazzaley A. Virtual reality video game improves high-fidelity memory in older adults. Sci Rep 2021 Jan 28;11(1):2552 [FREE Full text] [doi: 10.1038/s41598-021-82109-3] [Medline: 33510315]
73. Saredakis D, Keage HA, Corlis M, Ghezzi ES, Loffler H, Loetscher T. The effect of reminiscence therapy using virtual reality on apathy in residential aged care: multisite nonrandomized controlled trial. J Med Internet Res 2021 Sep 20;23(9):e29210 [FREE Full text] [doi: 10.2196/29210] [Medline: 34542418]
74. Cheng VY, Huang CM, Liao JY, Hsu HP, Wang SW, Huang SF, et al. Combination of 3-dimensional virtual reality and hands-on aromatherapy in improving institutionalized older adults' psychological health: quasi-experimental study. J Med Internet Res 2020 Jul 23;22(7):e17096 [FREE Full text] [doi: 10.2196/17096] [Medline: 32706660]
75. Chan JY, Chan TK, Wong MP, Cheung RS, Yiu KK, Tsoi KK. Effects of virtual reality on moods in community older adults. A multicenter randomized controlled trial. Int J Geriatr Psychiatry 2020 Aug 11;35(8):926-933. [doi: 10.1002/gps.5314] [Medline: 32346896]
76. Brimelow RE, Dawe B, Dissanayaka N. Preliminary research: virtual reality in residential aged care to reduce apathy and improve mood. Cyberpsychol Behav Soc Netw 2020 Mar 01;23(3):165-170. [doi: 10.1089/cyber.2019.0286] [Medline: 31829729]
77. Brito H, Pham T, Vicente B. Effect of sensorimotor rehabilitation based on an immersive virtual reality model on mental health. Int J Geriatr Psychiatry (Forthcoming) 2021 Apr 02;37(1). [doi: 10.1002/gps.5541] [Medline: 33797806]
78. Cook DM, Dissanayake D, Kaur K. Virtual reality and older hands: dexterity and accessibility in hand-held VR control. In: Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference. 2019 Presented at: CHIuXiD '19; April 1-9, 2019; Bali, Indonesia p. 147-151 URL: https://dl.acm.org/doi/10.1145/3328243.3328262 [doi: 10.1145/3328243.3328262]
79. Seifert A, Reinwand DA, Schlomann A. Designing and using digital mental health interventions for older adults: being aware of digital inequality. Front Psychiatry 2019 Aug 9;10:568 [FREE Full text] [doi: 10.3389/fpsyt.2019.00568] [Medline: 31447716]
80. Seifert A, Cotten SR, Xie B. A double burden of exclusion? Digital and social exclusion of older adults in times of COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):e99-103 [FREE Full text] [doi: 10.1093/geronb/gbaa098] [Medline: 32672332]

Abbreviations

ADL: activity of daily living
HMD: head-mounted device
Mini-ICF-APP: Mini-ICF-Rating for Impairment in Psychological Activities and Capacities
VR: virtual reality
WHO-5: World Health Organization–Five Well-Being Index

Edited by T Leung, G Eysenbach; submitted 12.07.23; peer-reviewed by S Isbel, E Kawamoto; comments to author 04.09.23; revised version received 27.10.23; accepted 02.02.24; published 29.03.24.

Please cite as:
Li Y, Wilke C, Shiyanov I, Muschalla B
Impact of Virtual Reality–Based Group Activities on Activity Level and Well-Being Among Older Adults in Nursing Homes: Longitudinal Exploratory Study
JMIR Serious Games 2024;12:e50796
URL: https://games.jmir.org/2024/1/e50796
doi:10.2196/50796
PMID:38551635

Original Paper

A Serious Game (MyDiabetic) to Support Children's Education in Type 1 Diabetes Mellitus: Iterative Participatory Co-Design and Feasibility Study

Daniel Novak[1], MSc, PhD

Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic

Corresponding Author:
Daniel Novak, MSc, PhD
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague
Technicka 1902
Prague, 16627
Czech Republic
Phone: 420 224357314
Email: xnovakd1@fel.cvut.cz

Abstract

Background: Serious games, which are gaming applications used for purposes beyond entertainment to educate users on, and address, specific issues, may present a timely approach to promote healthy diabetes management behaviors among children with type 1 diabetes mellitus (T1DM). The lasting benefits associated with these serious games encompass improved patient education; enhanced glycemic control; the reinforcement of bonds within the community of people with diabetes; the facilitation of meaningful dialogues with caregivers, especially within the familial setting; and a significant reduction in the economic burdens associated with subsequent complications.

Objective: This paper primarily aims to provide a detailed overview of the iterative design process and the associated evaluation methods used in the development of the educational game. Furthermore, this study aims to enhance motivation for sustained and extended engagement with the game over time. The MyDiabetic game design aims to educate children on various aspects, including the connections among food, insulin, and physical activity. Furthermore, it seeks to impart knowledge related to the operation of a glucometer and an insulin pen, as well as more advanced technologies such as administering glucagon, measuring ketoacidosis, and continuous glucose monitoring.

Methods: The co-design methodology was applied, involving interviews, design workshops, and prototype feedback sessions. A combination of several approaches, such as tailoring, observational learning, social and family support, decision-making practice, and reward systems, was used to support children's compliance. Moreover, incorporating the literature, guidelines, and current practices into the design ensured that the game was aligned with established health care pathways and included relevant information and best practices for diabetes management.

Results: The game was tested on 32 children in 3 iterations. Positive responses were received from children who tested the game as well as their parents. The game was also presented to 5 schoolmates of children with T1DM who appreciated a better understanding of the disease and the opportunity to support their friends more efficiently in T1DM compensation. The involvement of children and clinicians in participatory co-design contributed to to the game's high acceptance. With regard to the game's impact on education, 1 week of testing revealed an enhancement in educational outcomes.

Conclusions: The game is especially suitable for children newly diagnosed with T1DM because it acquaints them in a fun way with new terminology; for example, they can try to measure glycemia levels in an interactive way. The game also caters to children who still need to develop reading skills by including an audio guide. The guide ensures that children of all literacy levels can benefit from the game's educational content and interactive experiences. The game is available for download on Google Play and the Apple App Store.

(JMIR Serious Games 2024;12:e49478) doi:10.2196/49478

KEYWORDS

diabetes mellitus; serious games; mobile app; co-design; user-centered design; serious game; gaming; diabetes; child with diabetes; child; children; insulin; glucometer; glucose; patient education; insulin; mobile phone

Introduction

Background

Type 1 diabetes mellitus (T1DM) is the most common form of diabetes in children. Poor blood glucose control increases the risk of chronic microvascular complications, including renal and retinal complications. When T1DM starts in childhood or young adulthood, the course of the disease is long, and metabolic control is essential to prevent such complications [1]. Advances in the treatment of T1DM have decreased the risk of complications and delayed their occurrence, with a resultant overall increase in the quality of life of patients. Nutritional education, with systematic assessments of carbohydrate intake and the use of the insulin-to-carbohydrate ratio, has allowed for optimizing insulin dosage [2]. The basal-bolus scheme with multiple-dose insulin injections, continuous subcutaneous insulin infusion, and multiple capillary blood glucose measurements allow for better metabolic control. For this, adequate and continued diabetes education for patients and families is required [3].

In general, mobile phones are a natural choice for use in increasing the efficiency of medical care and patient motivation [4,5]. The development of monitoring technologies and designing systems for improving diabetes mellitus compensation based on a mobile platform is increasing and widespread [6,7]. In the field of diabetes mellitus, there are already several pilot studies that validate the effectiveness of this new technology [1,8-10]. The biggest problem so far is the motivation of patients, as many stop cooperating after a few weeks or months. Furthermore, some authors state that the noncooperation of patients is one of the most common reasons why we encounter failure in education outcomes [4,11,12].

Well-designed serious games can improve children's learning, skills development, attitudes, emotions, motivation, and many other factors that encourage children to work together with family members and health professionals on treatment [13]. The long-term benefits of the serious game are improved patient education; better diabetes compensation; increased connections with the community of people with diabetes; stimulation of discussion with caregivers, especially within the family; and a significant reduction in the economic costs of subsequent complications. Furthermore, due to the widespread use of technology among children, using serious games to educate and support health behavior for children with diabetes self-management is an emerging and promising practice [14,15].

Some examples of previous approaches are Packy & Marlon (Super Nintendo) [16], Balance [17], Mario Brothers [18], Monster Manor (Nintendo) [19], and mySugr Junior [20]. Carb Counting with Lenny [21] is geared toward teaching users valuable information about healthy food choices and allowing them to apply that knowledge during gameplay. Jerry the Bear [22] seeks to educate children about T1DM by getting them to take care of the game's avatar, check the avatar's blood glucose level, manage insulin dosage by administering the doses using a pen or a pump, and feed the avatar with various food items. Except in the case of Jerry the Bear, the games show only the relationship among food, blood glucose level and physical activity. Some diabetes management games tested in small populations have never been brought to market and have no scalability. In addition, no studies were identified examining T1DM educational games' impact on user behavior, knowledge, or clinical indicators. Further research is needed to better understand the sustainability of T1DM gaming as a tool for promoting adherence and the effect of education.

Objectives

This paper primarily aims to provide a thorough overview of the iterative design process and the accompanying evaluation methods applied in the development of the educational game, MyDiabetic. Furthermore, this study aims to enhance motivation for sustained and prolonged engagement with the game.

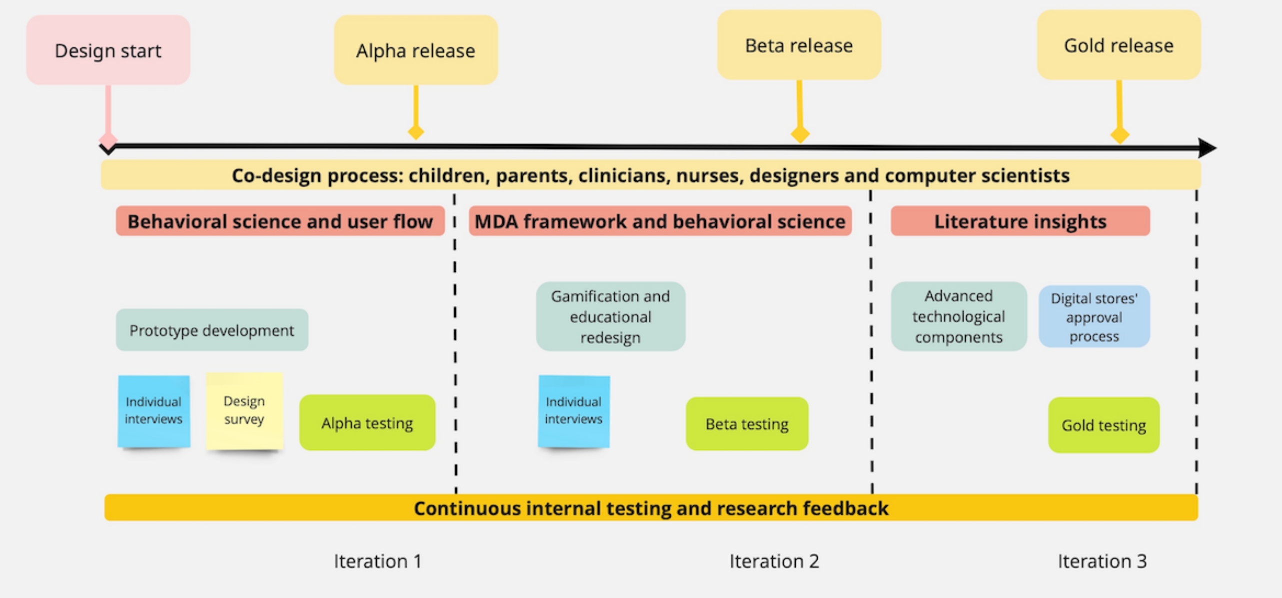
MyDiabetic aims to not only teach children the relationship among food, insulin, and physical activity but also pass on knowledge related to working with a glucometer and an insulin pen, as well as more advanced technology such as continuous glucose monitoring (CGM), insulin pump, glucagon administration, and ketones urine test. The game is designed for children aged between 5 and 12 years. It is especially suitable for children newly diagnosed with T1DM because it acquaints them in a fun way with new terminology; for example, they can try to measure glycemia levels in an interactive way.

Methods

Overview

A participatory iterative co-design approach was adopted. Participatory design in each iteration was guided by the fundamental principles from both traditional game design elements (eg, user flow and the mechanics, dynamics, and framework approach) and behavioral theory tailored for diabetes support. Figure 1 illustrates an approximate timeline for the MyDiabetic project, depicting key components of the participatory co-design methodology, along with the predominant principles used in each iteration indicated by the red bars.

Figure 1. Approximate timeline for the MyDiabetic project showing the principal elements of the participatory co-design methodology along with the predominant principles used in each iteration indicated by the red bars. MDA: mechanics, dynamics, and aesthetics.



Participatory Iterative Co-Design

The game was continually developed, resulting in the 3 iterations of the co-design phases applying participatory co-design principles. Figure 1 presents the timeline of the co-design process. Co-design process is intrinsically incorporated into the research. Co-design is a research methodology that actively engages diverse stakeholders to initiate, create, and validate solutions by adding creative and participative principles and tools [23]. Participatory co-design establishes that the necessary features of person-centered design, clinical acceptability, and health IT feasibility are accounted for, with each process needed for the ultimate success of the serious game. Table 1 describes user statistics during the co-design process.

The design team included several types of participants:

- Children in the user group and their parents, who were recruited through collaborating institutions, such as patient organizations, or Facebook groups
- Researchers with backgrounds in behavior change, informatics, and design
- Designers and developers (external and from the in-house IT system development group)
- Clinicians with vast expertise in T1DM management and nutrition nurses who are part of the project's advisory group

The first iteration of the MyDiabetic game was based on qualitative and quantitative research with children with T1DM diabetes, which led to the requirement for a game to help educate these children and help them cope with their new diagnosis. The outcome of the co-design research was the concept of continuous care of an avatar (based on the Tamagotchi principle) that is represented as a 3D avatar, including technical modules illustrating glucometer and insulin pen use. Subsequently, the co-design of the second iteration focused on incorporating educational material, such as an educational library and audio guidance provided by a physician's avatar, as well as developing other gamification features, such as game levels, minigames, and storyline extension. In addition, a model for measuring blood glucose levels was added. The third iteration addressed new technologies and methods, such as CGM and insulin pump modules, glucagon administration, and an explanation of ketoacidosis. Each iteration ended with usability testing. We followed software release cycle terminology. The alpha usability testing was focused on the evaluation of the main game concept, while the beta usability testing was the most detailed, concentrating on educational and adherence aspects. The gold usability testing evaluated new technological components only. The depiction of screens can be seen in Figure S1 in Multimedia Appendix 1.

Detailed information regarding the characteristics of the children for each iteration can be found in Tables S1-S6, S9-S10, and S12-S13 in Multimedia Appendix 2.

The clinical team was composed of experts who possessed diverse experience in various aspects of T1DM. Throughout the development process, these clinicians consistently provided invaluable feedback during presentation workshops conducted at the conclusion of each iteration.

Table 1. Description of iterative design and user statistics.

Iteration and activity	Users	Age (years), mean (SD)	T1DM[a] duration (years), mean (SD)	Stage of development	Features introduced
Iteration 1					Basic gaming concept, design of control elements, taking care of the avatar concept, carbohydrate counting, insulin pen administration, measurement of glycemia level, and performing exercise
Interviews	Boys: 4; girls: 0	12.8 (1.3)	5.5 (2.6)	Mock-up	
Design survey	Boys: 15; girls: 12	10.2 (2.1)	4.8 (3.0)	Mock-up	
Alpha testing	Boys: 8; girls: 4[b]	10.5 (2.7)	3.7 (2.2)	Alpha release	
Iteration 2					Education elements, such as an educational library or simulations of blood glucose, minigames, storyline, and levels design
Beta testing	Boys: 2; girls: 4	9.6 (4.5)	4.8 (2.9)	Beta release	
Iteration 3					Advanced topics such as CGM[c] and insulin pump illustration, measurement of ketoacidosis, and glucagon administration
CGM testing	Boys: 7; girls: 2	12.8 (1.7)	7.0 (3.7)	Gold release	
Glucagon testing	Boys: 4; girls: 4	13.2 (2.6)	6.4 (1.8)	Gold release	
Keto-acidosis testing	Boys: 2; girls: 5	12.2 (2.7)	8.9 (3.2)	Gold release	

[a]T1DM: type 1 diabetes mellitus.

[b]Of the 12 users, 5 (42%; n=2, 40% boys and n=3, 60% girls) did not have diabetes.

[c]CGM: continuous glucose monitoring.

Serious Games and Behavioral Theory Foundations

Methodological aspects of user flow [24]; the mechanics, dynamics, and aesthetics framework [25]; and behavioral theory (tailoring [26,27], observational learning [27-29], decision-making practice [13], social and family support [30,31], and reward systems [27,32]) were used during the design of MyDiabetic. Theoretical backgrounds are summarized in the Theoretical Background and Game Design section in Multimedia Appendix 1.

On the basis of literature research, guidelines, and current practices [13,19,21,22,33], the characteristics that are important for a successful serious game were further identified (Textbox 1).

Textbox 1. Characteristics that are important for a successful serious game.

Important characteristics

1. The main avatar should be empathetically connected to the player. One possibility is to use avatars that reflect the player in the game [34]. Customizing the avatar in game should also be possible (refer to the Avatar Tailoring subsection under Theoretical Background and Game Design and Figure S2A in Multimedia Appendix 1).
2. The game should have an incentive and customizable reward system that supports the player’s learning. As a reward, the game can offer trophies or new *unlocked* game content (refer to the Level Progress Design subsection under Theoretical Background and Game Design and Figure S2B in Multimedia Appendix 1).
3. The game should aim to develop the player’s skills by setting clear but challenging goals related to changing the player’s behavior in real life (mainly all educational and technical features developed in iterations 1-3: iteration 1 [main game concept], iteration 2 [educational and gamification features], and iteration 3 [advanced features]).
4. The difficulty should gradually increase and adapt to gradually improving the player’s skills, giving the player further opportunities for improvement (refer to the Level Progress Design subsection under Theoretical Background and Game Design in Multimedia Appendix 1).
5. The game should have a realistic and health-related story. This attractive design includes high-quality graphics, sounds, and animations to immerse the player in the game (refer to the Scenes Description subsection under Theoretical Background and Game Design and Figure S1 in Multimedia Appendix 1).

A summary of the game and the methodology is provided in Multimedia Appendix 3. In addition, Multimedia Appendix 4 contains a video highlighting the game's main features.

Participant Recruitment

The whole concept was iteratively validated throughout the 3-stage iteration design process. Children were recruited through the project website [35], at diabetes summer camps, and via Facebook groups and through nonprofit organizations dealing with treatment support. The eligibility requirements were as follows: (1) fluency in Czech, (2) possession of a smartphone for a minimum of 2 weeks, (3) a confirmed diagnosis of T1DM for at least 1 year, and (4) aged between 5 and 16 years.

Ethical Considerations

The design process was approved by the committee for research ethics at the Czech Technical University in Prague, Czech Republic (0000-01/24/51902/EKČVUT). Written study information was provided, and written informed consent was obtained from parents. All users agreed to provide anonymized data for research and data analysis during the sign-up process, which was required for app use.

Development Approach

Agile principles [36] were used for the rapid development of various components and for iteratively integrating. Undergraduate and graduate students (computer science and biomedical engineering students) participated in developing new functionalities over 8 semesters. The game was prereleased on Google Play at the end of each iteration design process for feasibility testing. User feedback and bug fixes were communicated in weekly core team meetings and daily written discussions (via email and Slack). During the academic term, student meetings were also held once a week. A cross-platform solution, Unity software (Unity Technologies) [37], was used for game development; a free and open-source platform, Blender, was used for 3D modeling [38]; the GitHub tool [39] was used for versioning; and the Trello project management tool (Atlassian) [40] was used to track bugs and for project management. Furthermore, each task was developed in a separate branch on GitHub before being tested and approved by another team member and then merged into the master branch.

Results

Herein are presented the results of the design and development phase that gave rise to the MyDiabetic game as a person-centered education tool, along with details of the contributions provided by each iteration of the participatory co-design methodology.

Iteration 1: The Main Game Concept

Overview

In the first phase, the main game framework was outlined, focusing on carbohydrate counting and elementary technological tools for the measurement of glycemia levels and insulin administration. In addition, the basic gamification concept was proposed. The main aim of the usability study was the determination of user experience.

Individual Interviews and Design Survey

The co-design methodology was applied throughout the whole game design process. The main gaming concept was built in the first co-design iteration. First, qualitative (individual interviews) and quantitative research (design survey) methods were used to collect game design requirements. The additional documents for interviews are described in Multimedia Appendix 1 (refer to the Screener, Session Guide, and Design Survey subsections under Alpha Release). In addition, research was performed among parents, who are very important stakeholders in T1DM management. The questionnaire for parents is presented in the Questionnaire for Parents subsection under Alpha Release in Multimedia Appendix 1; a summary is presented in Table S1 in Multimedia Appendix 2.

User needs were specified, for which an informal interview was conducted. Four users with T1DM were invited to the interviews. Two observers were present during the interviews. One observer moderated the session, while another took notes, kept track of the time, and made observations. This was followed by rapid prototyping using the user-centered design methodology. The prototype used the basic paradigm mentioned in the Theoretical Background and Game Design section in Multimedia Appendix 1. The summary of the interviews is presented in Table S2 in Multimedia Appendix 2.

Parents play a big role in diabetes compensation. In this study, in children diagnosed at a young age, parents took care of insulin administration, the measurement of blood glucose levels, meal preparation, and writing in the diabetes diary. Parents would check their children regularly, some excessively, especially if they were physicians themselves. All participants used paper diabetes diaries because it was almost always the participants' mothers who took up the responsibility of maintaining the diaries, and they did not seem to be very confident about using mobile app technology (if the children were to enroll, they would have preferred to use a mobile app). The children learned about managing the disease by observing their parents' actions. Initially, when they were at school, they had to call their parents, who told them the appropriate insulin dose to inject. In time, the children became confident, and consulting their parents was no longer necessary. Quantitative research was carried out at diabetes summer camps (refer to the Design Survey subsection under Alpha Release in Multimedia Appendix 1 as well as Table S3 in Multimedia Appendix 2). In total, 27 questionnaires were collected among children aged 7 to 13 years. The main interest centered around where children obtained most of their information about diabetes and if they had ever encountered a game about diabetes. Other topics of interest were current favorite games among the children and how independent the children were in terms of counting bread units (BUs) and insulin administration.

On the basis of the results (refer to Tables S1-S3 in Multimedia Appendix 2), the main requirements for the game can be summarized as follows:

- Teaching the relationship among food, insulin, and physical activity to children with diabetes
- Teaching children with diabetes how to count bread exchange units
- Demonstrating the symptoms of hypoglycemia and hyperglycemia and their solutions
- Demonstrating insulin administration, including technical skills and dose adjustment, and explaining how to assess diabetes compensation by blood testing
- Explaining the nature of the disease
- Motivating children to exercise
- Being available free of charge and to as many people as possible
- Being both educational and fun so that children can continue playing it for as long as possible

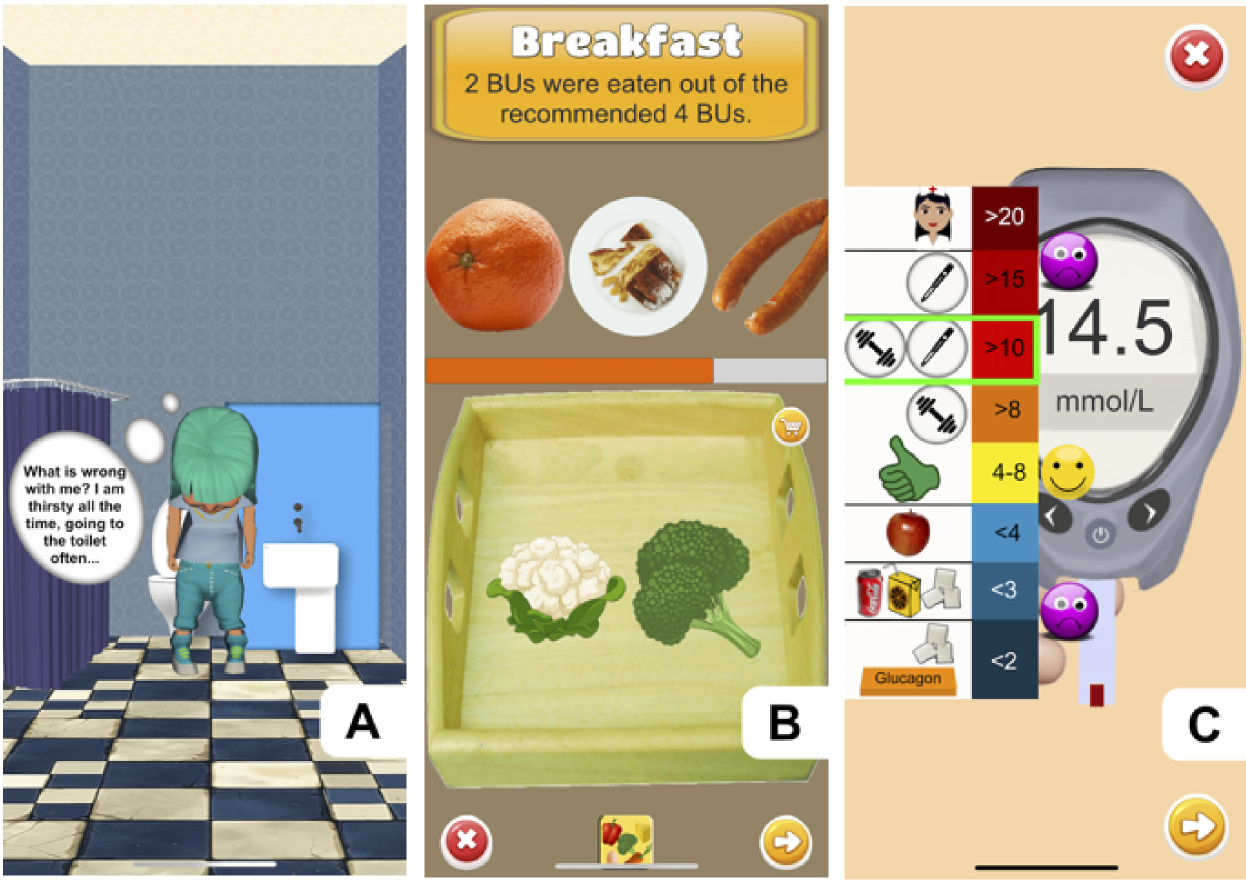
When designing the game, the choice was driven by the experience of children from quantitative research (refer to Table S3 in Multimedia Appendix 2): the most played game was Pou, which involves caring for a simulated creature, which is a proven gameplay principle of other successful games such as Moy, My Talking Tom, My Talking Angela, The Sims, and the older Tamagotchi. These games are based on the player taking care of a pet. Games of this type have had great success in the past and in recent times [34,41].

The primary design concept of the game divides the avatar's day into 6 parts, representing 6 meals, similar to the routine of a person with diabetes. The player's tasks include measuring the avatar's blood glucose level, administering insulin, and feeding the avatar. For completing these tasks, the player is rewarded with virtual coins, which can be used to enhance the avatar as well as buy new furniture, food, and other items. Neglecting the avatar's care can result in the avatar experiencing hypoglycemia or hyperglycemia, which the player can identify based on the virtual glucometer reading and the avatar's symptoms.

Onboarding and Tutorial

The game includes an introduction in which the avatar experiences the initial symptoms of diabetes and is transferred to the hospital in an ambulance, where it is diagnosed by a physician (Figure 2A).

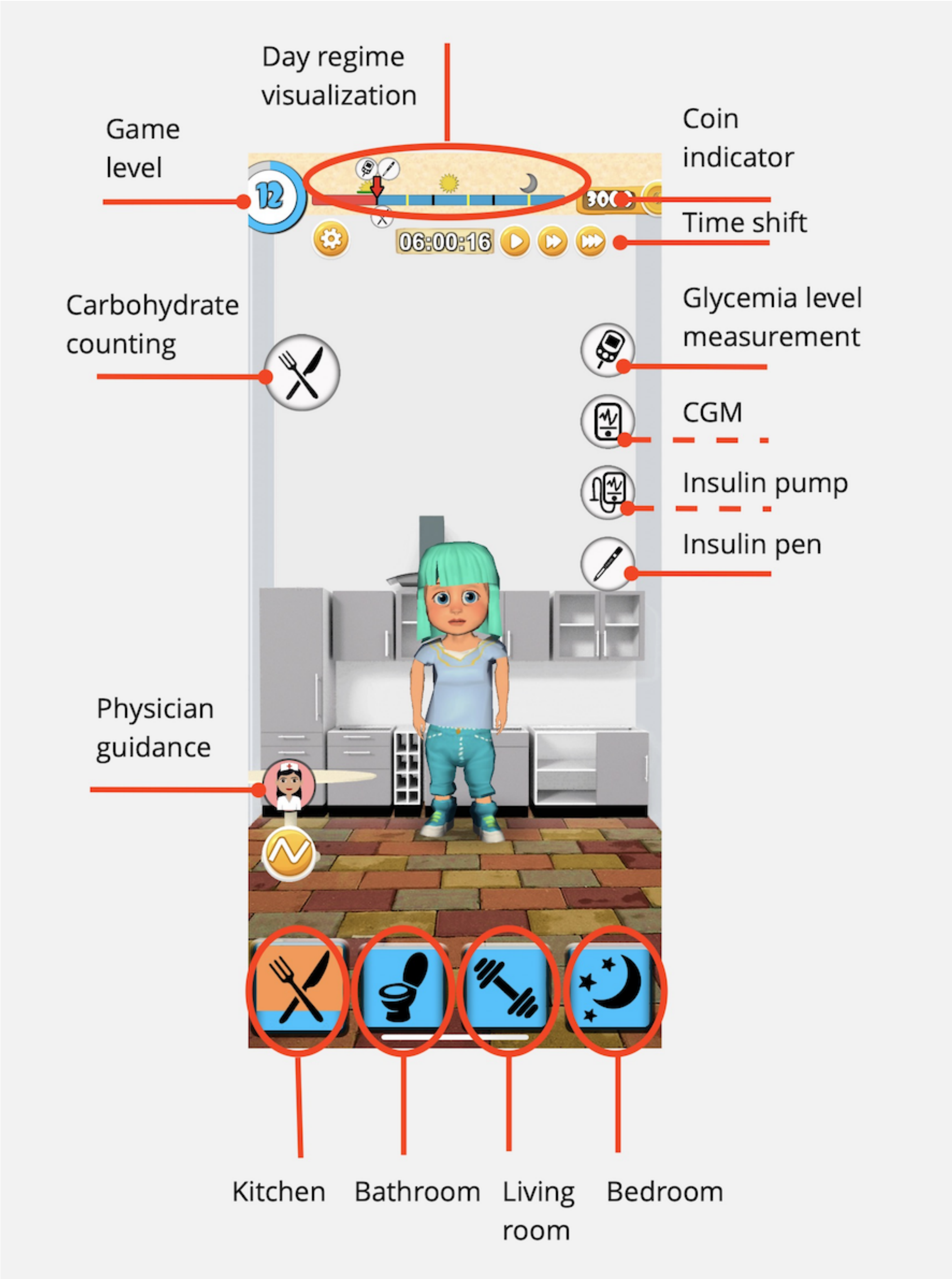
Figure 2. (A) Onboarding, including the first symptoms and diagnosis. (B) Carbohydrate counting. (C) Measuring blood glucose level using a glucometer. BU: bread unit.



Control Elements

Herein, the control elements of the game are described using the example of a kitchen scene (Figure 3). The different rooms are represented by an icon in the bottom panel. Each icon serves not only as a switch to access that room but also to indicate whether one of the avatar's basic needs, such as sleep or food, is met. In the upper left corner, the player is reminded of the level the avatar is currently in. The key is to follow the daily routine, which can be seen on the timeline. The timeline is the same every day. During the repetition of basic actions, the player learns about the problem.

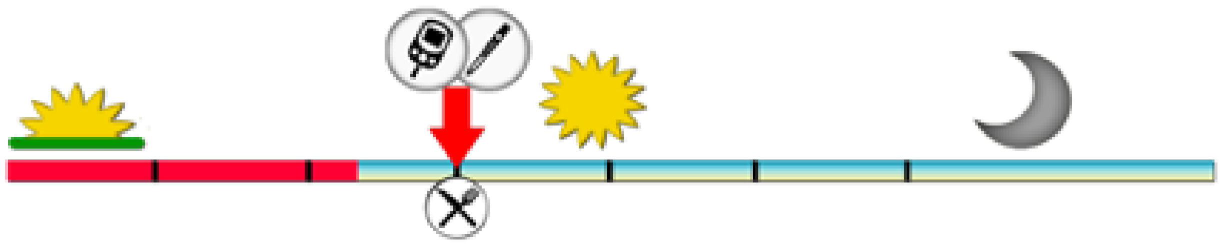
Figure 3. Description of the control elements. Continuous glucose monitoring (CGM) and insulin pen features were added in the final iteration (iteration 3) and appear in level 5.



If needed, time can be accelerated using the designated button. The amount of coins currently available in the avatar's cashbox is displayed on the right side.

To remind players of basic tasks such as measuring blood glucose levels, injecting insulin, and eating, a timeline was designed that always shows 1 task ahead, and all tasks are marked on the timeline so that the player knows approximately when the next task will occur (Figure 4).

Figure 4. Illustration of the timeline for the day regime of a person with diabetes.



Carbohydrate Counting

The player must choose the proper combination from the offered food items to achieve the specified number of BUs as closely as possible. The offered food items will be selected randomly from the purchased food so that the player cannot cheat (eg, by learning the BU component of only 1 food item and repeatedly giving it to the avatar). The player can see food items as a picture of food on a plate to estimate the amount of food based on the plate size (Figure 2B).

Glycemia-Level Measurement

There is no blood glucose–level indicator in the game. The player determines the exact blood glucose level by measuring it, for which they receive a reward. The virtual measurement mimics the real measurement (Figure 2C). The needle of the lancet becomes dull over time and needs to be replaced. A limited number of test strips are inserted into the pen and run out over time. The player must buy new needles and test strips in the virtual store with the collected coins. Estimating the blood glucose level based on the avatar's current mood is possible; for example, in the case of hypoglycemia or hyperglycemia, the avatar appears sick and in a bad mood (Figure 5C).

Figure 5. (A) Performing exercise. (B) Administration of insulin with an insulin pen. (C) Visualization of the symptoms of acute hyperglycemia.

Administration of Insulin

Insulin administration too mimics real-world settings. The player learns certain habits, such as pointing the needle skyward, giving the pen a little flick or tap to loosen any bubbles to the top, and performing an *airshot* (expelling air, thus priming the needle for injecting and avoiding possible air bubbles in the cartridge). The user is first required to assemble an insulin pen by removing the cap from the pen, taking a new needle, removing it from the sterile packaging, inserting it in the pen, and removing the needle cap and needle shield (Figure 5B). The game only allows the user to inject insulin once the needle has been primed for injecting. The game advises the user on the appropriate dose of insulin, as well as the type of insulin to administer to the avatar. The user administers short-acting insulin before main meals and long-acting insulin at night. Changing the needle and changing the insulin cartridge (the amount of insulin in it gradually decreases) are also incorporated.

Performing Exercise

The avatar requires regular exercise; otherwise, the avatar will be in a bad mood. The player determines the duration of the exercise (Figure 5A). If the exercise is too long, the avatar goes into hypoglycemia. The muscles tend to be more insulin sensitive for 1 to 2 days after exercise, leading to an increased risk of hypoglycemia, but this is not implemented in the game. To motivate the player to be active, the game offers them bonus coins for performing physical activity. If the player owns a Fitbit fitness bracelet, it is possible to connect it to the game and receive a virtual coin for every 2 steps taken in real life.

Alpha Usability Testing

The main aim of alpha usability testing was to determine users' experience and assess their understanding of the main game concept. In total, 12 children participated in alpha testing (refer to Table S4 in Multimedia Appendix 2): 7 (58%) with diabetes (n=5, 71% boys and n=2, 29% girls; aged between 6 and 13 years) and 5 (42%) without diabetes (n=2, 40% boys and n=3, 60% girls; aged between 6 and 13 years). The involvement of children without diabetes is important to address the reduction of stigma and to provide education to friends and schoolmates. The recruitment was performed through the Association of Parents and Friends of Diabetic Children. The testing was organized at participants' premises and consisted of completing the pretesting screener (refer to the Screener subsection under Alpha Release in Multimedia Appendix 1) and testing (onboarding and going through all 4 basic screens). One observer was always present to take notes.

Regarding user experience and game mechanics, all participants were satisfied; they liked the design and were able to find their way around in all scenes, although only after several attempts. Most users asked for help, both written and audio. Some users were looking for basic educational information about diabetes. The majority of users mentioned the issue of long-term motivation to use the game regularly. The main feedback is summarized in the Feedback Summary subsection under Alpha Release in Multimedia Appendix 1. Most shortcomings were addressed in future development in iterations 2 or 3.

Iteration 2: Educational and Gamification Features

Overview

On the basis of the outcomes of the usability alpha testing and continual feedback from the clinical team (nurses and physicians), in the next iteration phase, the focus of the main design effort was on increasing adherence, game experience, and educational impact by incorporating (1) game level design (Figure S2B in Multimedia Appendix 1); (2) minigames concept and storyline (Figures S3A, S8A, and S8B in Multimedia Appendix 1); (3) virtual guidance; and (4) an educational library, including the simulation of glycemia [42] (refer to the Theoretical Background and Game Design section in Multimedia Appendix 1 for a description of the new add-ons). Incorporating the audio guidance helps preschool children who cannot usually read yet to better orient themselves within the game. The main aim of the feasibility follow-up testing was the determination of educational and adherence effects.

Beta Usability Testing

The recruitment was organized via Facebook groups and targeted children aged 5 to 15 years willing to play the game for at least 1 week. Children without diabetes were not included because this feasibility test aimed to determine the educational and adherence effects in children with T1DM. In addition, adherence was tested by inspecting game statistics implemented by applying the Google Analytics framework.

Individual Interviews

The interviews were conducted in person with the participant and their parents. The first questionnaire was the screener (refer to the Screener subsection under Beta Release in Multimedia Appendix 1), which helped to place the children in the correct category corresponding to the desired target group of the participants. This was followed by a general questionnaire (refer to the General Questionnaire subsection under Beta Release in Multimedia Appendix 1) to estimate the participant's interest in games. Furthermore, the participant was asked questions regarding their management of their diabetes (refer to the Diabetes Management subsection under Beta Release in Multimedia Appendix 1). The user was also presented with a questionnaire on diabetes education (refer to the Pretesting Knowledge of Diabetes subsection under Beta Release in Multimedia Appendix 1). Subsequently, the game was installed on their mobile device. This was followed by observation of the participant and their behavior as they undertook the first steps in the game. The participant was required to go through the whole game tutorial and later try to find their way around the game for a few minutes. The gameplay was then interrupted, and the participant was asked about first impressions (refer to the First Impression of the Game subsection under Beta Release in Multimedia Appendix 1). The participant's parents were then interviewed (refer to the Questionnaire With Parents subsection under Beta Release in Multimedia Appendix 1). The main findings of the interviews are summarized in Table S5 in Multimedia Appendix 2. Seven participants took part in the study, of whom 1 (14%) did not complete the usability testing and was excluded. Of the remaining 6 participants, 1 (17%) was a boy and 5 (83%) were girls, and they were aged 5 to 15 years. The average duration of diabetes was 4.8 (SD 2.9) years; 4 (67%) of the 6 participants were using an insulin pump.

Beta Usability Testing

The participant was asked to play the game for at least 1 week, after which they were contacted by telephone, and the second part of the test questionnaire was discussed. First, the participant was asked about their impressions of the app after the 1-week testing (refer to the Beta usability testing subsection under Beta

Release in Multimedia Appendix 1), followed by a knowledge questionnaire on diabetes (refer to the Posttesting Knowledge of Diabetes subsection under Beta Release in Multimedia Appendix 1). The educational effect of the game is outlined in the next subsection. Regarding the usability testing results, the main findings are summarized in Table S6 in Multimedia Appendix 2.

In summary, the participant group consisted of children in different age ranges. Of the 7 participants, 2 (29%) were aged <7 years and could not read; 1 participant (14%) could not complete the usability testing because he required hospitalization. Of the 4 participants aged 8 to 12 years, 3 (75%) expressed high levels of enjoyment and found the game entertaining. Evidently, all children in this category understood the game well, confirming the study hypothesis that this age group should be the primary focus. Of the 6 participants who completed the testing, the remaining 2 (33%), aged 13 to 15 years, also understood the game well, although they commented that the entertainment aspect could have been more engaging. They highly recommended the game to younger audiences and stated that they wished such an educational tool had been available when they had been first diagnosed with diabetes. In general, the girls particularly enjoyed shopping, while the boy was more interested in minigames.

Educational Effect

Table 2 summarizes the educational impact based on the answers to the questionnaires, showing the level of education before and after the testing. The specific percentage response rates are given in Tables S7 and S8 in Multimedia Appendix 2. This score is expressed as a percentage depending on the quality of the correctly answered questions and the degree of confidence in answering them (assessed subjectively). If the answer is accurate and complete (correct), then 100% is given for this question. If the answer is partially correct or incorrect, a percentage measure of accuracy is estimated, where 0% is given for an outright nonsensical answer or if the participant does not know the answer.

Table 2. Effect of education.

	Educational score before testing (%)[a]	Educational score after testing (%)[b]	Difference (%)[c]
Participant 1	39	44	5
Participant 2	96	98	2
Participant 3	27	55	28
Participant 5	80	93	13
Participant 6	84	95	11
Participant 7	89	94	5

[a]Average 69.2.

[b]Average 79.8.

[c]Average 10.7.

Table 2 depicts the results of the educational effect after 1 week of usability testing. On average, the increase in education knowledge was approximately 10%; in the particular case of participant 3, the gain was considerably high: 28%.

Statistics Summary

Regarding game analytics, a significant improvement in skills was observed throughout the gameplay, specifically in the time taken to perform certain procedures. The time taken for glycemia-level measurement decreased by 42%, and the time taken for insulin administration decreased by 45% from the first interaction (day 1) to the last interaction (day 7).

Analyzing the purchase behavior of players, the most frequently bought food items were the large basket (46%), followed by the medium-sized basket (26%), and the small basket (28%). As for medical supplies, the majority of purchases consisted of glucometer strips (59%), followed by lancet packs (18%), injectable lancets (12%), and insulin (11%).

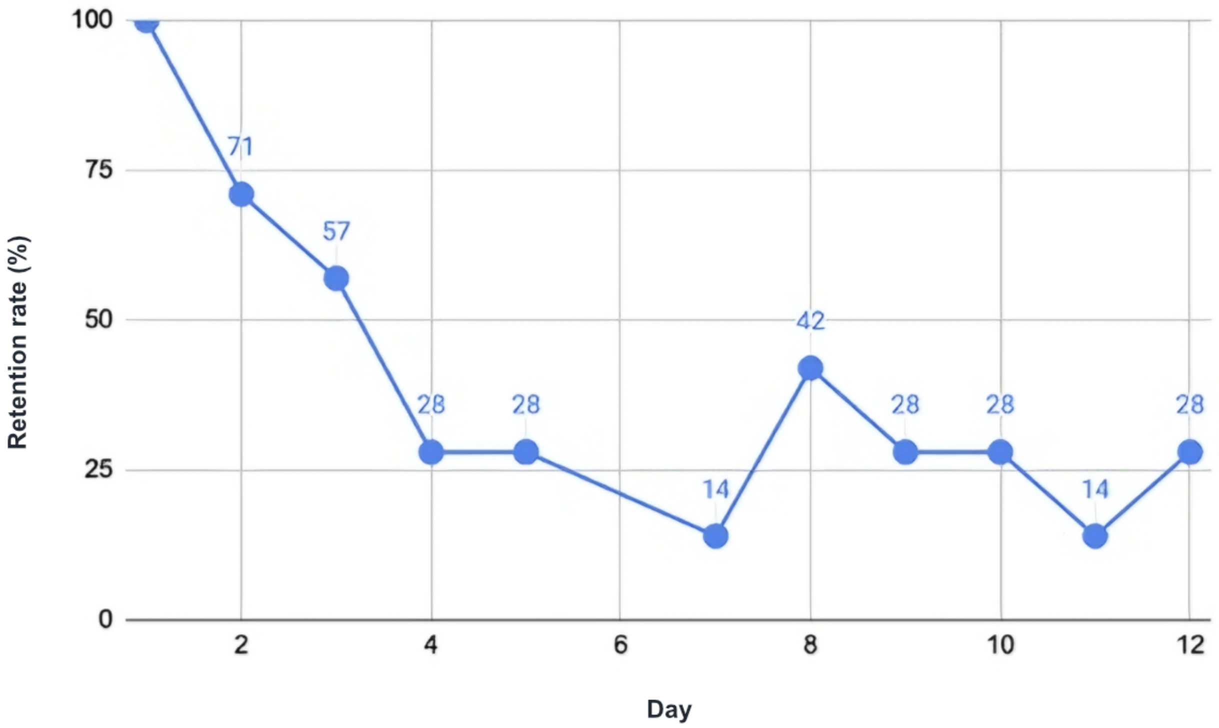
Examining the distribution of glycemia statistics, glucose measurements were most commonly applied on the middle finger (50%), followed by the index finger (22%), ring finger (19%), little finger (6%), and thumb (3%). Short-term insulin injections were most frequently applied to the abdomen (29%), left biceps (12%), and left thigh (7%). Long-lasting insulin injections were predominantly applied to the left and right thighs (11% and 12%, respectively) and to the chest (6%). The less frequent injection site was the right thigh 0.78%, of the 129 injections recorded.

During the gameplay, the avatar was hospitalized 31 times, with 15 (48%) of the cases involving hypoglycemia and 16 (52%) involving hyperglycemia. The average blood glucose level during these hospitalizations was 15.7 mmol/L. In addition, the avatar engaged in exercise 72 times, with 36 (50%) of the sessions lasting for 30 minutes, 19 (26%) for 90 minutes, and 17 (24%) for 60 minutes. However, the exercise was completed in only 32 (45%) of the cases, with early terminations occurring in the remaining 40 instances (55%).

The final phase of the usability testing focused on measuring adherence. At the beginning of the game, the day and time were recorded to mark the initial interaction. Subsequently, for each following day, whether the player returned to the game was noted. The retention rate, depicting the percentage of players who continued playing the game over time, is illustrated in Figure 6.

In the first week, the participants tested the game without any intervention. The telephone interview was performed on day 7. As shown in Figure 6, the adherence rate was >14%, and on the following day, day 8, the rate increased to 42% after the interview. The time spent playing the game on day 1 was approximately 39 minutes per user; on day 7, it was approximately 25 minutes. Some participants even played the game for 1.5 hours on 1 day. The daily average time for all participants was approximately 17 minutes.

Figure 6. Retention rate of participants.



Participant Feedback

Given the high number of participants (70%) using insulin pumps, it was unsurprising that there were requests to add insulin pump functionality and CGM in the game. Some users also expressed a desire for more advanced tools, such as the ability to administer glucagon and measure ketoacidosis. In response to this feedback, these features were included in iteration 3 of the game.

Of the 6 participants, 2 (33%) indicated that the game became routine after a certain period and that they desired some variation. Of these 2 participants, 1 (50%) suggested that a solution could be to introduce complications related to diabetes, such as cardiovascular issues, neuropathy, or retinopathy, which the player would need to manage.

In the real world, various types of insulin are available, each designed with different properties regarding onset, peak, and duration. Participants expressed the desire to be able to purchase and use the specific insulin types that they are familiar with and to understand their characteristics within the game.

Iteration 3: Advanced Technological Components

On the basis of the feedback received from the beta usability testing, in this iteration, more advanced tools for diabetes management were implemented. Subsequently, the final usability study briefly evaluated the new modules.

CGM and Insulin Pump Implementation

This module illustrates the use of the application of the CGM sensor (Figure S4 in Multimedia Appendix 1) and tethered insulin pump (Figures S5 and S6 in Multimedia Appendix 1). When the player's blood glucose levels reach a certain threshold (>5mmol/L), the advanced technology of using a CGM sensor and an insulin pump becomes available, unlocking new features and options in the game (see insulin pen and insulin pump icons in Figure 4).

Glucagon Administration

Glucagon administration is an essential treatment for severe hypoglycemia, particularly when a person with diabetes cannot consume food or fluids because they are in a hypoglycemic coma [12]. Another individual must administer the glucagon injection in such cases. The selected design closely resembles the real-life GlucaGen HypoKit (Novo Nordisk A/S) and visually represents the glucagon injection process (Figure S7B in Multimedia Appendix 1). The glucagon administration icon will be displayed when the avatar experiences hypoglycemia, indicating the appropriate action to take.

Measuring Ketoacidosis

Specifically, patients treated with continuous subcutaneous insulin infusion have a greater risk of developing diabetic ketoacidosis because there is no subcutaneous depot of insulin, and therefore ketoacidosis can occur much faster [1]. As the measurement of ketone bodies from urine using diagnostic strips is the most common and most accessible method in the home environment, the focus is on designing the scene using the home kit (Figure S7C in Multimedia Appendix 1). To simulate the level of ketoacidosis, the simplify model from the work by Fabietti et al [43] was applied (refer to the Ketoacidosis Module Implementation section in Multimedia Appendix 1).

Gold Usability Testing

The primary objective of the usability testing was to evaluate the technical feasibility of the advanced modules. Participants were recruited from a summer camp for children with diabetes. In the first phase, a general questionnaire was administered to assess the participants' knowledge about diabetes compensation. A total of 30 participants (n=12, 40% boys and n=18, 60% girls) aged between 7 and 15 years completed the questionnaire (refer to Table S9 in Multimedia Appendix 2). In the second phase, the 24 participants completed the tutorial (refer to the Onboarding and Tutorial subsection) together. Finally, the participants were divided into 3 groups to test the insulin pump and CGM sensor module (9/24, 37%), the ketoacidosis measurement module (7/24, 30%), and the glucagon administration module (8/24, 33%). Before testing, a pretesting questionnaire (refer to the Pretesting Questionnaire subsection under Gold Release in Multimedia Appendix 1) was administered. The usability testing session was conducted by a researcher, who explained the main goals to be achieved in each group in several steps.

A cohort of 9 participants (n=2, 22% boys and n=7, 78% girls) were enrolled to test the insulin pump and CGM sensor module. The study required the participants to complete several tasks, including applying the glucose sensor and insulin pump infusion set, calibrating the sensor, and spending a portion of the day using the sensor and the pump. The summary of the testing is provided in Table S10 in Multimedia Appendix 2. Overall, the insulin pump and CGM sensor module were rated positively for their realism; however, participants reported that certain complications were not accounted for in the game, such as disconnecting the infusion set from the reservoir, the possibility of the pump tubing becoming kinked and clogged, and adding insulin to the reservoir. During the pump assembly (refer to step 2 in Table S11 in Multimedia Appendix 2), participants had difficulty connecting the reservoir to the infusion set using tubing and squeezing the air out of the tubing by long pressing the button on the tank. In addition, confusion arose during the insulin infusion scene because participants needed clarification on how many units of insulin they should inject. Improvements in this scene could help better convey the necessary information to the player.

Eight participants were involved (n=4, 50% boys and n=4, 50% girls) in the testing of the glucagon administration module (refer to Table S12 in Multimedia Appendix 2). All participants were able to administer glucagon; however, sometimes the game took longer than expected due to minor bugs or nonintuitiveness, such as using a click gesture instead of a drag-and-drop gesture, which caused problems in cases of drawing the solution back into the syringe and then removing the syringe from the vial or removing the protective cover of the syringe. All participants praised the realistic handling of the game.

The ketoacidosis measurement module was successfully tested by all 7 participants (n=2, 28% boys and n=5, 72% girls), who found the scene good and did not suggest any improvements (refer to Table S13 in Multimedia Appendix 2). Only 1 (14%) of the 7 testers did not observe an increase in ketone bodies when insulin was not administered to the avatar. When asked about alternative methods of measuring ketones, most participants expressed openness to new options beyond diagnostic strips, especially for children accustomed to different measurement methods.

Discussion

Principal Findings

The main objective of this paper was to provide education mainly to children newly diagnosed with T1DM by using mobile technologies and gamification theory for educational purposes. The inclusion of gamification in diabetes management offers additional benefits compared to current methods, particularly in terms of improving patient motivation [4,5]. Given the widespread use of mobile technologies among young children, a mobile app is an ideal platform for introducing this serious topic to them entertainingly and engagingly. This approach involves combining various techniques, such as gamification theory [24] as well as topics from behavioral theory, such as tailoring [11], observational learning [29], decision-making practice [13], social and family support [30,31], and reward systems [32], to support patient motivation and compliance.

The game provides comprehensive education for children on various aspects of diabetes, including technical skills such as measuring blood glucose levels and administering insulin, as well as explanations of the nature of the disease. As a result, the game should be considered an additional resource for children newly diagnosed with T1DM to learn about the disease, along with books and brochures. The game's virtual avatar can also serve as a supportive friend for children unable to attend diabetes camps.

In addition, the game has the potential to raise awareness of diabetes among children without diabetes as well as parents [1]. By increasing knowledge of the disease and its symptoms, the game can aid in earlier diagnosis and treatment of diabetes. Initial feedback from children with T1DM and their parents on the game has been positive, with many expressing enjoyment and a willingness to participate in further testing.

Concerning the educational effect of the game, after a week of testing, there was a noticeable improvement in the educational outcomes (refer to Table 2). When compared to a previous study [16], it can be inferred that the improvement rate in educational outcomes is similar (33%). However, it is important to consider that the game in that study was tested for a year, while in this study, testing only lasted a week. In comparison, another study [18] reported an improvement rate in educational outcomes of 7% only. MyDiabetic aims to enhance traditional patient education formats (such as patient information booklets [44]) rather than replace them.

Recent literature reviews on gameful eHealth and mobile health tools have indicated that the most commonly used game elements are externally oriented, such as points and rewards [45,46]. The concept of avoiding excessive use of externally oriented motivational features has also been explored in theoretical works on designing engaging and gameful experiences [47]. Nevertheless, in a review of similar apps targeting gaming in T1DM (refer to Table S14 in Multimedia Appendix 2), authors found that the archetypal game elements of points, badges, and competitions were the primary approaches used [48]. MyDiabetic involves using rewards to purchase food, goods, and healthy supplements, which appeal to most children. The shopping aspect serves as an educational tool that prepares young children to increase self-efficacy in managing their T1DM, consequently reducing their dependence on their parents [5,17].

One of the most effective aspects of the co-design process in achieving design goals was the participation of clinicians, patients, family caregivers, developers, and game designers. Participants brought their unique expertise to the table and contributed to the design process; for example, health care researchers provided knowledge and feedback on diabetes compensation, while the developers and designers worked on making the tool user-friendly, and the external game designer suggested more interactive and immersive solutions. Research indicates that collaborative, team-based approaches are recommended for developing mobile health interventions [49]. The strategic coordination of stakeholder involvement at each stage of development was a key benefit of this design approach.

Although there is still no consensus on the optimal combination of game mechanics for serious games [34], it is acknowledged that MyDiabetic currently leans too heavily toward education, leading to lower retention rates than anticipated. Drawing from the insights gained during the second and third rounds of the iterative usability studies, it is proposed to redesign the game by incorporating multiple storylines, such as school experiences, dating, and career development, to enhance the entertainment factor and introduce new topics as the avatar ages. The existing educational components will be seamlessly integrated into the game mechanics to improve overall gameplay and increase engagement. This new direction should also provide an easier transfer of knowledge and skills from this serious game to real-world situations [34].

Limitations

Initially, all participants involved in the study voluntarily opted to participate or were contacted by the project team or collaborating institutions. Consequently, the participating groups may have had higher motivation, resourcefulness, and better chronic illness management, thus creating a potential bias that may only represent a portion of the user group. Nonetheless, this is a common limitation in this type of research.

While the sample sizes were indeed modest in the alpha, beta, and gold testing phases (12, 6, and 24 participants, respectively), the significance of the results lies in the qualitative insights and detailed observations that were gathered through iterative usability testing. This process enabled deep delving into the experiences and perceptions of each participant, yielding rich and meaningful data on the game's usability and educational impact. Furthermore, the results that were obtained from the limited numbers of participants were consistent and provided actionable insights for further refinement and development of this serious game.

The design outcomes were influenced by the stakeholders involved in the project, who had the potential to both positively impact and limit the final design. One example is the nutrition nurse who facilitated the design activities but had limited design experience and had to learn co-design processes as the project progressed.

The MyDiabetic app was developed and evaluated solely by Czech users, although an English version is now accessible. There may still be cultural discrepancies that have not been fully addressed. The project team constantly receives feedback and makes necessary adjustments to improve the app's usability and relevance.

It is important to note that this paper focuses solely on the technical development of the game and user acceptance testing. The clinical validation of the game, using metabolic control markers such as mean glycated hemoglobin levels, is beyond the scope of this paper. While it is important to evaluate the effectiveness of the game in improving diabetes management, this will be the subject of future research.

Comparison With Prior Work

After examining the survey presented in Table S14 in Multimedia Appendix 2 [13,16,18,19,21,24-31,33,42,43,50-65], it was observed that all games, except Packy & Marlon [16], lacked clinical validation. Furthermore, assessing adherence was not a focus in any of the games. In addition, of the 18 games, 3 (17%) were previously available on Nintendo or mobile game stores, but currently only 2 (11%; MyDiabetic and Jerry the Bear) are available for download on mobile game stores. Most games were developed as prototypes or concept studies as part of academic projects, with only a few aimed at preschool children [50-52]. Only 1 (6%) of the 18 games addresses both basic (carbohydrate counting, blood glucose–level measurement, and insulin administration) and advanced (CGM and insulin pump) educational self-management skills [50]. Nevertheless, no game tackles the complexity of diabetes management from such a wide perspective as MyDiabetic, in which unique features such as glucagon administration and ketoacidosis modules are included.

One of the challenges that serious game designers face is the *uncanny principle*, also known as the *uncanny valley*. This phenomenon occurs when a representation of a human or an animal looks and behaves almost but not exactly like the real thing [66]. The result is a feeling of eeriness, discomfort, or even revulsion in the observer. The *uncanny principle* can affect the learner's immersion, engagement, and emotional connection with the serious game. The *uncanny valley* can be useful for creating more effective and engaging serious games. By carefully navigating the *valley* and finding the right balance between realism and abstraction, game designers can create experiences that are both educational and emotionally compelling [67]. MyDiabetic aimed to strike this balance by using cartoonish avatar designs and environments, while carefully incorporating technology features that closely resemble real-world settings, setting it apart from most other games [16,50].

Future Directions

To enhance the game's realism, it would be beneficial to incorporate real-time elements. This would involve the avatar aging as time passes In addition to managing diabetes symptoms, the avatar would also engage in typical daily activities such as attending school/work, participating in outdoor sports, visiting friends, and partaking in entertainment (Figures S8A and S8B in Multimedia Appendix 1). In addition, chronic retinopathy or diabetes foot problems would also be demonstrated (Figure S8C in Multimedia Appendix 1). By including these elements, the game could provide a more comprehensive experience for users and better prepare them for managing their diabetes in real-life situations.

Including social support in the app could be an additional valuable feature [68]. Games that allow users to connect with others in a community-type setting or through social networks could be particularly beneficial because research suggests that increased social support is linked to improved self-efficacy practices and better clinical outcomes in children with diabetes. A review of studies also found that social media were commonly used to facilitate self-care in patients and caregivers, with 77.1% of the identified studies reporting such use [69].

Conclusions

This paper has shown how participants involved in co-design activities played a creative and productive role in shaping the content and design of the MyDiabetic app.

The main objective of creating a serious educational game was to captivate children with T1DM and also make it accessible to those who are interested in learning about the disease. On the basis of the testing described earlier, the educational goal was achieved because significant enhancements in children's understanding were evident after a mere week of gameplay. The game was well-received by the participants, who expressed a willingness to recommend it to their friends or siblings who are not affected by diabetes but are curious about the disease and would like to understand it better.

To summarize, this research suggests that the target audience for this game is children aged between 5 and 12 years, with those in the 8- to 12-year range being able to fully engage with and benefit from all the features the game offers.

Acknowledgments

The author expresses his gratitude to his students who designed and implemented MyDiabetic in their final projects (2016-2022): Veronika Cernohorska for the main concept, Jana Kejvalova for carrying out the user testing, Lukas Rubes for integrating the onboarding, Dusan Jencik for designing the levels, Natalie Zubkova for integrating the educational library, Victoria Eykhmann for the glucagon administration module implementation, Eva Uliarikova for implementing the measurement of ketoacidosis, Petra Salficka for storyline development, and Benjamin Hejl for continuous glucose measurement and insulin pump deployment as well as project management. The author's gratitude also goes to the clinicians who provided invaluable insight and support for the game during the design phase—Ludmila Brazdova (Brothers of Mercy Hospital) and Jaroslav Skvor (Masaryk Hospital) contributed significantly to the first iteration by offering insights into common T1DM principles. In addition, Jaroslav Skvor provided input regarding the role of physical exercise. In the second iteration, Katerina Stechova (University Hospital in Motol) shared valuable comments concerning the blood glucose simulator and adherence principles in children. For the third iteration, Milos Mraz (Institute for Clinical and Experimental Medicine) contributed his expertise on the CGM module and insulin pump, while Zdenek Sumnik (University Hospital in Motol) offered critical feedback on gamification principles. The author also thanks to patients' organizations for facilitating access to children with type 1 diabetes mellitus during the participatory design and feasibility studies—Vaclav Letocha (Association of Parents and Friends of Diabetic Children), Sarka Nosalkova (Center for Children with Diabetes), and Zdenka Stankova (Diacel)—and finally to game designers Michal Berlinger (Amanita Design) and Lucie Workova. The study was supported by the Research Centre for Informatics (CZ.02.1.01/0.0/16_019/0000765) and by the biomedical data acquisition, processing, and visualization (SGS22/165/OHK3/3T/13) grant. The author used generative artificial intelligence (ChatGPT, OpenAI) for proofreading this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional theory background, questionnaires, and details of implementation.
[PDF File (Adobe PDF File), 12529 KB - games_v12i1e49478_app1.pdf]

Multimedia Appendix 2
Results of the feasibility testing.
[XLSX File (Microsoft Excel File), 154 KB - games_v12i1e49478_app2.xlsx]

Multimedia Appendix 3
PowerPoint (Microsoft Corp) presentation showing key aspects of MyDiabetic and related material.
[PDF File (Adobe PDF File), 14994 KB - games_v12i1e49478_app3.pdf]

Multimedia Appendix 4
MyDiabetic video.
[MP4 File (MP4 Video), 19844 KB - games_v12i1e49478_app4.mp4]

References

1. Chiang JL, Maahs DM, Garvey KC, Hood KK, Laffel LM, Weinzimer SA, et al. Type 1 diabetes in children and adolescents: a position statement by the American Diabetes Association. Diabetes Care 2018 Sep 09;41(9):2026-2044 [FREE Full text] [doi: 10.2337/dci18-0023] [Medline: 30093549]
2. Bayram S, Kızıltan G, Akın O. Effect of adherence to carbohydrate counting on metabolic control in children and adolescents with type 1 diabetes mellitus. Ann Pediatr Endocrinol Metab 2020 Sep;25(3):156-162 [FREE Full text] [doi: 10.6065/apem.1938192.096] [Medline: 32871653]
3. Lyons SK, Ebekozien O, Garrity A, Buckingham D, Odugbesan O, Thomas S, et al. Increasing insulin pump use among 12- to 26-year-olds with type 1 diabetes: results from the T1D exchange quality improvement collaborative. Clin Diabetes 2021 Jul;39(3):272-277 [FREE Full text] [doi: 10.2337/cd21-0027] [Medline: 34421202]
4. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012 May 08;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
5. Yang Q, Van Stee SK. The comparative effectiveness of mobile phone interventions in improving health outcomes: meta-analytic review. JMIR Mhealth Uhealth 2019 Apr 03;7(4):e11244 [FREE Full text] [doi: 10.2196/11244] [Medline: 30942695]
6. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: a systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018 Aug;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
7. Novak D, Uller M, Rousseaux S, Mraz M, Smrz J, Stepankova O, et al. Diabetes management in OLDES project. Annu Int Conf IEEE Eng Med Biol Soc 2009;2009:7228-7231. [doi: 10.1109/IEMBS.2009.5335256] [Medline: 19965284]
8. Burda V, Mráz M, Schneider J, Novák D. Managing diabetes using Mobiab: long-term case study of the impact of a mobile app on self-management. JMIR Diabetes 2022 Apr 20;7(2):e36675 [FREE Full text] [doi: 10.2196/36675] [Medline: 35442201]
9. Alaslawi H, Berrou I, Al Hamid A, Alhuwail D, Aslanpour Z. Diabetes self-management apps: systematic review of adoption determinants and future research agenda. JMIR Diabetes 2022 Jul 28;7(3):e28153 [FREE Full text] [doi: 10.2196/28153] [Medline: 35900826]
10. Shan R, Sarkar S, Martin SS. Digital health technology and mobile devices for the management of diabetes mellitus: state of the art. Diabetologia 2019 Jun;62(6):877-887. [doi: 10.1007/s00125-019-4864-7] [Medline: 30963188]
11. Desrochers HR, Schultz AT, Laffel LM. Use of diabetes technology in children: role of structured education for young people with diabetes and families. Endocrinol Metab Clin North Am 2020 Mar;49(1):19-35 [FREE Full text] [doi: 10.1016/j.ecl.2019.11.001] [Medline: 31980118]
12. Lindholm Olinder A, DeAbreu M, Greene S, Haugstvedt A, Lange K, Majaliwa ES, et al. ISPAD Clinical Practice Consensus Guidelines 2022: diabetes education in children and adolescents. Pediatr Diabetes 2022 Dec 20;23(8):1229-1242 [FREE Full text] [doi: 10.1111/pedi.13418] [Medline: 36120721]
13. Novak D, Tulu B, Brendryen H. Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice. Hershey, PA: IGI Global; 2015.
14. Lesiak AJ, Griswold JC. Blood sugar balance: a glucose metabolism web game for diabetes education. Am Biol Teach 2023 May 01;85(5):270-277. [doi: 10.1525/abt.2023.85.5.270]
15. Martos-Cabrera MB, Membrive-Jiménez MJ, Suleiman-Martos N, Mota-Romero E, Cañadas-De la Fuente GA, Gómez-Urquiza JL, et al. Games and health education for diabetes control: a systematic review with meta-analysis. Healthcare (Basel) 2020 Oct 14;8(4):399 [FREE Full text] [doi: 10.3390/healthcare8040399] [Medline: 33066372]
16. Lieberman DA. Management of chronic pediatric diseases with interactive health games: theory and research findings. J Ambul Care Manage 2001 Jan;24(1):26-38. [doi: 10.1097/00004479-200101000-00004] [Medline: 11189794]
17. Fuchslocher A, Niesenhaus J, Krämer N. Serious games for health: an empirical study of the game “Balance” for teenagers with diabetes mellitus. Entertain Comput 2011;2(2):97-101. [doi: 10.1016/j.entcom.2010.12.001]
18. Baghaei N, Nandigam D, Casey J, Direito A, Maddison R. Diabetic Mario: designing and evaluating mobile games for diabetes education. Games Health J 2016 Aug;5(4):270-278. [doi: 10.1089/g4h.2015.0038] [Medline: 27304882]

19. Kamel Boulos MN, Gammon S, Dixon MC, MacRury SM, Fergusson MJ, Miranda Rodrigues F, et al. Digital games for type 1 and type 2 diabetes: underpinning theory with three illustrative examples. JMIR Serious Games 2015 Mar 18;3(1):e3 [FREE Full text] [doi: 10.2196/games.3930] [Medline: 25791276]
20. Martinez M, Park SB, Maison I, Mody V, Soh LS, Parihar HS. iOS appstore-based phone apps for diabetes management: potential for use in medication adherence. JMIR Diabetes 2017 Jul 11;2(2):e12 [FREE Full text] [doi: 10.2196/diabetes.6468] [Medline: 30291096]
21. Jimenez G, Lum E, Car J. Examining diabetes management apps recommended from a Google search: content analysis. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11848 [FREE Full text] [doi: 10.2196/11848] [Medline: 30303485]
22. Alsalman D, Bu Ali ZM, Alnosaier ZF, Alotaibi NA, Alanzi TM. Gamification for diabetes type 1 management: a review of the features of free apps in Google play and app stores. J Multidiscip Healthc 2020 May 14;13:425-432 [FREE Full text] [doi: 10.2147/JMDH.S249664] [Medline: 32523349]
23. Kildea J, Battista J, Cabral B, Hendren L, Herrera D, Hijal T, et al. Design and development of a person-centered patient portal using participatory stakeholder co-design. J Med Internet Res 2019 Feb 11;21(2):e11371 [FREE Full text] [doi: 10.2196/11371] [Medline: 30741643]
24. Sweetser P, Wyeth P. GameFlow: a model for evaluating player enjoyment in games. Comput Entertain 2005 Jul 01;3(3):3. [doi: 10.1145/1077246.1077253]
25. Hunicke R, LeBlanc M, Zubek R. MDA: a formal approach to game design and game research. In: Proceedings of the AAAI Workshop on Challenges in Game AI. 2004 Presented at: AAAI-04 Conference; July 25-29, 2004; San Jose, CA.
26. Ganesh A, Ndulue C, Orji R. Tailoring a persuasive game to promote secure smartphone behaviour. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany. [doi: 10.1145/3544548.3581038]
27. Thompson D. Designing serious video games for health behavior change: current status and future directions. J Diabetes Sci Technol 2012 Jul 01;6(4):807-811 [FREE Full text] [doi: 10.1177/193229681200600411] [Medline: 22920806]
28. Yari A, Hassani L, Ziapour A, Toghroli R, Mohammadkhah F, Yazdi F. Prediction of self-care behaviors among diabetic patients based on social cognitive theory. J Public Health Res 2023 Feb 01;12(1):22799036221150330 [FREE Full text] [doi: 10.1177/22799036221150330] [Medline: 36755578]
29. Thompson D, Baranowski T, Buday R. Conceptual model for the design of a serious video game promoting self-management among youth with type 1 diabetes. J Diabetes Sci Technol 2010 May 01;4(3):744-749 [FREE Full text] [doi: 10.1177/193229681000400331] [Medline: 20513343]
30. Brorsson AL, Leksell J, Viklund G, Lindholm Olinder A. A multicentre randomized controlled trial of an empowerment-inspired intervention for adolescents starting continuous subcutaneous insulin infusion--a study protocol. BMC Pediatr 2013 Dec 20;13:212 [FREE Full text] [doi: 10.1186/1471-2431-13-212] [Medline: 24354899]
31. Almeida AC, Leandro ME, Pereira MG. Adherence and glycemic control in adolescents with type 1 diabetes: the moderating role of age, gender, and family support. J Clin Psychol Med Settings 2020 Jun;27(2):247-255. [doi: 10.1007/s10880-019-09662-y] [Medline: 31538310]
32. Miranda JJ, Lazo-Porras M, Bernabe-Ortiz A, Pesantes MA, Diez-Canseco F, Cornejo SD, et al. The effect of individual and mixed rewards on diabetes management: a feasibility randomized controlled trial. Wellcome Open Res 2019 Feb 05;3:139 [FREE Full text] [doi: 10.12688/wellcomeopenres.14824.3] [Medline: 30662958]
33. D'Aprile G, Ligorio MB, Ulloa Severino A. How serious games for health enhance empowerment-related patient education: the impact of gender. Technol Knowl Learn 2017 Dec 20;24:325-340. [doi: 10.1007/s10758-017-9344-x]
34. Nørlev J, Sondrup K, Derosche C, Hejlesen O, Hangaard S. Game mechanisms in serious games that teach children with type 1 diabetes how to self-manage: a systematic scoping review. J Diabetes Sci Technol 2022 Sep;16(5):1253-1269 [FREE Full text] [doi: 10.1177/19322968211018236] [Medline: 34024156]
35. My diabetic homepage. My Diabetic. URL: https://www.my-diabetic.com/ [accessed 2024-01-05]
36. Shore J, Warden S. The Art of Agile Development. Sebastopol, CA: O'Reilly Media; 2007.
37. Unity real-time development platform. Unity. URL: https://unity.com [accessed 2023-05-11]
38. Home of the Blender project. Blender. URL: https://www.blender.org/ [accessed 2023-05-11]
39. Let's build from here. GitHub. URL: https://github.com/ [accessed 2023-05-11]
40. Manage your team's projects from anywhere. Trello. URL: https://trello.com/ [accessed 2023-05-11]
41. Tuah NM, Yoag A, Ahmedy F. MyDiabetes—the gamified application for diabetes self-management and care. Computers 2021 Apr 13;10(4):50. [doi: 10.3390/computers10040050]
42. Lehmann ED. Why people download the freeware AIDA v4.3a diabetes software program: a proof-of-concept semi-automated analysis. Diabetes Technol Ther 2003;5(3):477-490. [doi: 10.1089/152091503765692027] [Medline: 12828835]
43. Fabietti PG, Canonico V, Orsini Federici M, Sarti E, Massi Benedetti M. Model based study on monitoring ketone bodies to improve safety in intensive insulin therapy. Int J Artif Organs 2006 Jun;29(6):596-601. [doi: 10.1177/039139880602900609] [Medline: 16841289]
44. Lebl J, Průhová Š, Šumník · Z. Abeceda Diabetu. Prague, Czech Republic: Maxdorf; 2018.
45. Xie H. A scoping review of gamification for mental health in children: uncovering its key features and impact. Arch Psychiatr Nurs 2022 Dec;41:132-143. [doi: 10.1016/j.apnu.2022.07.003] [Medline: 36428041]

46. Suleiman-Martos N, García-Lara RA, Martos-Cabrera MB, Albendín-García L, Romero-Béjar JL, Cañadas-De la Fuente GA, et al. Gamification for the improvement of diet, nutritional habits, and body composition in children and adolescents: a systematic review and meta-analysis. Nutrients 2021 Jul 20;13(7):2478 [FREE Full text] [doi: 10.3390/nu13072478] [Medline: 34371989]
47. Kim TW, Werbach K. More than just a game: ethical issues in gamification. Ethics Inf Technol 2016 May 12;18:157-173. [doi: 10.1007/s10676-016-9401-5]
48. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
49. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
50. Ullman D, Phillips E, Aladia S, Haas P, Fowler HS, Iqbal IS, et al. Evaluating psychosocial support provided by an augmented reality device for children with type 1 diabetes. Proc Int Symp Hum Factors Ergon Health Care 2021;10(1):126-130. [doi: 10.1177/2327857921101117]
51. Ebrahimpour F, Najafi M, Sadeghi N. The design and development of a computer game on insulin injection. Electron Physician 2014 May 10;6(2):845-855 [FREE Full text] [doi: 10.14661/2014.845-855] [Medline: 25763157]
52. Calle-Bustos AM, Juan MC, García-García I, Abad F. An augmented reality game to support therapeutic education for children with diabetes. PLoS One 2017 Sep 28;12(9):e0184645 [FREE Full text] [doi: 10.1371/journal.pone.0184645] [Medline: 28957355]
53. Glasemann M, Kanstrup AM, Ryberg T. Making chocolate-covered broccoli: designing a mobile learning game about food for young people with diabetes. In: Proceedings of the 8th ACM Conference on Designing Interactive Systems. 2010 Presented at: DIS '10; August 16-20, 2010; Aarhus, Denmark. [doi: 10.1145/1858171.1858219]
54. Brox E, Hirche J, Evertsen G, Yliräisänen-Seppänen P, Bomark P. User centric social diabetes game design for children. In: Proceeding of the 16th International Academic MindTrek Conference. 2012 Presented at: MindTrek '12; October 3-5, 2012; Tampere, Finland. [doi: 10.1145/2393132.2393196]
55. Peters R, Oleari E, Sardu F, Neerincx MA. Usability of the PAL objectives dashboard for children's diabetes self-management education. In: Proceedings of the 5th International Conference on e-Society, e-Learning and e-Technologies. 2019 Presented at: ICSLT '19; January 10-12, 2019; Vienna, Austria. [doi: 10.1145/3312714.3312736]
56. Rønningen IC, Årsand E, Hartvigsen G. Exploring in-game reward mechanisms in diaquarium – a serious game for children with type 1 diabetes. In: Proceedings of the 6th International Work-Conference on Bioinformatics and Biomedical Engineering. 2018 Presented at: IWBBIO 2018; April 25-27, 2018; Granada, Spain. [doi: 10.1007/978-3-319-78759-6_40]
57. Pesare E, Roselli T, Rossano V. Intelligent Agent and Virtual Game to support education in e-health. In: Proceedings of the 21st International Conference on Distributed Multimedia Systems. 2015 Presented at: DMS 2015; August 31-September 2, 2015; Vancouver, Canada. [doi: 10.18293/DMS2015-35]
58. Lauritzen J, Årsand E, Horsch A, Fernandez-Luque L, Chomutare T, Bellika JG, et al. Social media and games as self-management tools for children and adolescents with type 1 diabetes mellitus. In: Proceedings of the International Conference on Health Informatics - Volume 0BIOSTEC. 2012 Presented at: BIOSTEC 2012; February 1-4, 2012; Algarve, Portugal. [doi: 10.5220/0003874104590466]
59. Bomark P, Evertsen G, Brox E, Hirche J, Yliräisänen-Seppänen P. A prototype social learning platform for children with diabetes type 1. In: Proceeding of the 16th International Academic MindTrek Conference. 2012 Presented at: MindTrek '12; October 3-5, 2012; Tampere, Finland. [doi: 10.1145/2393132.2393174]
60. Moosa A, Al-Maadeed N, AlJa'am JM. A simple health-based game for children. In: Proceedings of the International Conference on Computer and Applications. 2018 Presented at: ICCA 2018; August 25-26, 2018; Beirut, Lebanon. [doi: 10.1109/COMAPP.2018.8460213]
61. Sparapani VC, Fels S, Kamal N, Nascimento LC. Conceptual framework for designing video games for children with type 1 diabetes. Rev Lat Am Enfermagem 2019 Mar 18;27:e3090 [FREE Full text] [doi: 10.1590/1518-8345.2764.3090] [Medline: 30916223]
62. Ling C, Seetharaman S, Mirza L. Roles of serious game in diabetes patient education. Simul Gaming 2022;53(5):513-537. [doi: 10.1177/10468781221120686]
63. Kudva YC, Ahmann AJ, Bergenstal RM, Gavin JR3, Kruger DF, Midyett LK, et al. Approach to using trend arrows in the FreeStyle libre flash glucose monitoring systems in adults. J Endocr Soc 2018 Dec 01;2(12):1320-1337 [FREE Full text] [doi: 10.1210/js.2018-00294] [Medline: 30474069]
64. Aleppo G, Laffel LM, Ahmann AJ, Hirsch IB, Kruger DF, Peters A, et al. A practical approach to using trend arrows on the Dexcom G5 CGM system for the management of adults with diabetes. J Endocr Soc 2017 Dec 01;1(12):1445-1460 [FREE Full text] [doi: 10.1210/js.2017-00388] [Medline: 29344577]
65. Barnett CR, Barnett YA. Ketone bodies. In: Caballero B, editor. Encyclopedia of Food Sciences and Nutrition. San Diego, CA: Academic Press; 2003.

66. Sierra Rativa A, Postma M, van Zaanen M. The uncanny valley of a virtual animal. Comput Anim Virtual Worlds 2022 Mar 16;33(2):e2043. [doi: 10.1002/cav.2043]
67. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
68. Elnaggar A, Ta Park V, Lee SJ, Bender M, Siegmund LA, Park LG. Patients' use of social media for diabetes self-care: systematic review. J Med Internet Res 2020 Apr 24;22(4):e14209 [FREE Full text] [doi: 10.2196/14209] [Medline: 32329745]
69. Muangsrinoon S, Boonbrahm P. Game elements from literature review of gamification in healthcare context. J Technol Sci Educ 2019 Feb;9(1):20. [doi: 10.3926/jotse.556]

Abbreviations

BU: bread unit
CGM: continuous glucose monitoring
T1DM: type 1 diabetes mellitus

Edited by G Eysenbach, T Leung; submitted 02.06.23; peer-reviewed by D Cuesta, F Babic, V Sparapani; comments to author 29.06.23; revised version received 24.08.23; accepted 13.02.24; published 07.05.24.

Please cite as:
Novak D
A Serious Game (MyDiabetic) to Support Children's Education in Type 1 Diabetes Mellitus: Iterative Participatory Co-Design and Feasibility Study
JMIR Serious Games 2024;12:e49478
URL: https://games.jmir.org/2024/1/e49478
doi:10.2196/49478
PMID:38713496

Review

Virtual Reality Applications for the Implementation of Domestic Respiratory Rehabilitation Programs for Patients With Long COVID and Post-COVID Condition: Scoping Review

Katharina Dalko[1], MA; Hlynur Andri Elsuson[2], MSc; Ivonne Kalter[3], MA; Max Zilezinski[4], MSc; Sebastian Hofstetter[1,3], MA; Dietrich Stoevesandt[1], MD; Denny Paulicke[3,5*], Prof Dr; Patrick Jahn[3*], Prof Dr

[1]Dorothea Erxleben Lernzentrum, Medical Faculty Halle, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[2]LichterSchatten Therapiezentrum GmbH, Berlin, Germany

[3]Health Service Research Working Group, Acute Care, Department of Internal Medicine, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[4]Institute for Clinical Nursing Science, Charité - Universitätsmedizin Berlin, Berlin, Germany

[5]Department of Medical Education, Akkon University of Human Sciences, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Katharina Dalko, MA
Dorothea Erxleben Lernzentrum
Medical Faculty Halle
Martin-Luther-University Halle-Wittenberg
Magdeburger Straße 12
Halle (Saale), 06112
Germany
Phone: 49 345 557 5493
Email: katharina.dalko@medizin.uni-halle.de

Abstract

Background: Due to a high number of patients affected by long COVID or post-COVID condition, an essential step to address the long-term effects of COVID-19 lies in the development and implementation of flexible and accessible rehabilitation programs. Virtual reality (VR) technologies offer the potential to support traditional therapies with individualized at-home programs.

Objective: This study aims to provide an overview of existing scientific evidence on the development and implementation of VR-assisted respiratory rehabilitation programs for patients with long COVID and post-COVID condition and to synthesize the results.

Methods: We conducted a scoping review of studies from 6 databases. PubMed, CINAHL, Cochrane, ScienceDirect, Web of Science Social Sciences Citation Index, and PEDro were searched using an exploratory search strategy. The search, which was last updated in February 2024, included peer-reviewed studies on immersive VR applications providing respiratory rehabilitation programs for patients with chronic obstructive pulmonary disease and long COVID or post-COVID condition. Exclusion criteria were studies in clinical or inpatient settings, telemedicine, nonimmersive VR applications, and gray literature. Nine publications were included in this review. Findings were extracted and summarized from the studies according to the JBI (Joanna Briggs Institute) method and thematically categorized. Topics covered were study characteristics, physiotherapeutic concept, clinical parameters, as well as usability and acceptability.

Results: The 9 publications included in the qualitative analysis were published in 2019-2023. Eight empirical studies were included: 4 followed a mixed methods design, 3 were qualitative studies, and 1 followed a quantitative method. One scoping review was included in the data analyses. Four of the included studies were on patients with chronic obstructive pulmonary disease. The 9 studies demonstrated that VR-supported respiratory rehabilitation programs result in positive initial outcomes in terms of physical as well as psychological parameters. Particularly noteworthy was the increased motivation and compliance of patients. However, adverse effects and lack of usability are the barriers to the implementation of this innovative approach.

Conclusions: Overall, VR is a promising technology for the implementation of individualized and flexible respiratory rehabilitation programs for patients with long COVID and post-COVID condition. Nevertheless, corresponding approaches are still under development and need to be more closely adapted to the needs of users. Further, the evidence was limited to pilot studies or a small number of patients, and no randomized controlled trials or long-term studies were part of the study selection. The included studies were performed by 4 groups of researchers: 3 from Europe and 1 from the United States.

(JMIR Serious Games 2024;12:e52309) doi:10.2196/52309

KEYWORDS

long COVID; post-COVID; rehabilitation; virtual reality; implementation; respiratory; respiratory rehabilitation; scoping review; development; accessibility; support; physical; psychological; motivation; compliance; usability; COVID-19; COVID

Introduction

Background

In the wake of the COVID-19 pandemic, the rehabilitation of patients in the postacute phase of this disease is an important measure to address the long-term effects since a significant number of patients experience the condition commonly known as long or post-COVID [1]. This condition is characterized by symptoms that persist or develop after the acute phase of infection—starting from the fourth week after infection—and that cannot be explained by an alternative diagnosis [2,3]. The number of affected patients varied in the studies depending on the methodology, symptoms, as well as the population included in the analysis. Although studies provide heterogeneous results and case numbers, they show that a significant number of patients are affected by persistent symptoms after a COVID-19 infection. Patients who report an impact on everyday functioning up to 3 months after testing negative account for 10%-50% of the study participants [4]. According to a study by Peter et al [5] conducted in 2020 and 2021, including 11,710 patients from Germany, 28.5% of the patients reported persistent symptoms for 6-12 months after infection with COVID-19. That study further estimated that at least 6.5% of the adult patients in the general population who had recovered from COVID-19 infection were affected by long-term symptoms such as fatigue, dyspnea, neurocognitive impairments, and chest pain [5]. As per the World Health Organization, in the European region approximately 20% of the patients developed symptoms mentioned above continuing for at least 3 months after recovery according to a meta-analysis [6]. In addition, psychological symptoms such as anxiety and stress can have a negative impact on the quality of life caused by the abovementioned long-term effects [4,7].

After the treatment of acute symptoms is completed, patients need postacute rehabilitation, where physical therapy plays an important role in the treatment of lung-specific symptoms. For the best treatment possible, outpatient programs as well as solutions for the implementation of therapy programs in the home environment have to be established [8,9]. Pulmonary therapy approaches for chronic diseases such as chronic obstructive pulmonary disease (COPD) designed to normalize respiratory function have been well-established and are guiding the development of therapies for patients with long COVID or post-COVID condition. Therapy approaches include mobilization exercises, endurance, as well as strength training [10-14]. Since many patients experience psychological symptoms from the effects of impaired respiratory function, it is necessary to guide, counsel, and train patients in the use of appropriate strategies and coping skills when acute respiratory distress occurs [11].

Although the number of people affected by long and post-COVID symptoms remains high even as the pandemic situation continues to ease, there is still insufficient knowledge about the disease and a shortage of specialists and therapy programs [4,15]. In addition, the physical limitations of patients further impact their access to traditional physical therapy services. One way to address the shortage of adequate programs is to develop digital therapy solutions and assistive devices that are applicable in a home setting and can be individually applied without the constant supervision of specialist staff.

Initial findings suggest that digital approaches enable a more accessible implementation of therapies in the home environment and, at the same time, can contribute to increased motivation and adherence to therapy on the part of patients [16]. Previously established digital applications, for example, for patients with COPD have led to an increase in the quality of life, particularly with regard to emotional control and reduction in fatigue and dyspnea [17]. Virtual reality (VR) technologies are a solution to implement individual and flexible physical therapies in virtual space through the virtual representation of therapeutic measures and therapy situations. Approaches to integrate immersive VR applications already exist in various areas of rehabilitation as well as in psychotherapy, for example, to alleviate respiratory symptoms [18,19] and reduce anxiety and stress [20]. However, the implementation of respiratory therapy approaches for the home environment and especially programs for the target group of patients with long COVID/post-COVID condition are still under development. Further, a comprehensive review of the initial evidence on the development and implementation of appropriate VR applications does not yet exist.

Objectives

The aim of our literature review was to (1) obtain an overview of the findings in international research regarding VR-based respiratory rehabilitation programs for patients with long COVID/post-COVID condition and (2) obtain criteria for the development and implementation of respective VR applications for the home environment. Our scoping review addresses the research question: what scientific evidence exists on the development and implementation of VR-assisted rehabilitation programs for patients with long COVID/post-COVID condition that are implementable in a home setting? The selection and analysis of the studies were based on the following subquestions:

1. Which guidelines exist for the design of VR respiratory rehabilitation programs?
2. What are the enabling aspects? What are the barriers to implementation?

3. What clinical outcomes have been reported?

Methods

Study Design

The methodological approach of the JBI (Joanna Briggs Institute) method according to von Elm et al [21] was adopted as the basis of this scoping review to give a broad overview on the existing findings and identify established criteria in international research for the implementation of VR-assisted rehabilitation programs for patients with long COVID/post-COVID condition. Since only a small number of studies was expected regarding the patient group and the focus was on respiratory rehabilitation programs, conditions with a comparable symptomatic spectrum—such as patients with COPD—were also included. Results will be reported using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines (Multimedia Appendix 1). In preparation for this, a review protocol was developed but not published or registered.

Search Strategy

A sensitive database search was conducted using the databases PubMed, CINAHL, Cochrane, ScienceDirect, Web of Science Social Sciences Citation Index, and PEDro. The search was last updated in February 2024 in order to incorporate newly released studies. According to the search components—population, concept, and context—search terms were applied using Boolean operators, truncations, and proximity operators (see Textbox 1 and Multimedia Appendix 2).

Textbox 1. Search string using the example of the search in PubMed.

(((covid-19[MeSH Terms]) OR (respiratory*[Title/Abstract])) OR (pulmonary*[Title/Abstract])) AND (Rehabilitation[Title/Abstract]) AND (VR[Title/Abstract] OR virtual reality[Title/Abstract])

The identified studies were merged in the web-based tool “rayyan” [22] and screened by titles, abstracts, and full texts in regard to the research question, which was conducted independently by 3 researchers. Additionally, the reference lists of all the publications included in the full text screening were searched for further evidence. The inclusion and exclusion criteria for the studies were decided by team consensus (KD, IK, and HAE).

Study Selection

All types of studies published in the period between 2012 and 2023 that were available in English or German and provided with an abstract were included. The following inclusion and exclusion criteria were applied (see Textbox 2).

Textbox 2. Inclusion and exclusion criteria for the studies in this review.

Inclusion criteria

- Studies on immersive virtual reality applications providing respiratory rehabilitation programs, including breathing exercises, physical training, education, and programs introducing psychological counseling such as stress reduction for the home environment
- Studies including patients with long COVID, post-COVID condition, or chronic obstructive pulmonary disease
- Peer-reviewed empirical studies
- Mixed methods, qualitative, and quantitative studies

Exclusion criteria

- Clinical inpatient setting
- Telemedicine
- Virtual reality applications that are nonimmersive (applications for personal computers, augmented reality, etc)
- Gray literature (conference papers, opinion papers, study protocols)

Data Extraction and Synthesis

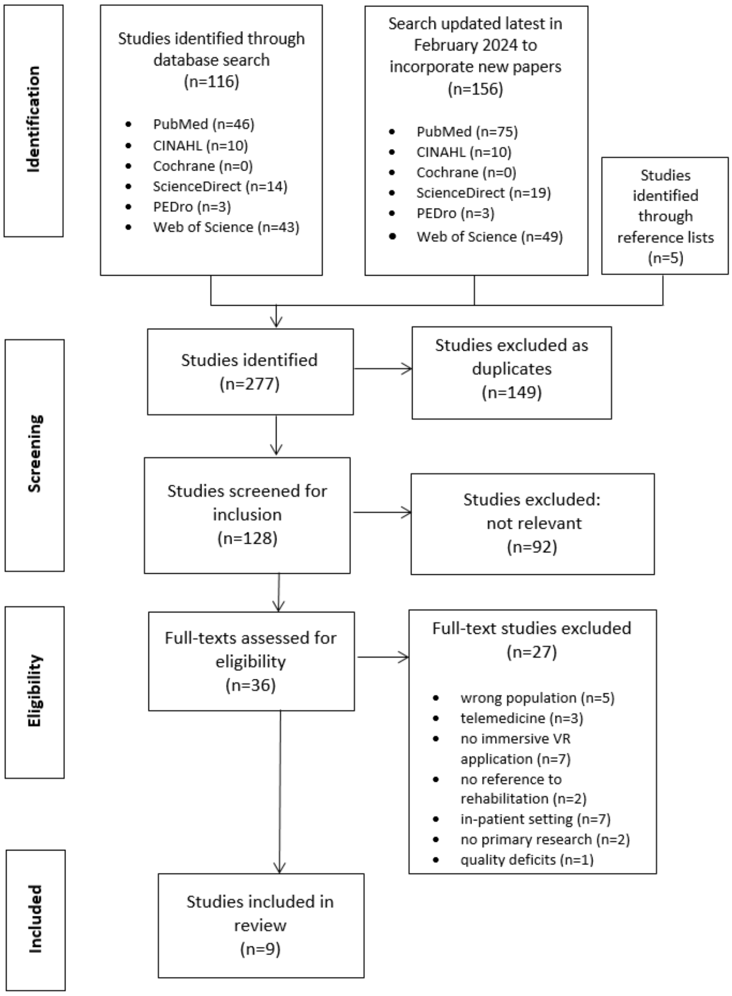
The characteristics of the identified studies were mapped in a preconsented data form (KD) and summarized narratively. Three researchers (KD, IK, and HAE) derived evidence on the development and implementation as well as the clinical outcomes of the studies concerning VR-based rehabilitation programs from the literature included in this analysis. The aspects identified were then categorized thematically, while study results covered the areas of study characteristics, physiotherapeutic concept, and outcomes in terms of clinical parameters as well as usability and acceptability aspects and were clustered and summarized according to the JBI method [21]. The categorization was then discussed within the research team.

Results

Overview

After duplicates were removed, 128 identified abstracts according to the above listed criteria were independently reviewed by 3 authors (KD, IK, and HAE). A full-text screening of the resulting 36 publications led to 9 studies, which were included in our review. Figure 1 shows the process of study selection in a PRISMA flowchart.

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart of the study selection process.



Characteristics of the Included Studies

The included studies were published between 2019 and 2023. The majority of the studies [23-28] were published since 2022. Six empirical studies were conducted in Europe and were limited to the Netherlands, United Kingdom, and Slovakia [19,25-29]. Further, 2 publications that refer to the same study conducted in the United States were included [23,24]. Four studies followed a mixed methods design [19,23,25,27], 3 studies were qualitative studies [24,28,29], and 1 study applied only quantitative methods [26] (Table 1). The data analyses included 1 coping review: in 2020, Colombo et al [30] reported VR applications and exergaming for pulmonary rehabilitation of patients with COPD. Both immersive and nonimmersive approaches were included. However, an assessment of the results with regard to the quality of the included studies was not performed. There was no systematic review on immersive VR rehabilitation programs for home settings and targeting patients with long COVID/post-COVID condition.

Table 1. Full list of the publications included in this scoping review.

Study, year	Objective	Study design	Population type	Sample size, n
Colombo et al [30], 2020	Literature review exploring findings on virtual reality and exergaming applications for pulmonary rehabilitation with focus	Scoping review	Patients with COPD[a]	N/A[b]
Gabriel et al [23], 2023	Qualitative evaluation of the feasibility of a pulmonary rehabilitation program, including educational content for patients with COPD	Mixed methods study	Patients with COPD	18
Gabriel et al [24], 2023	Qualitative evaluation of a pulmonary rehabilitation program for patients with COPD	Qualitative study	Patients with COPD	9
Groenveld et al [25], 2022	Evaluating self-administered virtual reality exercises at home for post–COVID-19 condition	Mixed methods study	Patients with post-COVID	48
Jung et al [19], 2020	Investigating whether virtual reality provides a credible alternative to traditional pulmonary rehabilitation programs and improves compliance among patients with COPD	Mixed methods study	Patients with COPD (Medical Research Council Dyspnea scale 4-5)	10
Lacko and Ruzický [26], 2022	Analyzing the use of virtual reality devices to rehabilitate patients in a controlled outpatient environment as well as in the home environment	Quantitative study	Patients with long-term COVID or post-COVID syndrome	16
Moorhouse et al [29], 2019	Evaluating a virtual reality pulmonary rehabilitation for patients with COPD	Qualitative study	Patients with COPD (Medical Research Council Dyspnea scale 4-5)	10
Ruzicky et al [27], 2022	Investigating the prevention, diagnosis, and treatment of patients after COVID-19 while using artificial intelligence and virtual reality in combination with traditional approaches to patient rehabilitation	Mixed methods study	Patients with post-COVID	10
Smits et al [28], 2022	Developing an evidence-based “Guidance ethics in context” for virtual reality development	Qualitative study	Patients with long COVID (n=20), physical therapists (n=15)	35

[a]COPD: chronic obstructive pulmonary disease.

[b]N/A: not applicable.

Synthesis of Results

The following section presents a synthesis of evidence from the included studies as the result of the qualitative analysis. The narrative description is based on the aspects of study characteristics, rehabilitation program, technical implementation, and evaluative and clinical outcomes.

Study Characteristics

The studies were first divided into 2 groups. The English research group led by Jung et al [19] and Moorhouse et al [29] and a US-American research group [23,24] tested VR applications designed for patients with COPD. They refer to the significance of continuous pulmonary rehabilitation, education of patients regarding the characteristics of their disease, and useful behavioral interventions. They also reported low compliance due to depression related to the condition, low awareness of the potential therapeutic approaches among patients, and lack of knowledge about the benefits of continuous therapy for chronic conditions. Immersive VR applications are intended to create an innovative motivating rehabilitation approach in this context [19,23,24,29]. The other publications refer to the development or evaluation of immersive digital rehabilitation programs for patients with long COVID or post-COVID condition [25-28]. Although the post-COVID condition, as already described in the introduction, refers to persistent symptoms from 4 weeks after a COVID-19 infection [26,27], other studies do not distinguish between long COVID and post-COVID condition at all [25]. Long COVID is usually described as a long-term consequence of infection with SARS-CoV-2 with various physical, psychological, and cognitive symptoms [25-28].

Ruzicky et al [27] collected data on symptoms and severity of the disease during COVID-19 infection and after recovery among a group of students and professors via a questionnaire. Based on the results, which included fatigue, fever, shortness of breath, and depression as common post-COVID symptoms, 2 groups of patients were included in the study. Ten patients had mild muscle pain and shortness of breath after mild exertion and another 6 patients were included who reported severe muscle pain along with shortness of breath [27]. None of the studies required a medical diagnosis or the submission of test results to verify if participants were actually infected with COVID-19; instead, studies relied solely on patient reports. Moreover, none of the studies provided further information regarding the characteristics of participants apart from gender and age or the cultural diversity of populations included. Studies have characterized VR as a way of facilitating access to therapeutic measures in the context of the pandemic [25,27]. In addition, 2 studies implemented individualized and multimodal therapy programs, which address both physical and psychological factors such as stress and anxiety [25,28]. Gabriel et al [23,24] further implemented educational content regarding pulmonary rehabilitation as part of the VR program [23,24].

Collection of Evaluative Data and Assessment

All the studies mentioned dealt with the evaluation of pulmonary VR rehabilitation applications by patients with regard to technology acceptance, usability, and criteria for implementation [19,23-29]. Only 1 publication also included the viewpoint of physical therapists in the assessment [28]. The majority of the studies used interview procedures as the qualitative data collection method [19,24,27-29]. Focus groups in which those affected by COPD were able to discuss their experiences in using the VR application were also used in this context [19,29]. The quantitative survey instruments used in the studies by Jung et al [19] and Groenveld et al [25] included both standardized questionnaires and assessments of physical performance. The questionnaires included the Chronic Respiratory Disease Questionnaire, Patient Health Questionnaire-9 [19], and the 11-point Borg Scale [25], as well as questionnaires regarding psychological and cognitive factors such as Generalized Anxiety Disorder-7 items [19], Hospital Anxiety and Depression Score, and Cognitive Failure Questionnaire [25]. Gabriel et al [23] further applied questionnaires to assess the usability of the application such as the System Usability Scale [23]. Notably, Groenveld et al [25] collected physical performance parameters. These include the 6-Minute Walk Test, Timed Up and Go Test, or 30-Second Chair to Stand Test. In addition, sensors such as smart bracelets (heart rate, pedometer, hand movements, sleep cycles) and pulse oximeters were used to measure the progress of therapy in the studies [27]. Overall, clinical outcomes (physical as well as psychological) and the evaluation of the application used in terms of acceptance and usability were analyzed in the studies.

Rehabilitation Program

The physiotherapy programs in the studies described in Table 2 primarily involve respiratory physiotherapy aimed at improving patients' functional ability [19,23-26,28,29].

Table 2. Characteristics of the rehabilitation programs evaluated in the studies included in this review.

Study, year	Type of training	Length of training	Setting of training	Virtual scenario
Gabriel et al [23], 2023	Physical training (pulmonary rehabilitation), educational content	Not reported	Home setting	Custom-made minigames and multiple choice
Gabriel et al [24], 2023	Physical training (pulmonary rehabilitation)	Not reported	Home setting	Custom-made minigames
Groenveld et al [25], 2022	Physical training, cognitive training, psychological exercises (meditation), independent training: self-management by patients	6-week trial	Home setting	Custom-made applications (minigames) for different exercises
Jung et al [19], 2020	Educational material, physical training (traditional pulmonary rehabilitation + focus on lower extremity), independent training: self-management by patients	8-week trial	Home setting, remotely supervised by health practitioners	Digital avatar
Lacko and Ruzický [26], 2022	Breathing exercises, physical training (upper limb), cognitive training	6-10 weeks	Diverse, remotely supervised by health practitioners	Digital avatar, photorealistic environments
Moorhouse et al [29], 2019	Physical training (pulmonary rehabilitation), educational material, independent training: self-management by patients	8-week trial	Home setting	Digital avatar
Ruzicky et al [27], 2022	Breathing exercises, physical training (upper limb)	Minimum of 3-4 weeks trial up to more than 5 months	Diverse, remotely supervised by health practitioners	Digital avatar, photorealistic environments
Smits et al [28], 2022	Physical training, cognitive training, psychological exercises (meditation), independent training: self-management by patients	6-week trial	Diverse	Custom-made applications (minigames) for different exercises

The focus of the studies was on various aspects such as relief of respiratory distress [19,23,24], rehabilitation of the upper limbs [26,27], and strengthening of the respiratory support muscles [19,29]. For this purpose, physiotherapeutic programs consisting of endurance, strength, and respiratory training were implemented. In addition, mental rehabilitation in light of the impact of the disease [28] as well as long-term goals such as feeling confident leaving the house [19] and quality of life [29] are addressed. Three studies also included exercises to improve cognitive skills [25,26,28]. Ruzicky et al [27] claimed that they addressed prevention, diagnosis, and treatment after COVID-19 but did not provide any further details on that. The length of the therapy programs ranged from 6 to 10 weeks [19,25,26,28,29]. Only in the study of Ruzicky et al [27], a minimum duration of 3-4 weeks was indicated, while in individual cases, the rehabilitation may last several months. In each case, the training program was designed for implementation in the home setting [19,23-29], but half of the publications also reported hybrid use cases within the study. For example, patients were able to try out the VR application in a therapy practice.

Moorhouse et al [29] as well as Jung et al [19], Smits et al [28], and Groenveld et al [25] implemented respiratory rehabilitation measures that were applied fully independently by patients. However, Ruzicky et al [27], Lacko and Ruzický [26], and Jung et al [19] implemented measures to monitor exercise progress. Patients were remotely assisted in setting up the program, for example, via telerehabilitation methods, while the exercises themselves were then performed independently [26,27]. Jung et al [19] focused on ensuring patient safety in their study on patients with COPD. For this reason, data such as heart rate and oxygen saturation were continuously measured in order to be able to intervene in the event of respiratory distress [19]. The selection of exercises was either defined by the patients themselves [29] or discussed with the therapist and adjusted to the requirements of the patients [26-28]. Ruzicky et al [27] and Lacko and Ruzický [26] specified that they included age, gender, and personality type in the selection process but did not specify what the criteria for the personality type is referring to and how age and gender influence the selection of the program.

The interdisciplinary research team consisting of researchers, medical doctors, physical therapists, designers, and VR developers as part of the study by Smits et al [28] developed a toolkit with resources and games. The program consisted mainly of pre-existing games and apps for physical, mental, and cognitive rehabilitation [28]. Three studies stated that a 3D avatar guides the exercises in the immersive VR environment [19,26,29]. Thereby, only 2 studies (from the same research team in Slovakia) chose a realistic representation of the avatar and environment [26,27]. Most commonly used VR headsets were Oculus Quest 2 [23,24,26,27] and Oculus Quest [25,28]. Other headsets used were Pico Interactive Goblin [19] and HTC VIVE Pro EYE [26]. Only Moorhouse et al [29] did not specify the headset used in their research.

Usability and Acceptance

VR-assisted digital respiratory rehabilitation was found to be generally acceptable and feasible in the reviewed studies. In this context, both the enabling factors and barriers for the development and implementation of corresponding applications could be derived from these studies. The benefits of virtual therapy include the aspects of immersion, motivation, as well as autonomy, flexibility, and the possibility of monitoring by therapists. In contrast, barriers include the initial adverse effects related to VR technology and the technical problems and lack of accessibility or usability of the VR applications.

Enabling Factors

In general, study participants described the tested applications as easy to use and enjoyable [19,23-25,28-30]. Studies in the context of respiratory rehabilitation of patients with COPD, in particular, also addressed the immersion in a virtual world, which among other things represents a distraction from the disease. In addition, the avatar guiding the exercises was considered as a social element to a certain extent [19,30]. VR is also considered easier to apply compared to traditional options such as instructions from printed material and booklets for the home environment [29]. Patients with COPD as well as patients with long COVID described the VR application as engaging and pleasurable [23,24,29]. They emphasized experiencing increased motivation to engage in therapeutic measures due to the stimulating or even calming nature of the virtual world, depending on the exercise [19,23-26,29,30]. The increased motivation leads to an increased frequency in usage [19,26,27]. Additionally, Smits et al [28] concluded that the gamification of exercises, in particular, contributes to the motivation of patients. Gamification refers to design elements that reproduce game elements and logics. This includes, for example, the integration of exercises into a game environment, earning scores through correct performance, and competitive approaches such as playing against each other.

Groenveld et al [25] distinguished between users according to their age and found that the duration of VR application increases with age. A possible explanation could be older persons' lesser familiarity with VR technologies because of which they are slower in navigating through the application and they lose interest in the interactive environment and the immersion less quickly [25]. Other studies reported patient groups without any previous experience with VR technology showing difficulties at the beginning of the program [26,27]. Lacko and Ruzický [26] and Ruzicky et al [27] also address the so-called "WOW-effect" in their study, describing first-time users' initial curiosity and great interest in VR [26,27]. However, according to this logic, boredom could also set in after a certain time of using the rehabilitation measure. In 2 studies, comments from study participants, including those who dropped out early, also indicated this same issue [25,26]. The digital program was described as boring, and doubts were expressed about the usefulness of the therapy [25].

Finally, the flexibility and autonomy in the implementation of VR therapy measures is emphasized in the reviewed studies [19,23,24,28,29]. This includes the feasibility of rehabilitation independent of time or location restrictions [23,24,28]. Jung et al [19] concluded that VR reduces the barriers for compliance by increasing the accessibility of rehabilitation programs, which are applicable in the home environment [19]. Patients mentioned that implementation in the home environment, in particular, can contribute to a feeling of comfort and security. In addition, patients reported that monitoring by therapists also gave them confidence [19,23,24]. Nevertheless, VR is perceived as a way to complement traditional therapies offering the advantage of social contact, which VR applications cannot fully compensate for [29].

All rehabilitation programs were customized for the patients in question. However, a distinction was made whether the program was determined by the therapist [19,26,27,29] or by the patients themselves [28]. Smits et al [28] also found that the ability to adapt the therapy to one's own level of rehabilitation and to select exercises individually offered a high added value and contributed to patients' autonomy. Likewise, Jung et al [19] and Moorhouse et al [29] corroborated the same findings in their evaluations.

Barriers

In addition to the benefits of VR-assisted respiratory therapy, barriers to VR implementation were described in these studies. First, the potential adverse effects of VR therapy are motion sickness or overextension due to immersion in an interactive virtual environment [25,28]. Motion sickness is a condition characterized by symptoms of nausea, vomiting, and dizziness caused by conflicting sensations related to motion. Motion sickness is commonly experienced when using VR devices, as the immersive nature of VR can create a sensory conflict between the visual input of a virtual environment and the lack of corresponding physical movement [31]. Further, dizziness, headache, or neck pain were among the most frequent reasons for patients to discontinue the studies [25,28]. Second, immersion can also cause anxiety through a realistic representation of an environment that does not match one's own setting. For example, some patients were afraid of falling while performing the exercises [28]. Third, the VR application was sometimes perceived as overwhelming [25,28]. Smits et al [28] stated the assumption that cognitive impairment as part of long/post-COVID condition may also make the use of VR difficult or impossible for certain patients. This also means that VR-assisted therapy programs should be used individually depending on the condition of each patient. Smits et al [28] therefore recommend a close physiotherapeutic supervision, tracking of training sessions, as well as the monitoring of vital signs. Furthermore, the headset was felt to be too heavy [19,23,29,30].

Studies with a focus on the technical user-friendliness of the application provided information on how to improve the tested applications [26,27,29]. These included, in particular, information on the navigation of the programs (ie, pause or fast forward button [29], one-click solution to start program [26,27]), which states that the technical accessibility and intuitive environment that allows handling even by less technically experienced people, are of great importance. In addition, some studies recommend providing the virtual environment with a simple graphical user interface [26,27] and using clear instructions [30] to facilitate use. Smits et al [28] point out that lack of usability can not only prevent usage and acceptance but also influence the results of studies aiming toward the evaluation of efficacy of programs. An appropriate design could also reduce the abovementioned adverse effects. Therefore, an interdisciplinary design process is recommended to ensure usability [28]. Furthermore, therapists will need proper training and logistical support in regard to the use and implementation of respective VR technology to adequately supervise the training of patients [28].

Clinical Outcomes

In addition to usability and acceptance assessments, all studies listed here also collected clinical parameters for the evaluation of respiratory rehabilitation. These included both physical and psychological outcomes. In regard to the respiratory function and fatigue, an overall improvement was reported in 3 studies [19,25,27]. Jung et al [19] reported that outcomes of female participants with regard to dyspnea and fatigue were even better than the results from male participants. In addition, a significant increase in patients' physical abilities such as strength and mobility was also observed [19,27-29]. Groenveld et al [25] found significant improvements in the 6-Minute Walk Test, grip strength, and 30-Second Chair to Stand Test. Ruzicky et al [27] mentioned that the program focused on developing upper limb mobility and cognitive skills through interactive tasks in the VR environment. However, information on the exact content of the tasks is lacking.

Surveys of the health-related quality of life and the Positive Health questionnaire and 12-Item Short Form Survey showed significant improvement in the quality of life in patients with COPD and patients with long COVID [19,25,29]. Patients felt fitter and were more likely to participate in social activities [19,29]. In the study of Groenveld et al [25], the improvement occurred already after 6 weeks. Further, stress and anxiety were reduced during the rehabilitation [19,25,29]. However, this was partly only true for patients who used specific mental health applications [25]. Participants also felt more confident in dealing with their own disease and in everyday tasks [19,28,29]. Some patients said that they were more mindful of their own health as a result of the program, making time for meditation on their own and setting preventive boundaries in their daily lives [28]. These experienced benefits of VR therapy as well as increased patient motivation also led to improved compliance [19,26-29].

Discussion

Principal Findings

The results of the reviewed studies show internationally available and initial evidence with regard to the development as well as the feasibility of respiratory VR rehabilitation for patients with long COVID in particular. The topic addressed in this review is a very new field of research. Approaches to VR therapy for other therapeutic needs such as for patients who had a stroke or Parkinson disease have already been implemented in the past 15 years [25,26,29]. However, lung-specific virtual physiotherapy appears to be still under development even in regard to the condition of COPD, which is already well-studied. The reported results are promising for VR applications, as the tested applications were described as enjoyable, pleasurable, and motivating when correctly introduced [19,23-27,29]. Furthermore, they can offer a more flexible rehabilitation program without restrictions of time and location [19,23,24,28,29]. Nevertheless, the studies also showed that the implementation of VR therapy interventions cannot generally be considered appropriate for every patient or in every setting because some patients could not complete the training due to motion sickness [25,28], and VR poses various hurdles in terms of the digital literacy of patients and therapists [23,24,28].

Comparison With Prior Work

A previous study that is the most similar to ours is the comprehensive review of VR for pulmonary rehabilitation by Pittara et al [32] in 2023. They offer a broader overview of VR applications in pulmonary rehabilitation by including all types of VR experiences, ranging from nonimmersive to fully immersive, and various populations, including healthy individuals as well as patients with COPD and asthma. In comparison to Pittara et al [32], our review takes a more focused approach by specifically filtering publications to include only immersive VR experiences within pulmonary rehabilitation programs for outpatients. The main strength of our review therefore lies in its clearly defined goals to provide an overview of immersive VR experiences for outpatients experiencing long COVID and COPD. Our review highlights the findings and shortcomings of existing research specifically in relation to the implementation of home-based rehabilitation programs for the target group. In particular, the hurdles of digital literacy identified for implementation at home and the need for training to ensure adequate use and guidance can be highlighted in this regard [27-29].

Our study findings also show the necessity to include the needs and prior knowledge of the target groups in the development of appropriate therapies. Only when appropriate programs achieve added value (eg, through individualized programs, monitoring) and are at the same time easily implemented, they can be applied in practice. The approaches adopted to involve patients in the evaluation of the applications, in terms of usability and acceptance, have shown that patients can provide important information for the development and implementation of VR-supported therapies [19,28,29].

Participatory approaches to technology development, which involve patients already during the development of applications and therapy programs, could help to adapt the applications even more precisely to the needs and requirements of users. The World Health Organization, for instance, recommends a patient-centered development of rehabilitation measures, digital services, and devices in order to support the self-care competence, especially of patients with long/post-COVID condition [4]. Regarding the novelty of the postacute condition, the involvement of patients seems even more crucial because researchers, practitioners, and patients are still undergoing a learning process on how to address and manage the symptoms reported [15]. Therapists, who are to integrate the applications into their therapy services and train the patients in their use, can also provide concrete information on their feasibility in practice. Of the studies analyzed in this review, however, only 1 study design included therapists [28].

Limitations

Although this scoping review was supported by steps, including refinement of the protocol through team discussion, blinded searching, and selection of papers by 3 researchers, several limitations have to be mentioned. First of all, limitations in the scope of the reported study results must be explained, as only 4 groups of researchers (from Slovakia, the Netherlands, United States, and United Kingdom) were involved in the studies analyzed, which represents quite a Eurocentric perspective. This scoping review was limited to studies published in English and German—the spoken languages of the researchers in this scoping review. However, the scoping review method and the explorative search strategy were deliberately chosen in order to be able to explore an overview of the topic while maintaining the focus of the research, and a critical evaluation of the studies was not intended. The second important aspect to note is that study designs implemented by the included studies and the quality in regard to scientific standards were very heterogeneous. Due to the diverse applied assessments and questionnaires, the comparability remains limited. Furthermore, evidence is still limited to pilot studies or a small number of patients, and no long-term studies or randomized controlled studies could be integrated—only cross-sectional surveys.

Conclusion

The results of this scoping review show that VR applications are well accepted by users, especially due to their flexible and individual applicability. Particularly mentioned by patients was the possibility of individualizing training plans and schedules as well as monitoring functions for remote monitoring by therapists. The implementation of rehabilitation measures in a playful, immersive setting contributed to motivating patients and increasing their compliance in respective studies. Initial feasibility studies also show an improvement in physical performance as well as psychological parameters such as confidence in managing the disease and quality of life. At the same time, hurdles arise with regard to the technical feasibility of VR therapies. Virtual applications must be as accessible and easy to use as possible so that patients without prior knowledge can also benefit from the therapy options. Furthermore, scientific research has to further develop empirical reliable findings for the sustainable long-term implementation of support programs for patients with long/post-COVID condition in the years ahead. In particular, the question on how to implement these findings into practice with regard to financing, further education of therapists, technical support, and the alignment of traditional and innovative autonomous approaches to therapy have to be a priority.

Acknowledgments

This research was funded by Federal Ministry of Education and Research Germany (Bundesministerium für Bildung und Forschung; grant: 03COV26A). The authors would also like to acknowledge the support of Translationsregion für digitale Gesundheitsversorgung and thank Anne-Marie Lachmund for proofreading.

Authors' Contributions

KD contributed to conceptualization, methodology, data collection/analysis, and writing the original draft. HAE and IK contributed to data collection/analysis and writing the original draft. SH contributed to funding acquisition, reviewing, and editing. DS contributed to supervision, project administration, funding acquisition, and reviewing. DP contributed to methodology, conceptualization, reviewing, and editing. PJ and MZ contributed to methodology, reviewing, and editing. All authors have read and agreed to the published version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 157 KB - games_v12i1e52309_app1.pdf]

Multimedia Appendix 2
Search strategy.
[DOCX File , 14 KB - games_v12i1e52309_app2.docx]

References

1. Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine 2021 Aug;38:101019 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101019] [Medline: 34308300]
2. COVID-19 rapid guideline: managing the long-term effects of COVID-19. National Institute for Health and Care Excellence guideline [NG188]. 2024. URL: https://www.nice.org.uk/guidance/ng188 [accessed 2024-05-21]
3. Koczulla A, Ankermann T, Behrends U, Berlit P, Berner R, Böing S, et al. [German S1 Guideline Long-/Post-COVID]. Pneumologie 2022 Dec;76(12):855-907 [FREE Full text] [doi: 10.1055/a-1946-3230] [Medline: 36479679]
4. Rajan S, Khunti K, Alwan K, et al. In the wake of the pandemic: Preparing for long COVID. Policy Brief 2021:1-39. [Medline: 33877759]
5. Peter RS, Nieters A, Kräusslich HG, Brockmann SO, Göpel S, Kindle G, EPILOC Phase 1 Study Group. Post-acute sequelae of COVID-19 six to 12 months after infection: population based study. BMJ 2022 Oct 13;379:e071050 [FREE Full text] [doi: 10.1136/bmj-2022-071050] [Medline: 36229057]
6. At least 17 million people in the WHO European Region experienced long COVID in the first two years of the pandemic; millions may have to live with it for years to come. WHO. 2022. URL: https://tinyurl.com/2p9wdmun [accessed 2023-06-23]
7. Li Z, Zheng C, Duan C, Zhang Y, Li Q, Dou Z, et al. Rehabilitation needs of the first cohort of post-acute COVID-19 patients in Hubei, China. Eur J Phys Rehabil Med 2020 Jun;56(3):339-344 [FREE Full text] [doi: 10.23736/S1973-9087.20.06298-X] [Medline: 32672029]
8. Candan SA, Elibol N, Abdullahi A. Consideration of prevention and management of long-term consequences of post-acute respiratory distress syndrome in patients with COVID-19. Physiother Theory Pract 2020 Jun;36(6):663-668. [doi: 10.1080/09593985.2020.1766181] [Medline: 32419564]
9. Kiekens C, Boldrini P, Andreoli A, Avesani R, Gamna F, Grandi M, et al. Rehabilitation and respiratory management in the acute and early post-acute phase - "Instant paper from the field" on rehabilitation answers to the COVID-19 emergency. Eur J Phys Rehabil Med 2020;56(3):323-326 [FREE Full text] [doi: 10.23736/s1973-9087.20.06305-4]
10. Kisner C, Colby L. Vom Griff zur Behandlung: Physiotherapie Grundsätzlich 3, überarb und erw Aufl. Stuttgart: Georg Thieme Verlag; 2010.
11. Spruit MA, Holland AE, Singh SJ, Tonia T, Wilson KC, Troosters T. COVID-19: interim guidance on rehabilitation in the hospital and post-hospital phase from a European Respiratory Society and American Thoracic Society-coordinated international task force. Eur Respir J 2020 Aug 13;56(6):1-13 [FREE Full text] [doi: 10.1183/13993003.02197-2020] [Medline: 32817258]
12. Yang L, Yang T. Pulmonary rehabilitation for patients with coronavirus disease 2019 (COVID-19). Chronic Dis Transl Med 2020 Jun;6(2):79-86 [FREE Full text] [doi: 10.1016/j.cdtm.2020.05.002] [Medline: 32411496]
13. Curci C, Pisano F, Bonacci E, Camozzi DM, Ceravolo C, Bergonzi R, et al. Early rehabilitation in post-acute COVID-19 patients: data from an Italian COVID-19 Rehabilitation Unit and proposal of a treatment protocol. Eur J Phys Rehabil Med 2020 Oct;56(5):633-641 [FREE Full text] [doi: 10.23736/S1973-9087.20.06339-X] [Medline: 32667150]
14. Glöckl R, Buhr-Schinner H, Koczulla AR, Schipmann R, Schultz K, Spielmanns M, et al. [Recommendations from the German Respiratory Society for pulmonary rehabilitation in patients with COVID-19]. Pneumologie 2020 Aug;74(8):496-504 [FREE Full text] [doi: 10.1055/a-1193-9315] [Medline: 32583378]
15. Coping with the unknown: A family's story of living with long COVID. WHO. 2022. URL: https://www.who.int/europe/news-room/feature-stories/item/coping-with-the-unknown---a-family-s-story-of-living-with-long-covid [accessed 2023-06-23]
16. Rinn R, Gao L, Schoeneich S, Dahmen A, Anand Kumar V, Becker P, et al. Digital interventions for treating post-COVID or long-COVID symptoms: scoping review. J Med Internet Res 2023 Apr 17;25:e45711 [FREE Full text] [doi: 10.2196/45711] [Medline: 36943909]

17. Rassouli F, Boutellier D, Duss J, Huber S, Brutsche MH. Digitalizing multidisciplinary pulmonary rehabilitation in COPD with a smartphone application: an international observational pilot study. Int J Chron Obstruct Pulmon Dis 2018;13:3831-3836 [FREE Full text] [doi: 10.2147/COPD.S182880] [Medline: 30538444]
18. Naranjo C, Velasco P, Quevedo W, Naranjo G, Rivas-Lalaleo D, Silva F, et al. Virtual reality system for assistance in treating respiratory disorders. In: Augmented Reality, Virtual Reality, and Computer Graphics. Cham: Springer International Publishing; 2018.
19. Jung T, Moorhouse N, Shi X, Amin MF. A virtual reality-supported intervention for pulmonary rehabilitation of patients with chronic obstructive pulmonary disease: mixed methods study. J Med Internet Res 2020 Jul 07;22(7):e14178 [FREE Full text] [doi: 10.2196/14178] [Medline: 32673224]
20. Freeman D, Haselton P, Freeman J, Spanlang B, Kishore S, Albery E, et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. Lancet Psychiatry 2018 Aug;5(8):625-632 [FREE Full text] [doi: 10.1016/S2215-0366(18)30226-8] [Medline: 30007519]
21. von Elm E, Schreiber G, Haupt CC. [Not Available]. Z Evid Fortbild Qual Gesundhwes 2019 Jun;143:1-7. [doi: 10.1016/j.zefq.2019.05.004] [Medline: 31296451]
22. Hirt J, Nordhausen T. [Not Available]. Pflege 2019;32(5):277-278. [doi: 10.1024/1012-5302/a000690] [Medline: 31538877]
23. Gabriel A, Tsai T, Xhakli T, Finkelstein J. Mixed-methods assessment of a virtual reality-based system for pulmonary rehabilitation. Stud Health Technol Inform 2023 Oct 20;309:245-249. [doi: 10.3233/SHTI230789] [Medline: 37869851]
24. Gabriel A, Tsai T, Xhakli T, Finkelstein J. Patient perceptions of a virtual reality-based system for pulmonary rehabilitation: a qualitative analysis. Stud Health Technol Inform 2023 Jun 29;305:406-409. [doi: 10.3233/SHTI230517] [Medline: 37387051]
25. Groenveld T, Achttien R, Smits M, de Vries M, van Heerde R, Staal B, COVID Rehab Group. Feasibility of virtual reality exercises at home for post-COVID-19 condition: cohort study. JMIR Rehabil Assist Technol 2022 Aug 15;9(3):e36836 [FREE Full text] [doi: 10.2196/36836] [Medline: 35858254]
26. Lacko J, Ruzický E. Possibilities of rehabilitation - telerehabilitation of patients with moderate severe course of COVID-19 disease using virtual reality. In: Jung T, tom Dieck MC, Correia Loureiro SM, editors. Extended Reality and Metaverse. Basel: Springer International Publishing; 2023.
27. Ruzicky E, Sramka M, Sramka M, Lacko J, Masan J, Rottermund J, et al. Providing prevention, diagnosis, and treatment of patients after COVID-19 using artificial intelligence. Neuro Endocrinol Lett 2022 Apr;43(1):9-17. [Medline: 35786805]
28. Smits M, Ludden GDS, Verbeek P, van Goor H. Responsible design and assessment of a SARS-CoV virtual reality rehabilitation programme: guidance ethics in context. Journal of Responsible Innovation 2022 May 24;9(3):344-370. [doi: 10.1080/23299460.2022.2076986]
29. Moorhouse N, Jung T, Shi X, Amin F, Newsham J, McCall S. Pulmonary rehabilitation in virtual reality for COPD patients. In: tom Dieck MC, Jung T, editors. Augmented Reality and Virtual Reality. Basel: Springer International Publishing; 2019.
30. Colombo V, Aliverti A, Sacco M. Virtual reality for COPD rehabilitation: a technological perspective. Pulmonology 2022;28(2):119-133 [FREE Full text] [doi: 10.1016/j.pulmoe.2020.11.010] [Medline: 33358425]
31. Nesbitt K, Nalivaiko E. Cybersickness. In: Lee N, editor. Encyclopedia of Computer Graphics and Games. Basel: Springer International Publishing; 2018.
32. Pittara M, Matsangidou M, Pattichis CS. Virtual reality for pulmonary rehabilitation: comprehensive review. JMIR Rehabil Assist Technol 2023 Oct 02;10:e47114 [FREE Full text] [doi: 10.2196/47114] [Medline: 37782529]

Abbreviations

COPD: chronic obstructive pulmonary disease
JBI: Joanna Briggs Institute
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
VR: virtual reality

Edited by T Leung; submitted 01.09.23; peer-reviewed by A Money, M de Vries; comments to author 03.01.24; revised version received 17.02.24; accepted 31.03.24; published 31.05.24.

Please cite as:
Dalko K, Elsuson HA, Kalter I, Zilezinski M, Hofstetter S, Stoevesandt D, Paulicke D, Jahn P
Virtual Reality Applications for the Implementation of Domestic Respiratory Rehabilitation Programs for Patients With Long COVID and Post-COVID Condition: Scoping Review
JMIR Serious Games 2024;12:e52309
URL: https://games.jmir.org/2024/1/e52309
doi:10.2196/52309
PMID:38819890

Review

Extended Reality for Mental Health Evaluation: Scoping Review

Olatunji Mumini Omisore[1], BTech, MTech, PhD; Ifeanyi Odenigbo[2], BSc, MSc; Joseph Orji[2], BSc, MSc; Amelia Itzel Hernandez Beltran[2], BSc, MSc; Sandra Meier[3], PhD; Nilufar Baghaei[4], BSc, MSc, PhD; Rita Orji[2], BSc, MSc, PhD

[1]Research Centre for Medical Robotics and Minimally Invasive Surgical Devices, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[2]Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada

[3]Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

[4]School of Electrical Engineering and Computer Science, University of Queensland, St Lucia, Australia

Corresponding Author:
Olatunji Mumini Omisore, BTech, MTech, PhD
Research Centre for Medical Robotics and Minimally Invasive Surgical Devices
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences
1068 XueYuan Avenue
Xili, NanShan District
Shenzhen, 518055
China
Phone: 86 13172482240
Email: omisore@siat.ac.cn

Abstract

Background: Mental health disorders are the leading cause of health-related problems worldwide. It is projected that mental health disorders will be the leading cause of morbidity among adults as the incidence rates of anxiety and depression grow worldwide. Recently, "extended reality" (XR), a general term covering virtual reality (VR), augmented reality (AR), and mixed reality (MR), is paving the way for the delivery of mental health care.

Objective: We aimed to investigate the adoption and implementation of XR technology used in interventions for mental disorders and to provide statistical analyses of the design, usage, and effectiveness of XR technology for mental health interventions with a worldwide demographic focus.

Methods: In this paper, we conducted a scoping review of the development and application of XR in the area of mental disorders. We performed a database search to identify relevant studies indexed in Google Scholar, PubMed, and the ACM Digital Library. A search period between August 2016 and December 2023 was defined to select papers related to the usage of VR, AR, and MR in a mental health context. The database search was performed with predefined queries, and a total of 831 papers were identified. Ten papers were identified through professional recommendation. Inclusion and exclusion criteria were designed and applied to ensure that only relevant studies were included in the literature review.

Results: We identified a total of 85 studies from 27 countries worldwide that used different types of VR, AR, and MR techniques for managing 14 types of mental disorders. By performing data analysis, we found that most of the studies focused on high-income countries, such as the United States (n=14, 16.47%) and Germany (n=12, 14.12%). None of the studies were for African countries. The majority of papers reported that XR techniques lead to a significant reduction in symptoms of anxiety or depression. The majority of studies were published in 2021 (n=26, 30.59%). This could indicate that mental disorder intervention received higher attention when COVID-19 emerged. Most studies (n=65, 76.47%) focused on a population in the age range of 18-65 years, while few studies (n=2, 3.35%) focused on teenagers (ie, subjects in the age range of 10-19 years). In addition, more studies were conducted experimentally (n=67, 78.82%) rather than by using analytical and modeling approaches (n=8, 9.41%). This shows that there is a rapid development of XR technology for mental health care. Furthermore, these studies showed that XR technology can effectively be used for evaluating mental disorders in a similar or better way that conventional approaches.

Conclusions: In this scoping review, we studied the adoption and implementation of XR technology for mental disorder care. Our review shows that XR treatment yields high patient satisfaction, and follow-up assessments show significant improvement with large effect sizes. Moreover, the studies adopted unique designs that were set up to record and analyze the symptoms reported by their participants. This review may aid future research and development of various XR mechanisms for differentiated mental disorder procedures.

(JMIR Serious Games 2024;12:e38413) doi:10.2196/38413

KEYWORDS

extended reality; mental disorder; depression; anxiety; exposure therapy

Introduction

Background

Mental disorders are defined as behavioral or mental patterns that cause significant distress or impairment for an individual. These are highly prevalent and, currently, are the leading cause of disability worldwide. In the past decades, a worldwide increase in the incidence of mental disorders has been observed [1,2]. According to the World Health Organization (WHO), mental disorders are the leading cause of disability in the United States and the United Kingdom. The WHO predicted that mental disorders would account for 13% of the total burden of diseases by 2030 [3]. As an indication, around 20% of adults experience 1 type of mental health problem in the United States, the United Kingdom, and related high-income countries [4]. A recent survey shows that the acceleration of socioeconomic developments has increased the prevalence of mental disorders (17.5% adults) in China [5]. Meanwhile, adolescents are characterized with the highest incidence of mental disorders in Canada [6]. Although over 75% of people with mental disorders remain untreated in middle-income countries, 35%-50% of the corresponding range is also found in high-income countries [3].

The WHO estimated that these mental conditions will cost the worldwide economy about US $1 trillion in lost productivity annually [7]. Some of the mental disorders are interlinked. For instance, anxiety and depression remain the most common mental disorders in society [8]. Anxiety is closely linked to mood disorders, and individuals developing depression have often experienced anxiety disorder at some earlier point(s) in life [9]. Although the etiology of anxiety disorder and depression is complex, multiple causal factors, such as rapid social change, stressful work conditions, gender discrimination, social exclusion, an unhealthy lifestyle, physical ill health, human rights violations, and genetics, have been appropriately studied. Many times, mental health researchers have studied the positive effects of evaluating anxiety in combination with other mental conditions, such as pain and depression. For example, Bandelow and Michaelis [10] reported that 1 of every 13 mental disorders is anxiety with major depressive and specific phobia disorder. In general, reports mostly suggest that closer mental care should be addressed by increasing the accessibility and development of tools that patients can use on their own [3].

Conventional Assessment Approaches

Cognitive behavioral therapy (CBT) is a conventional approach that is shown to be effective in the treatment of a wide range of mental disorders, such as anxiety disorder, depression, phobia, and alcohol use problems [11]. CBT is based on the core principle that thoughts impact feelings and feelings impact behavior. During CBT, patients learn to change maladaptive thinking patterns and novel coping behaviors to become and stay healthy. CBT can be as effective as, or even more effective than, other forms of psychological therapy or even psychiatric medications, especially for patients diagnosed with anxiety disorder or depression. CBT is well supported by many clinical practice guidelines [11]. Studies have shown that it is an evidence-based therapy that reliably helps in overcoming depression. However, it involves aiding people to identify and change the bad lifestyles that negatively influence their behavior and emotions [12]. Rather than being a set method, CBT combines procedures that are developed on a certain disorder that has been unevaluated. For instance, the treatment procedure for depression is different from how CBT is used in evaluating phobia and anxiety disorder.

Exposure therapy is a major element of CBT that is more focused on certain mental disorders related to anxiety [13]. In this approach, participants, the subjects being assessed, are exposed to feared objects, activities, or situations in a safe environment, and this is known to reduce patients' fear and possibility of avoidance. With gradual follow-up, participants learn to overcome their anxiety [14]. The variations of exposure therapy can be majorly classified as conventional exposure and modern exposure, usually based on the application context. Conventional exposure includes both in vivo and imaginal exposure. During in vivo exposure, patients are intentionally faced with real-world objects or situations they fear to reduce their anxiety [15]; however, it only works in a small percentage of mental health cases. In contrast, imaginal exposure configures an alternative approach during which patients imagine the worst outcome scenarios to confront their fears within their mind. The effectiveness of imaginal exposure depends on a patient's motivation and their ability to generate fear-inducing imaginations. Exposure therapy is challenging as therapists require extensive training and multiple, long exposure sessions. Consequently, the conventional methods are time-consuming and costly. Recently, XR was evaluated as a new approach for delivering exposure-based therapy for mental disorders. The potential of XR for treating anxiety and depression has been reported [16,17].

Use of Extended Reality in Mental Health

"Extended reality" (XR) is an umbrella term referring to all real, virtual, and mixed environments, wherein interactions are generated by computer technology to engage humans [13]. XR is a rapidly growing technology and is playing prominent roles in different sectors, such as providing clear benefits in many aspects of work and business, including training, collaborative working, and marketing. The technology is rapidly gaining traction in creating imagination of real worlds through virtual reality (VR), augmented reality (AR), and mixed reality (MR). XR was recently conceived for carrying out the evaluation of mental health. Thus, patients with mental disorders can be virtually immersed to allow them to display and confront the disorders they have. It has been noted that the advances in XR tools can transform the health domain remarkably; however, an exciting issue is studying the adoption and implementation levels of current VR, AR, and MR techniques for evaluating mental health [18-20]. In the industry sector, reports showed that the XR medical market was estimated to reach US $1.7 billion in 2022, with a compound annual growth rate of 105.6% from 2018 to 2022 [21]. Thus, supporting XR-based solutions will play a crucial role in the future of mental health. As the market continues to grow, it is safe to assume that developing XR technologies for mental disorder interventions will continue to increase.

In mental health interventions, XR techniques involve the use of single or multiple base technologies to create exposure. The base technologies, namely VR, AR, and MR, involve using computer models to artificially design real-world environments with stimuli sensory features. Thus, the artificial environment can simulate typical contexts that induce mental disorders, such as anxiety, phobia, or pain, to enable users to interact with the environment. Typically, an artificial environment can be developed using 4 main components:

- A high-end graphics-rendering unit that is used to compute and render virtual scenes via a frame buffer
- A 3D stereo display unit that connects users' visual sensory system to the environment
- A tracking system that models users' movement in the virtual environment
- Other input interfaces, such as joysticks or sensory gloves, that provide tactile feedback

Currently, studies suggest that XR-based evaluations can be as effective as conventional exposure–based methods [11,22-24]. It is anticipated that XR technology will offer the greatest promise for mental health care [11]. This is because XR-based exposure therapies are found to be accessible and can offer lasting improvements for different mental health conditions. By analyzing many studies, we have found that a good number of XR techniques exist. These are used to evaluate different mental disorders via different software and hardware technologies [11,22,25,26]. XR systems have been successfully applied in individual, group-based, and internet-based mental health interventions [27-29]. The adoption of XR systems started around 2 decades ago, when Hoffman and coworkers [30] developed a VR gaming system called *SnowWorld* for exposure-based therapy in mental health care. The game provides a systematic way of reducing players' pain perception during burn wound care. Anderson et al [22] presented a follow-up of the first randomized clinical trial to test another format for delivering CBT for social anxiety disorder—VR exposure therapy. The study showed that VR and exposure group therapy has been well established as an effective strategy for evaluating social anxiety disorder.

The application of XR technologies (VR, AR, MR) for mental health care delivery provides opportunities and a greater degree of control for therapists to customize, reproduce, and tweak several evaluation parameters according to an individual patient's needs during mental health care. Such parameters include fan wind, stereo sound, a moving chair, a color display, and odor emitters [31]. This kind of customization may not be achieved in traditional exposure therapy [32,33]. In addition, the risks associated with privacy intrusion reduce as everything is transformed into a virtual environment [34]. Simulated and augmented environments are less scary than the use of in vivo and imaginal exposure in conventional therapy [30]. Exposure-based therapies defined on VR/AR/MR apps have been shown to be effective for evaluating different mental health conditions. This study presents the findings of a scoping review of the state-of-the-art XR systems used in mental health care.

Objective

XR-based mental health interventions have been advancing rapidly. It is critical to analyze the implementation and adoption levels of state-of-the-art XR techniques (ie, a combination of studies that have reported VR, AR, and MR) used for mental health care delivery. This review was conducted following the guidelines outlined by Arksey and O'Malley [35]. The main objective of this scoping review was to show the implementation and usage levels of XR-based therapy in providing care for different mental disorders worldwide. Thus, this review was set to provide a statistical analysis of studies that have recently focused on (1) technological design and usage of XR in mental health care with a worldwide demographic focus, (2) components that are found in different XR interventions used for mental disorders, and (3) effectiveness of XR technology in anxiety and depression as top mental disorders.

Methods

Eligibility Criteria

The adoption of VR, AR, and MR for mental disorder evaluation has evolved over time. The rapid advancement has occurred in a corresponding timeline with developments in the hardware and software used for implementing XR technologies. Hence, we decided to limit our data sources to papers published between August 2016 and December 2023 so as to analyze the state of the art in the study area. Advance search sections of 3 databases by the authors OM, IO, JO, and AB individually and the papers located were later combined to ensure a wide coverage of papers published in the search period. Further, only a limited set of search criteria were used to limit the papers extracted to more relevant ones. We only considered studies that were published in peer-reviewed journals and refereed conferences (with oral presentations). Ten records were identified via professional sourcing. Based on our search strategy and study goal, we decided to use a combination of 2 search rules: (1) all search terms must be present in the paper's title or abstract or both, and (2) the paper's publication year must be within the specified range between August 2016 and December 2023. Additionally, exclusion criteria were defined as (1) duplicate papers; (2) version updates; (3) papers written in a language other than English; (4) studies that reported anxiety or depression as a secondary aspect or induced illness; and (5) papers presenting strengths, weaknesses, opportunities, threats (SWOT) analysis, thesis and citations, and scoping reviews.

Information Sources

We formulated a search strategy used to explore multiple databases to find all recent and relevant studies on XR technologies. We focused on information from studies that focused on depression and anxiety and related mental disorders. The scoping search was conducted on 3 different databases that are library sources for research papers, gray literature, patents, and common information. Our choices of databases were (1) PubMed, (2) Google Scholar, and (3) the ACM Digital Library. These databases were chosen because they provide an interface to generate wild search queries across a variety of disciplines, databases, and journals. In addition, they have the most complete indexes of papers that focus on the theme of this review study. This aided us in simultaneously accessing a broad range of evidence, including technical and peer-reviewed studies reported from different parts of the world, different publishers, and over a long period. We defined our search period to filter out only papers published between August 2016 and December 2023 and indexed in any of the 3 databases. Overlapping papers were filtered to avoid duplication. Multilevel filtering was carried out following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [35] and the associated checklist in Multimedia Appendix 1. This limited the search outcomes to relevant studies that could provide the most valuable data to answer our research objective. The search strategy was set to limit data sources to studies that implemented or used VR, AR, or MR for different mental disorders.

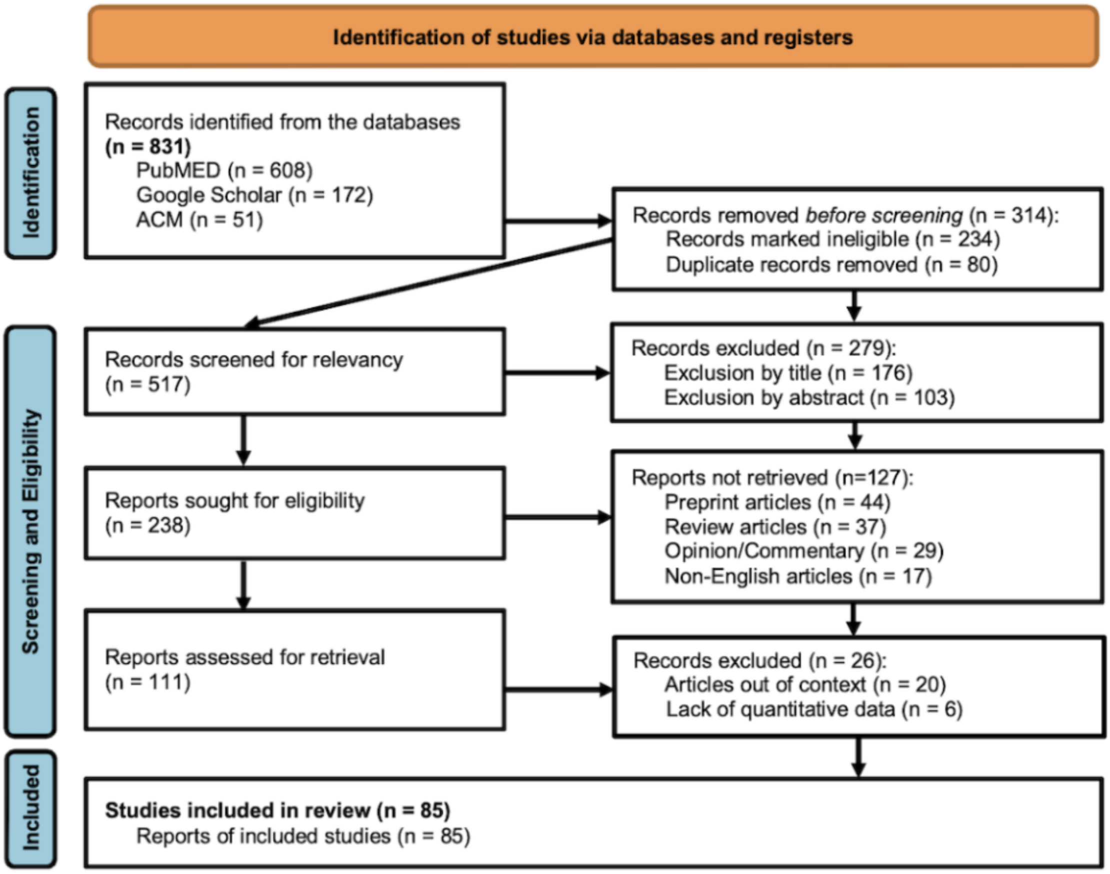
Search Strategy

Our paper search strategy was based on using an organized structure of search terms to retrieve the existing literature in the 3 databases. We combined the keywords in our research objective in order to retrieve relevant papers from the databases. The search terms were discussed among the research team members and were defined as "augmented reality," "mixed reality," "virtual reality," "depression," "anxiety," and "mental health." We chose the free-text search as it is more flexible, and targeting the free words in both the title and abstract fields to limit the final sets provided an efficient way to increase the specificity of our search. These keyword searches were carried out for each concept in our research objective, and we designed the search queries to include a combination of the Boolean operators "OR" and "AND" to reduce the omission of vital papers. These search terms were the most appropriate keywords that were reflected in the subject area and had the utmost relevance to our review objective. We exclusively used full terms during the search in order to avoid any potential conflicts with other terms; for example, VR for "virtual reality" might also be used for "voice recognition," which would make the filtering cumbersome and not necessarily generate additional useful resources. The selection criteria were carefully designed to consider papers that contained 1 or multiple search terms in the their title or abstract or both.

Study Selection

The search yielded 831 papers scraped from the 3 databases, and 10 papers were identified through a professional source. Specifically, by using the defined search filter criteria "anywhere in the article," the search results included 608 (73.16%), 172 (20.7%), and 51 (6.14%) papers scraped from PubMed, Google Scholar, and the ACM Digital Library, respectively. The papers were scraped and processed by following the set of items in the PRISMA checklist. First, authors IO, JO, and AB independently screened the papers retrieved, while author OM performed a quality check on all the final records. Next, irrelevant and duplicate papers (314/831, 37.79%) were removed; thus, a total number of 517 (61.47%) papers were left. The remaining papers were further screened for relevance. With title screening, 176 (30.04%) irrelevant papers were removed. Full abstract reviews was performed in situations in which a paper's relevance could not be resolved from its title. Thus, 103 (58.52%) papers were further screened out. Yet, authors carried out a review of the full text when certainty on a paper's relevance was still lacking in order to decide whether it was relevant. In total, 279 (53.97%) nonrelevant papers were screened out by the authors, leaving only 238 (46.03%) papers for retrieval. A second screening step was required to limit the scoping review to papers that fulfilled the eligibility criteria. Thus, further assessment was carried out, and another 127 (53.36%) papers were removed. The full texts of the remaining 111 (46.64%) papers were retrieved. Papers that were out of context (n=20, 18.02%) and those that lacked quantitative data (n=6, 5.4%) were also excluded. Finally, a total of 85 (76.58%) papers that meet the eligibility criteria were included in this review. All these procedures were performed in Microsoft Excel and without any form of automation. The paper selection process strictly followed the steps shown in Figure 1.

Figure 1. Literature screening and selection flowchart following PRISMA guidelines. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Data Collection and Information Extraction

Three of the authors performed data extraction, while data validity and accuracy were checked by a fourth author. The full texts of the 85 papers were downloaded and shared among the authors for review. The following specific details of the papers were extracted and processed in Excel to analyze the adoption of VR, AR, and MR in managing anxiety and depression and related mental disorders:

- Authors, year, and regions
- The study type and study design focus and health domain
- the methodology (eg, study duration, number of sessions, and duration in minutes)
- The methodology that the study was based on and the evaluation strategy
- The VR/AR/MR app and technology (eg, type of headset, toolkit) used for the study
- Study demographics, such as targeted population, sample size, and age distribution
- Motivational strategies, targeted outcome, and region
- Key findings on using the XR techniques for managing depression and anxiety

The details related to the abovementioned data were used to address the specific research objective guiding this scoping review. The useful insights provided by the data could help developers and researchers on future research on VR/AR/MR for the intervention of mental disorders. In addition, users can learn the importance of such systems, such as the use of XR-based exposure therapy, for mental health.

Results

Publications' Demographics by Country and Year

This scoping review was based on a total of 85 papers [26,32,36-118], which are described in Multimedia Appendix 2. We reported the statistics and meta-analysis of studies that addressed the technological design and usage of XR in mental health, the major components used in different XR interventions for the management of mental disorders, and the effectiveness of the XR technology in anxiety and depression as top mental disorders.

First, we analyzed the country of origin of the papers. The 85 studies were conducted in 27 countries worldwide, as presented in Table 1. Of these 85 studies, 14 (16.47%) were carried out in the United States, followed by 11 (12.94%) in Germany. Compared to a previous study [12], both countries dedicated a good amount of research funding and time to study how XR aids mental health care in the United States and Germany. The data also showed that a good number of studies were conducted in South Korea (n=8, 9.41%) and the Netherlands (n=6, 7.06%). Our study infers that compared to the remaining 21 countries, the aforementioned 4 countries invested a good amount of effort in domestic technological development toward creating XR-based tools for mental disorder. Thus, XR systems contribute immensely to the economic and health care systems of high-income countries. Meanwhile, our data also identified that mental studies are not yet prioritized in Africa. In terms of study frequency by year, Table 2 shows that the majority of papers (n=27, 31.76%) were published in 2021. This could underline a worldwide priority set to advance mental health care. However, none of the studies mentioned whether this was attributed to the ongoing COVID-19 pandemic [119]. Nonetheless, secondary studies identified that the prevalence of anxiety and depression increased by 25% in the first year of the pandemic, while the psychosocial effects of the pandemic varied by regions [120,121]. Hence, it is likely that the pandemic-related increase in mental disorders and the increased adoption of virtual treatment during the pandemic contributed to the rise in the number of XR-based mental health interventions in 2021. Furthermore, there was a great decline in the number of investigations reported in 2022 (n=12, 14.12%) and 2023 (n=16, 18.82%). This coincides with the time COVID-19's worldwide prevalence had decreased.

Table 1. List of countries that conducted mental health studies using XR[a] technologies.

Number	Country of study	Studies (N=85), n (%)
1	Armenia	1 (1.18)
2	Australia	2 (2.35)
3	Austria	2 (2.35)
4	Belgium	1 (1.18)
5	Canada	1 (1.18)
6	China	4 (4.71)
7	Denmark	2 (2.35)
8	France	1 (1.18)
9	Germany	11 (12.94)
10	Hong Kong	1 (1.18)
11	India	1 (1.18)
12	Iran	3 (3.53)
13	Israel	2 (2.35)
14	Japan	1 (1.18)
15	Jordan	2 (2.35)
16	South Korea	8 (9.41)
17	Netherland	6 (7.06)
18	Philippines	1 (1.18)
19	Poland	2 (2.35)
20	Portugal	2 (2.35)
21	Romania	2 (2.35)
22	Singapore	1 (1.18)
23	Spain	7 (8.24)
24	Sweden	2 (2.35)
25	Turkey	1 (1.18)
26	United Kingdom	4 (4.71)
27	United States	14 (16.47)

[a]XR: extended reality.

Table 2. Number of studies published per year.

Year of study	Studies (N=85), n (%)
2016	3 (3.53)
2017	8 (9.41)
2018	4 (4.71)
2019	8 (9.41)
2020	7 (8.24)
2021	27 (31.76)
2022	12 (14.12)
2023	16 (18.82)

Demographics of XR Usage: Age Population

We analyzed the age of the participants included in the studies and categorized the participants as *children*, *teenagers*, *adolescents*, *young adults*, *young and old adults*, and *old adults* based on the age ranges reported in the 85 studies included in this review. Studies that omitted such information were declared as *not specified*. A substantial age overlap was found among the groups of individuals included in the reviewed studies. Thus, *adults* were taken as participants between 18 and 65 years old in the reported studies, and *older adults* were above 65 years old. We observed that around half of the studies (n=38, 44.71%) were designed for adults (between 18 and 65 years old). However, a small number of studies focused on younger age groups: 10 (11.76%) of the 85 studies focused on participants between 0 and 12 years old, while 2 (2.35%) focused on participants in their teenage years as well. The poor representation of participants from each age group in the reviewed studies could be due to a lack of a standardized way of selecting a target audience when developing XR systems for managing mental disorders [122].

In Table 3, we indicated the common classification of participants' age ranges reported in the selected studies, in addition to the statistical information derived from the age groups. This is because the age ranges used for defining the categories of participants in the 85 papers were not unique. In addition, there was great overlap when comparing the categories and age ranges across studies. Thus, we refined the data to synthesize the mean age distribution of the participants included in each study. For this, the age ranges were set as given or generated as (mean – SD) to (mean + SD), when only the mean (SD) was given. The mean age distributions of the participants used for classification are reported in Table 3. The data indicated that the majority of studies (n=47, 55.29%) were designed for an audience with a mean age of 35.079 (SD 9.72) years. The age distribution in this group was particularly dominated with lower and upper values of 18 years (26/47, 55.32%) and 65 years (7/47, 14.89%), respectively, in the different participants' age ranges. The age range of the youngest participants who participated in an XR-based study on mental disorder [58] that investigated how VR reduces the perception of anxiety in infants were a group of children 4-8 years old.

Table 3. Table3. Participant categories found in the included studies by level of maturity.

Participant categories	Studies (N=85), n (%)
Audience group	
Children	10 (11.76)
Teenagers	2 (2.35)
Adolescents	27 (31.76)
Adults	38 (44.71)
Older adults	3 (3.53)
Not specified	5 (5.88)
Age range (years)	
1-10	23 (27.06)
11-20	47 (55.29)
21-60	7 (8.24)
61-85	1 (1.18)
≥85	2 (2.35)
Not specified	5 (5.88)
Participants enrolled	
1-10	10 (11.76)
11-20	9 (10.59)
21-50	23 (27.06)
51-100	22 (25.88)
>100	16 (18.82)
Not specified	5 (5.88)

Demographics of XR Usage: Study Sample Size

We analyzed the sample size of participants enrolled in the 85 studies included in this scoping review. We divided the sample size into 5 different categories and analyzed the number of studies, as reported in Table 3. It can be seen that most studies recruited 51-100 participants (n=22, 25.88%). Next came studies that recruited 21-50 participants (n=23, 27.06%) to evaluate mental health care with XR. Furthermore, 16 (18.82%) studies included over 100 participants, while a small sample size (≤20 participants) was considered in 19 (22.35%) of the 85 studies. Most of these 19 studies were more subtle in their findings and conclusions. Thus, it can be understood that having relatively more participants is helpful to reach better conclusions. Overall, each of the participant categories identified was reported in at least 5 (5.88%) different studies. It is worth mentioning that only 5 (5.88%) studies did not specify the sample size used. The participants' gender distributions were not analyzed, as these data were missing in most studies included in this scoping review.

Demographics of Design and Implementation Strategies

The application of XR systems for mental disorders requires vigorous study and implementation strategies. We analyzed different factors usually considered when designing or evaluating XR systems for mental health interventions. The 3 major considerations found in the 85 selected studies were the *type* of study performed, *design* factors, and the *evaluation method* used to assess each study. With a focus on anxiety and depression, the 4 main types of studies that were are carried out were (1) *discussions*, which are studies with a narrative focus; (2) *experimental*, which are studies that are conducted to investigate the effect of XR techniques on certain groups of subjects or other factors that aid or affect such a setup; (3) *modeling*, which are studies that are conducted to develop new models or setups and validate these on limited subjects or data; and, lastly, (4) *analysis*, which are studies performed without any particular experimental study but relying on the data of previous *experimental* studies.

As reported in Table 4, it was found that 67 (78.82%) of the 85 studies investigated experimentally to report how XR aids in interventions for mental disorders. Meanwhile, 8 (9.41%) studies were based on modeling and analysis each, while 1 (1.18%) study was based on narration (ie, discussion). This shows that most studies were conducted as experimental investigations. Typically, this enabled a direct comparison between mental conditions and their relationships with their causal factors in psychological cornerstone studies [123]. Furthermore, 50 (58.82%) studies investigated the effects of XR immersion. This shows that researchers in this domain are commonly fond of investigating how immersion can influence mental health care procedures. The other design factors of the XR systems found in the 85 studies were on the subject's process automation (n=14, 16.47%); these studies were majorly investigated to observe whether they well emulated real-world situations and environments. Similarly, 9 (10.59%) and 6 (7.06%) studies typically focused on cases of XR personalization and manual execution, respectively.

Table 4. Demographics of the study implementation factors.

Participant categories	Studies (N=85), n (%)
Type of study	
Discussion	1 (1.18)
Experimental	67 (78.82)
Modeling	8 (9.41)
Analysis	8 (9.41)
Design factor	
Personalization	9 (10.59)
Manual	6 (7.06)
Process automation	14 (16.47)
Immersion analysis	50 (58.82)
Evaluation method	
Quantitative method	32 (37.65)
Qualitative method	26 (30.59)
Mixed method	23 (27.06)
Not specified	4 (4.71)

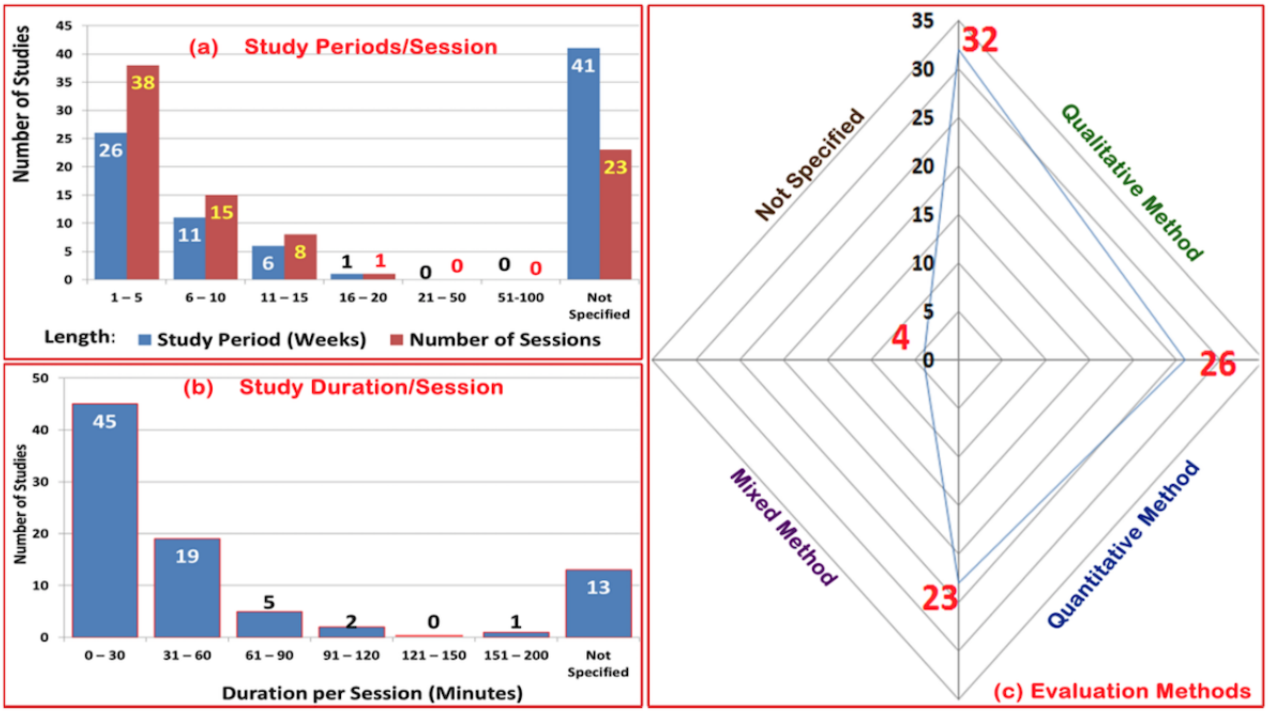
Relationship Between Study Periods and Duration per Session

Next, we analyzed the common evaluation categories reported in the 85 studies. First, categories of study periods (in weeks) and the duration per session (in minutes) were analyzed with respect to the number of sessions in each study. As presented in Figure 2a, some of the studies (n=27, 31.76%) were carried out in 1-5 weeks, 7 (8.24%) studies lasted for 6-10 weeks, and 8 (9.41%) studies lasted for 11-15 weeks . It is worth emphasizing that the longest study (n=1, 1.18%) lasted for 16-20 weeks. In addition, 39 (45.88%) studies evaluated their participants in 1-5 sessions, while 15 (17.65%) studies evaluated their participants in 6-10 sessions.

Furthermore, we analyzed the duration per session (in minutes) for sessions that were reported in each study. In more than half of the studies (n=44, 51.76%), participants used XR techniques for 0-30 minutes, followed by studies requiring 31-60 minutes (n=19, 22.35%) and 61-90 minutes (n=5, 5.88%) of user engagement per session. On the extreme end, the XR technique was used for a single session that lasted over 100 minutes in 4 (4.71%) studies. It was also observed that 13 (15.29%) studies did not specify session durations.

In addition, we analyzed the common evaluation method used in the 85 studies included in this review, and identified 3 main methods: quantitative, qualitative, and mixed. The qualitative assessment approach was applied in 31 (36.47%) of the 85 studies, and it was understood that the qualitative method reveals deeper insights into XR-based evaluation. Furthermore, the quantitative method was used in 9 (10.59%) studies, while the mixed methods approach was used in 18 (21.17%) studies. We also found that 2 (2.35%) studies did not report evaluation methods (see Figure 2c).

Figure 2. Categories of studies included in this review.



Extended Reality and Gamification Strategies for Mental Disorders

It is important to analyze the XR techniques used in interventions for depression and anxiety. First, we analyzed the major strategies found in the 85 studies included in this review. The XR tools used in each study were identified to be either gamified or nongamified. Gamification strategies were adopted in 26 (30.59%) studies, and these strategies were used across 10 mental disorders in all 85 papers. The few exceptions where a gamification strategy was not applied included negative thoughts, panic disorder, and pain and anxiety. Conversely, nongamified strategies were adopted in interventions for the remaining mental conditions, accounting for 59 (69.41%) of the 85 studies. However, alcohol use disorder and attachment behavior were only addressed using gamified XR systems. It may be right to think that gamification strategies are yet to mature for such conditions or, possibly, that existing gamification strategies are not suitable when evaluating such mental disorders using XR techniques or, perhaps, there are ongoing studies to show their applicability.

Extended Reality Development Tools

We also analyzed XR apps that were used in the 85 studies and found that specific systems, such as 3D Unity Pro, are commonly used by many authors in the development of XR systems. Only 68 (77.64%) of the 85 studies reported the name of the actual XR app they implemented or adopted. As reported in Table 5, the most frequently used tool for developing XR platforms was 3D Unity Pro (n=16, 23.53%). This is probably due to its powerful editor to create XR systems and its support for cross-platform development. Similarly, we observed that some studies (n=15, 22.06%) were carried out with custom VR systems. Such adaptive apps are either newly developed or adopted and evaluated for aiding mental health care. Meanwhile, Blender 3D and mobile virtual systems were used in 3 (4.41%) and 10 (14.71%) studies, respectively. Another 24 (28.24%) studies indicated using a development platform but did not specify it, while the remaining 17 (20%) studies did not mention the use of any development platform.

Table 5. Software VR[a] tools commonly used for XR[b] development.

XR app	Studies (N=85), n (%)
3D Unity Pro	16 (18.82)
Custom adaptive VR software	15 (17.64)
Blender 3D	3 (3.52)
Mobile virtual system	10 (11.76)
Others	24 (28.24)

[a]VR: virtual reality.

[b]XR: extended reality.

Hardware Technologies Used for XR in Mental Disorders

To further fulfill the aim of this study, we extracted information about the XR technologies used to deliver mental health care in the 85 studies. To be as inclusive as possible, we only reported the hardware components that were listed for setting up the XR environments in the studies (Table 6). The use of headsets was consistent in 46 (54.11%) of the 85 studies; thus, headsets are the most commonly used component when setting up XR for mental disorder interventions. Typically, it was found that the Oculus head-mounted display (HMD) and VR headsets were common in such studies. In addition, smartphones are a common technology used in setting up the XR environment. It was found that 7 (8.24%) studies included smartphones of different types. These hardware components (ie, HMDs, smartphones, and VR glasses) were increasingly popular in studies where the gamification strategy was adopted. We further analyzed the most popular types of headsets and found that they were headphones, earbuds, and VR HMDs. These last are a more advanced technology and a basic component in most XR studies. As reported in Table 7, there were 9 different types of HMDs used in the 85 studies. HTC Vive and Samsung Gear VR were the most used HMDs in setting up XR systems: these were found in 12 (14.12%) and 11 (12.94%) studies, respectively. The next such HMD was 3D VR glasses, which were used in 7 (8.24%) studies. In addition, Oculus Go and Google VR Box were used in 3 (3.53%) studies each; 2 (2.35%) papers reported to have used Oculus Rift; and different types of VR simulators, such as Oculus CV1, a custom electroencephalography (EEG) cap with a VR HMD, and the Windows MR headset were also used in 1 (1.18%) study each.

Table 6. Types of hardware technology used in XR[a] interventions for mental disorders.

XR technology used	Studies (N=85), n (%)
VR[b] HMD[c]	46 (54.11)
3D VR glasses	5 (5.88)
Smartphone	7 (8.24)
Google VR Box	3 (3.53)
EEG[d]/EMG[e] cap	3 (3.53)
Headphones	3 (3.53)
Biopac MP150	2 (2.35)
Earbuds	2 (2.35)
Location tracker	2 (2.35)
Directional microphone	1 (1.18)
Gamepad	1 (1.18)
Webcam	1 (1.18)
Unspecified	9 (10.59)

[a]XR: extended reality.

[b]VR: virtual reality.

[c]HMD: head-mounted display.

[d]EEG: electroencephalography.

[e]EMG: electromyography.

Table 7. Types of VR[a] headsets.

HMDs[b] used	Studies (N=85), n (%)
HTC Vive	12 (14.12)
Samsung Gear VR	11 (12.94)
3D VR glasses	7 (8.24)
Oculus Go	3 (3.53)
Google VR Box	3 (3.53)
Oculus Rift	2 (2.35)
Oculus CV1	1 (1.18)
EEG[c] VR HMD	1 (1.18)
Windows MR[d] headset	1 (1.18)
Unspecified	5 (5.88)

[a]VR: virtual reality.

[b]HMD: head-mounted display.

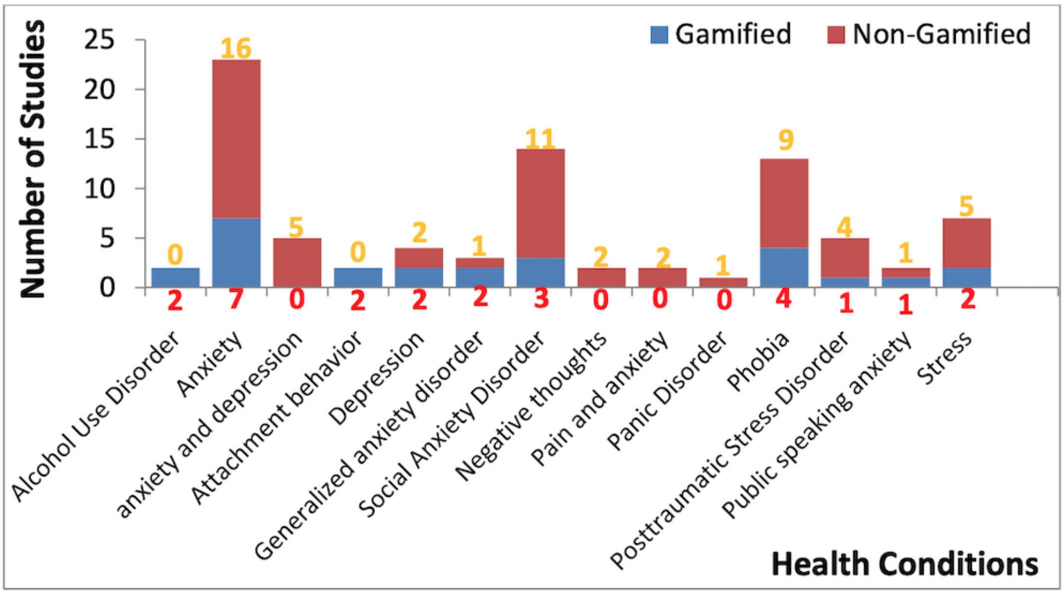
[c]EEG: electroencephalography.

[d]MR: mixed reality.

Extended Reality for Anxiety and Depressive Disorders

We identified 14 types of mental disorders in the 85 papers and illustrated the number of studies investigating each condition, as shown in Figure 3. We observed that most of the XR studies were centered around anxiety and depression (n=53, 62.35%). These included the use of XR for anxiety without depression in 43 (52.94%) studies, where the primary and secondary focus included anxiety, and depression was not considered at all. In 4 (4.71%) studies, we found that minimal attention was on depression, wherein anxiety of any kind was not considered. Typically, 23 (27.05%) studies focused on just anxiety, while the remaining 30 (35.29%) studies combined anxiety with other mental disorders. Anxiety and depressive disorders sometimes have ambiguous borderline definitions; thus, this scoping review focused more on them. We further looked into individual conditions, such as social anxiety disorder and generalized anxiety disorder, which were found in 14 (16.47%) and 3 (3.53%) studies, respectively. These conditions were combined with unclassified distress in 2 (2.35%) studies [66,79]. Of the 85 papers included in this review, 13 (15.29%) were found to have applied XR technologies for phobia-related mental disorders (fear disorders that are a clinical evaluation of anxiety).

Figure 3. Number of studies per clinical condition.



The adoption of XR for mental disorders without anxiety or depression was also studied. For instance, physiological disorders, such as emotion and stress issues, were investigated. Among these, phobias of different kinds (eg, acrophobia, claustrophobia, fear) were investigated in 13 (15.29%) of the 85 studies, and posttraumatic stress disorder was studied in 5 (5.88%) papers. The latter has a similar frequency as 1 of the top mental disorders, depression. We also found that negative thoughts and attachment behavior were investigated in 2 (2.35%) studies each. In contrast to Baghaei et al's [124] findings, we found that generalized anxiety disorder was investigated as a specific clinical condition in 3 (3.53%) studies. In addition, other mental disorders, such as alcoholic use disorder, attention disorder, attachment behavior, negative thoughts, pain and anxiety, and public speaking anxiety, were found in only 2 (2.35%) studies, while panic disorder was found in only 1 (1.18%) study. Thus, XR technology is commonly used for evaluating anxiety disorder. Finally, this scoping study shows that XR-based evaluations are distinctly applied for anxiety and other mental disorders that exclude depression. A typical case includes the development of an XR system for anxiety and phobia, as well as anxiety and psychiatric disorders [16,87]. The benefits of XR for the evaluation and management of mental disorders were identified in the 85 papers that were reviewed. Recent studies show that VR yields the same level of effectiveness as exposure-based therapy for reducing anxiety symptoms [125]. This section mostly uncovered the use of VR technology in anxiety; however, it showed that AR and MR have been recently emphasized as an add-on technology and not a substitute. It is clear that more studies are still needed for evaluating how AR and MR can singly improve mental health.

Discussion

Principal Findings

In the previous section, we focused on the demographics, technologies, and study designs found in existing XR systems used in mental disorder interventions. In this section, the effectiveness of the XR systems for anxiety and depression as top mental disorders are analyzed, as reported in the 85 studies [26,32,36-118] included in this review.

Effectiveness of XR Technology for Mental Disorder Intervention

Following the review of the literature included in this scoping study, it can be concluded that XR systems are commonly used for managing mental disorders. In this scoping review, we found that XR technologies have been majorly used for evaluating anxiety and depression separately, in combination with each other or with other common mental disorders. In the latter case, the majority of studies were targeted at cognitive and behavioral change (ie, subjective care) to improve patients' behavior or attitude or both. In addition, it was observed that among XR technologies, VR-based systems are mostly used. For instance, in some studies [26,77,82], VR was effectively used to evaluate anxiety and depressive symptoms in patients with mental disorders. Similarly, Li and Luo [84] established that gamified XR can reduce depressive disorders through cognitive empathy and mutual understanding among patients and caregivers. Many studies have reported that XR systems help reduce the symptoms of mental disorders. For instance, some authors [26,41,44] have strongly indicated that using XR technology in the psychotherapy process reduced anxiety and depression in their subjects. Similarly, Niharika et al [63] showed a significant decrease in subjects' anxiety scores when using VR eyeglasses during dental treatment.

XR intervention is a safe, noninvasive technique that does not require any previous education and training and has lasting effects. However, McLay et al [59] showed that statistically significant differences between XR-based treatment and conventional approaches may not be a constant thing when applying XR systems for mental disorder intervention. In comparison to standard CBT, some authors [60,73] have improved the psychotherapy of depressive disorder in young adults by developing effective VR-enhanced personal construct therapy. Arnfred et al [101], in the SoREAL study, investigated in vivo group CBT and compared its effects with those of VR exposure CBT on patients diagnosed with social anxiety disorder. Similarly, Shin et al [107] and Donker et al [110] investigated the efficacy of mobile-based VR CBT for panic disorder and phobia interventions. The app-based XR interventions were effective in managing disorder symptoms and restoring subjects' autonomic nervous system. This demonstrates the validity of using XR systems as self-guided and cost-effective therapeutic approaches. Taken together, these studies show that the recent development of XR technologies is gaining traction for mental disorder evaluation and treatment. Thus, some researchers have suggested that future XR interventions should consider providing multiuser experiences that can help increase social engagements for patients who are possibly confined due to disabilities. It can be concluded that virtual environments are as effective as exposure therapy for evaluating mental health. We found studies investigating whether gamified XR is also effective in reducing acrophobia, and the stimuli presented using AR, indeed, induced physiological alterations in the participants [43,44,80].

Effects of XR Design Factors on the Outcomes of Mental Disorder Interventions

Brás et al [78] showed that AR and VR offer high levels of immersion and are optimal solutions for counteracting the effects of in vivo exposure. Weerdmeester et al [74] showed that by engagement and cognitive biofeedback, gamified VR can reduce anxiety symptoms. Grieger et al [81] investigated whether XR-based interventions with multiple design factors can yield better results when used for mental disorders. In a randomized controlled trial, the authors found that personalized VR aids a general positive shift in thoughts and emotions, with increased relaxation and self-refection [81]. This shows that VR systems with multiple features, such as personalization, immersion and focus, interaction design and embodiment, and integration, can enhance treatment outcomes. Similarly, De Asis et al [82] developed a mobile VR-based system for promoting relaxation to reduce anxiety and alter stressful activities among class students. Studies conducted by El-Qirem et al [91] and Traister [79] have shown that XR technology can significantly lower students' anxiety and enhance them psychologically and physiologically with a safe and risk-free therapeutic experience. The related studies [56,57] focused on mental stress management in teenagers and adolescents. Brivio et al [39] compared the efficiencies of 360° panorama technology and a computer-simulated prototype in generating an XR sense of presence, emotions, and relaxation when treating mental disorders. In addition, Lundin et al [98] investigated whether filming virtual environments with a low-cost 360° film camera to produce VR CBT can offer a feasible and acceptable treatment for some kinds of phobia. These studies show that VR exposure therapy can produce long-lasting benefits for mental disorders, consistent with research on a variety of forms of short-term CBT for social anxiety disorder. The results showed that treatment satisfaction was high and that participants had significant improvement at 6-month follow-up, with large effect sizes [98]. Another study [87] showed that VR relaxation induces positive affective states and has short-term effects toward reducing psychiatric stress and anxiety disorder symptoms compared to standard relaxation exercises.

Limitations

XR-based systems have some unique advantages over traditional methods used for mental disorder management. Nevertheless, XR systems also have some limitations. Future developments should consider technological innovation and standardization of treatment options. The following limitations should be considered when interpreting the results of this review. The search strategy developed was limited to using PubMed, Google Scholar, and the ACM Digital Library databases for efficient and accurate search results. This may have excluded qualified papers from other databases. We also found that the number of weeks of evaluation in 40 studies and the number of sessions in 23 studies were not specified in the selected papers. Thus, it is difficult to assert the best number of weeks and sessions needed to validate the use of XR-based technology in mental disorder evaluation. This review identified various major methodological approaches and development tools used by studies. Another limitation of the study is the lack of scientific assessment of the quality of the publications that were included in the scoping review. Moreover, due to the large number of papers reviewed, there is a possibility of that we overlooked valid publications that might have met the inclusion criteria. Non-English papers were not included in this review either. Finally, considering the possibility of bias in the reported outcomes for many reasons, including due to self-reporting and publishing bias that tend to favor papers with positive outcomes, the findings of this scoping review should be applied with caution, especially regarding the effectiveness of XR-based intervention for mental disorders.

Conclusion

XR therapy has been widely used in the care of a variety of mental disorders. This scoping review investigated the adoption of XR for mental disorders, specifically anxiety and depression. The review covered 85 studies that used different types of VR, AR, and MR technologies for mental disorders, with a focus on anxiety and depression. We uncovered that the majority of reviewed papers reported a reduction in the symptoms of anxiety or depression with the use of XR. Moreover, the studies adopted unique designs that were set up to monitor the signs of mental disorders. The recorded signs can be used for formulating appropriate therapies. We also found that XR-based interventions have been shown to be effective approaches with a high level of user acceptability in 18 mental health conditions. Although a considerable number of studies (N=85) were included in this scoping review, some areas are still underinvestigated and, hence, not well represented in the reviewed studies. For instance, the adoption of nongamified strategies was found to have cut across 18 mental health conditions included in this review. However, studies investigating pain and anxiety, negative thoughts, autoimmune disorders, and acquired brain injury did not use any form of gamification strategies. This study was conducted to investigate the implementation and adoption levels of XR for mental health care delivery. Our study outputs indicate that many studies have focused on anxiety, either alone or in combination with other conditions. Meanwhile, a limited number of studies have solely focused on depression. In a previous study, Baghaei et al [124] also showed that supporting people with depression in XR settings is an interesting area to explore for mental health care. We recommend that future work should conduct controlled trials to investigate and compare the effectiveness of using XR-based intervention in mental health care and the benefits and costs of XR in mental disorder management.

Acknowledgments

NB acknowledges the University of Queensland for the startup grant. OMO was supported by the Shenzhen Engineering Laboratory for Diagnosis and Treatment of Interventional Surgical Robot Key Technology (Grant XMHT20220104009). RO acknowledges funding support from the Canada Research Chairs (CRC) program and the Natural Sciences and Engineering Research Council of Canada (NSERC) through the discovery grant.

Data Availability

The data used for this review have been included as Multimedia Appendices with the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.

[DOCX File , 34 KB - games_v12i1e38413_app1.docx]

Multimedia Appendix 2

Table of studies.
[DOCX File , 93 KB - games_v12i1e38413_app2.docx]

References

1. Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: a review. EXCLI J 2017;16:1057-1072 [FREE Full text] [doi: 10.17179/excli2017-480] [Medline: 28900385]
2. Oyebode O, Ndulue C, Adib A, Mulchandani D, Suruliraj B, Orji FA, et al. Health, psychosocial, and social issues emanating from the COVID-19 pandemic based on social media comments: text mining and thematic analysis approach. JMIR Med Inform 2021 Apr 06;9(4):e22734 [FREE Full text] [doi: 10.2196/22734] [Medline: 33684052]
3. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level. World Health Organization. 2011. URL: https://apps.who.int/gb/ebwha/pdf_files/EB130/B130_9-en.pdf [accessed 2024-06-06]
4. Whiteford H, Ferrari A, Degenhardt L. Global burden of disease studies: implications for mental and substance use disorders. Health Aff (Millwood) 2016 Jun 01;35(6):1114-1120. [doi: 10.1377/hlthaff.2016.0082] [Medline: 27269030]
5. Que J, Lu L, Shi L. Development and challenges of mental health in China. Gen Psychiatr 2019 Feb 20;32(1):e100053 [FREE Full text] [doi: 10.1136/gpsych-2019-100053] [Medline: 31179426]
6. Colizzi M, Lasalvia A, Ruggeri M. Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care? Int J Ment Health Syst 2020 Mar 24;14(1):23 [FREE Full text] [doi: 10.1186/s13033-020-00356-9] [Medline: 32226481]
7. Investing in treatment for depression and anxiety leads to fourfold return. World Health Organization. 2016. URL: https://www.who.int/news/item/13-04-2016-investing-in-treatment-for-depression-and-anxiety-leads-to-fourfold-return [accessed 2024-06-06]
8. Liang D, Mays VM, Hwang WC. Integrated mental health services in China: challenges and planning for the future. Health Policy Plan 2018 Jan 01;33(1):107-122 [FREE Full text] [doi: 10.1093/heapol/czx137] [Medline: 29040516]
9. Meier SM, Mattheisen M, Mors O, Mortensen PB, Laursen TM, Penninx BW. Increased mortality among people with anxiety disorders: total population study. Br J Psychiatr 2016 Sep 02;209(3):216-221 [FREE Full text] [doi: 10.1192/bjp.bp.115.171975] [Medline: 27388572]
10. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin Neurosci 2022 Apr 01;17(3):327-335. [doi: 10.31887/dcns.2015.17.3/bbandelow]
11. Andrews G, Bell C, Boyce P, Gale C, Lampe L, Marwat O, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment of panic disorder, social anxiety disorder and generalised anxiety disorder. Aust N Z J Psychiatr 2018 Nov 30;52(12):1109-1172. [doi: 10.1177/0004867418799453]
12. Baghaei N, Lehan S, Hlasnik A, Emanov K, Hach S, Naslund J, et al. individualised virtual reality applications for supporting depression: a feasibility study. 2021 Presented at: EICS '21: 13th ACM SIGCHI Symposium on Engineering Interactive Computing Systems; June 8-11, 2021; Virtual, Eindhoven, the Netherlands. [doi: 10.1145/3459926.3464761]
13. Gautam M, Tripathi A, Deshmukh D, Gaur M. Cognitive behavioral therapy for depression. Indian J Psychiatr 2020;62(8):223. [doi: 10.4103/psychiatry.indianjpsychiatry_772_19]
14. Joseph JS, Gray MJ. Exposure therapy for posttraumatic stress disorder. JBA-OVTP 2008;1(4):69-79. [doi: 10.1037/h0100457]
15. Rothbaum BO, Hodges L, Smith S, Lee JH, Price L. A controlled study of virtual reality exposure therapy for the fear of flying. J Consult Clin Psychol 2000 Dec;68(6):1020-1026. [doi: 10.1037//0022-006x.68.6.1020] [Medline: 11142535]
16. Boeldt D, McMahon E, McFaul M, Greenleaf W. Using virtual reality exposure therapy to enhance treatment of anxiety disorders: identifying areas of clinical adoption and potential obstacles. Front Psychiatr 2019;10(773):773 [FREE Full text] [doi: 10.3389/fpsyt.2019.00773] [Medline: 31708821]
17. Kjernisted K, McIntosh D. Venlafaxine extended release (XR) in the treatment of panic disorder. Ther Clin Risk Manag 2007 Mar;3(1):59-69 [FREE Full text] [doi: 10.2147/tcrm.2007.3.1.59] [Medline: 18360616]
18. Ligthart S, Meynen G, Biller-Andorno N, Kooijmans T, Kellmeyer P. Is virtually everything possible? The relevance of ethics and human rights for introducing extended reality in forensic psychiatry. AJOB Neurosci 2022 Mar 29;13(3):144-157. [doi: 10.1080/21507740.2021.1898489] [Medline: 33780323]
19. Andrews C, Southworth MK, Silva JNA, Silva JR. Extended reality in medical practice. Curr Treat Options Cardiovasc Med 2019 Mar 30;21(4):18 [FREE Full text] [doi: 10.1007/s11936-019-0722-7] [Medline: 30929093]
20. Parveau M, Adda M. Toward a user-centric classification scheme for extended reality paradigms. J Ambient Intell Human Comput 2019 Jun 18;11(6):2237-2249. [doi: 10.1007/s12652-019-01352-9]
21. Woods J. impact of XR in mental health. XR TODAY. 2020. URL: https://www.xrtoday.com/virtual-reality/the-impact-on-xr-in-mental-health/ [accessed 2024-06-06]
22. Anderson PL, Edwards SM, Goodnight JR. Virtual reality and exposure group therapy for social anxiety disorder: results from a 4–6 year follow-up. Cogn Ther Res 2016 Nov 15;41(2):230-236. [doi: 10.1007/s10608-016-9820-y]

23. Tsai C, Yeh S, Huang Y, Wu Z, Cui J, Zheng L. The effect of augmented reality and virtual reality on inducing anxiety for exposure therapy: a comparison using heart rate variability. J Healthc Eng 2018 Nov 25;2018:6357351-6357358 [FREE Full text] [doi: 10.1155/2018/6357351] [Medline: 30595830]
24. Lin AJ, Cheng F, Chen CB. Use of virtual reality games in people with depression and anxiety. 2020 Presented at: ICMIP '20: 5th International Conference on Multimedia and Image Processing; January 10-12, 2020; Nanjing, China. [doi: 10.1145/3381271.3381299]
25. Amaral CP, Simões MA, Mouga S, Andrade J, Castelo-Branco M. A novel brain computer interface for classification of social joint attention in autism and comparison of 3 experimental setups: a feasibility study. J Neurosci Methods 2017 Oct 01;290:105-115. [doi: 10.1016/j.jneumeth.2017.07.029] [Medline: 28760486]
26. Appel L, Appel E, Bogler O, Wiseman M, Cohen L, Ein N, et al. Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. Front Med (Lausanne) 2019 Jan 15;6:329 [FREE Full text] [doi: 10.3389/fmed.2019.00329] [Medline: 32010701]
27. Ong T, Wilczewski H, Paige SR, Soni H, Welch BM, Bunnell BE. Extended reality for enhanced telehealth during and beyond COVID-19: viewpoint. JMIR Serious Games 2021 Jul 26;9(3):e26520 [FREE Full text] [doi: 10.2196/26520] [Medline: 34227992]
28. Morimoto T, Kobayashi T, Hirata H, Otani K, Sugimoto M, Tsukamoto M, et al. XR (extended reality: virtual reality, augmented reality, mixed reality) technology in spine medicine: status quo and quo vadis. J Clin Med 2022 Jan 17;11(2):470 [FREE Full text] [doi: 10.3390/jcm11020470] [Medline: 35054164]
29. Eckert M, Volmerg JS, Friedrich CM. Augmented reality in medicine: systematic and bibliographic review. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e10967 [FREE Full text] [doi: 10.2196/10967] [Medline: 31025950]
30. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007 Oct;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
31. Srivastava K, Das R, Chaudhury S. Virtual reality applications in mental health: challenges and perspectives. Ind Psychiatr J 2014;23(2):83-85 [FREE Full text] [doi: 10.4103/0972-6748.151666] [Medline: 25788795]
32. Dechant M, Trimpl S, Wolff C, Mühlberger A, Shiban Y. Potential of virtual reality as a diagnostic tool for social anxiety: a pilot study. Comput Hum Behav 2017 Nov;76:128-134. [doi: 10.1016/j.chb.2017.07.005]
33. Kamkuimo Kengne S, Fossaert M, Girard B, Menelas B. Action-Centered Exposure Therapy (ACET): a new approach to the use of virtual reality to the care of people with post-traumatic stress disorder. Behav Sci (Basel) 2018 Aug 16;8(8):76 [FREE Full text] [doi: 10.3390/bs8080076] [Medline: 30115837]
34. Dick E. Balancing user privacy and innovation in augmented and virtual reality. Information Technology & Innovation Foundation. 2021 Mar 4. URL: https://tinyurl.com/yvmuybc8 [accessed 2024-06-06]
35. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
36. Bani Mohammad E, Ahmad M. Virtual reality as a distraction technique for pain and anxiety among patients with breast cancer: a randomized control trial. Pall Supp Care 2018 Sep 10;17(1):29-34. [doi: 10.1017/s1478951518000639]
37. Bilgin P, Agres K, Robinson N, Wai AAP, Guan C. A comparative study of mental states in 2D and 3D virtual environments using EEG. 2019 Presented at: SMC2019: 2019 IEEE International Conference on Systems, Man and Cybernetics; October 6-9, 2019; Bari, Italy. [doi: 10.1109/smc.2019.8914326]
38. Bjarling EA, Cicero R, Sankar A, Sekar A. Thought disposal: co-designing a virtual interaction to reduce stress in teens. 2019 Presented at: 18th ACM International Conference on Interaction Design and Children; June 12-15, 2019; Boise, ID. [doi: 10.1145/3311927.3325313]
39. Brivio E, Serino S, Negro Cousa E, Zini A, Riva G, De Leo G. Virtual reality and 360° panorama technology: a media comparison to study changes in sense of presence, anxiety, and positive emotions. Virtual Reality 2020 Jul 09;25(2):303-311. [doi: 10.1007/s10055-020-00453-7]
40. Chicchi Giglioli IA, Pravettoni G, Sutil Martín DL, Parra E, Raya MA. A novel integrating virtual reality approach for the assessment of the attachment behavioral system. Front Psychol 2017 Jun 28;8:959 [FREE Full text] [doi: 10.3389/fpsyg.2017.00959] [Medline: 28701967]
41. Dehghan F, Jalali R, Bashiri H. The effect of virtual reality technology on preoperative anxiety in children: a Solomon four-group randomized clinical trial. Perioper Med (Lond) 2019 Jun 4;8(1):5 [FREE Full text] [doi: 10.1186/s13741-019-0116-0] [Medline: 31171963]
42. Deppermann S, Notzon S, Kroczek A, Rosenbaum D, Haeussinger F, Diemer J, et al. Functional co-activation within the prefrontal cortex supports the maintenance of behavioural performance in fear-relevant situations before an iTBS modulated virtual reality challenge in participants with spider phobia. Behav Brain Res 2016 Jul 01;307:208-217. [doi: 10.1016/j.bbr.2016.03.028] [Medline: 26996315]
43. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. JAMA Psychiatry 2019 Jul 01;76(7):682-690 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]

44. Donker T, Van Esveld S, Fischer N, Van Straten A. 0Phobia - towards a virtual cure for acrophobia: study protocol for a randomized controlled trial. Trials 2018 Aug 09;19(1):433 [FREE Full text] [doi: 10.1186/s13063-018-2704-6] [Medline: 30092844]
45. Kim HE, Hong Y, Kim M, Jung YH, Kyeong S, Kim J. Effectiveness of self-training using the mobile-based virtual reality program in patients with social anxiety disorder. Comput Hum Behav 2017 Aug;73:614-619. [doi: 10.1016/j.chb.2017.04.017]
46. Fanger Y, Pfeuffer K, Helmbrecht U, Alt F. PIANX@ a platform for piano players to alleviate music performance anxiety using mixed reality. 2020 Presented at: MUM 2020: 19th International Conference on Mobile and Ubiquitous Multimedia; November 22-25, 2020; Essen, Germany. [doi: 10.1145/3428361.3428394]
47. Gamito P, Oliveira J, Matias M, Cunha E, Brito R, Lopes PF, et al. Virtual reality cognitive training among individuals with alcohol use disorder undergoing residential treatment: pilot randomized controlled trial. J Med Internet Res 2021 Jan 29;23(1):e18482 [FREE Full text] [doi: 10.2196/18482] [Medline: 33512329]
48. Ghiţă A, Hernández-Serrano O, Fernández-Ruiz J, Moreno M, Monras M, Ortega L, et al. Attentional bias, alcohol craving, and anxiety implications of the virtual reality cue-exposure therapy in severe alcohol use disorder: a case report. Front Psychol 2021 Feb 22;12:543586 [FREE Full text] [doi: 10.3389/fpsyg.2021.543586] [Medline: 33692713]
49. Gold JI, Annick ET, Lane AS, Ho K, Marty RT, Espinoza JC. "Doc McStuffins: Doctor for a Day" virtual reality (DocVR) for pediatric preoperative anxiety and satisfaction: pediatric medical technology feasibility study. J Med Internet Res 2021 Apr 19;23(4):e25504 [FREE Full text] [doi: 10.2196/25504] [Medline: 33730687]
50. Hong Y, Kim HE, Jung YH, Kyeong S, Kim J. Usefulness of the mobile virtual reality self-training for overcoming a fear of heights. Cyberpsychol Behav Soc Netw 2017 Dec;20(12):753-761. [doi: 10.1089/cyber.2017.0085] [Medline: 29211504]
51. Hur J, Shin H, Jung D, Lee H, Lee S, Kim GJ, et al. Virtual reality-based psychotherapy in social anxiety disorder: fMRI study using a self-referential task. JMIR Ment Health 2021 Apr 14;8(4):e25731 [FREE Full text] [doi: 10.2196/25731] [Medline: 33851931]
52. Inozu M, Çelikcan U, Trak E, Üzümcü E, Nergiz H. Assessment of virtual reality as an anxiety and disgust provoking tool: the use of VR exposure in individuals with high contamination fear. Cyberpsychology 2021 Feb 16;15(1):Article 7. [doi: 10.5817/CP2021-1-7]
53. Jeong HS, Lee JH, Kim HE, Kim J. Appropriate number of treatment sessions in virtual reality-based individual cognitive behavioral therapy for social anxiety disorder. J Clin Med 2021 Feb 26;10(5):915 [FREE Full text] [doi: 10.3390/jcm10050915] [Medline: 33652739]
54. Kothgassner OD, Felnhofer A, Hlavacs H, Beutl L, Palme R, Kryspin-Exner I, et al. Salivary cortisol and cardiovascular reactivity to a public speaking task in a virtual and real-life environment. Comput Hum Behav 2016 Sep;62:124-135. [doi: 10.1016/j.chb.2016.03.081]
55. Levy F, Leboucher P, Rautureau G, Jouvent R. E-virtual reality exposure therapy in acrophobia: a pilot study. J Telemed Telecare 2015 Aug 06;22(4):215-220. [doi: 10.1177/1357633x15598243]
56. Liszio S, Basu O, Masuch M. A universe inside the MRI Scanner: an in-bore virtual reality game for children to reduce anxiety and stress. 2020 Presented at: CHI PLAY '20: The Annual Symposium on Computer-Human Interaction in Play; November 2-5, 2020; Ottawa, Canada. [doi: 10.1145/3410404.3414263]
57. Liszio S, Masuch M. Virtual reality MRI: playful reduction of children's anxiety in MRI exams. 2017 Presented at: 2017 Conference on Interaction Design and Children; June 27-30, 2017; Stanford, CA. [doi: 10.1145/3078072.3079713]
58. McInerney J, Brown P, Bird JC, Nickless A, Brown G, Freeman D. Does raising heart rate prior to a behavioural test enhance learning in cognitive therapy for anxiety? An experimental test for the treatment of fear of heights using virtual reality. Behav Res Ther 2021 Sep;144:103928 [FREE Full text] [doi: 10.1016/j.brat.2021.103928] [Medline: 34246887]
59. McLay RN, Baird A, Webb-Murphy J, Deal W, Tran L, Anson H, et al. A randomized, head-to-head study of virtual reality exposure therapy for posttraumatic stress disorder. Cyberpsychol Behav Soc Netw 2017 Apr;20(4):218-224. [doi: 10.1089/cyber.2016.0554] [Medline: 28394217]
60. Montesano A, Medina JC, Paz C, García-Mieres H, Niño-Robles N, García-Grau E, et al. Does virtual reality increase the efficacy of psychotherapy for young adults with mild-to-moderate depression? A study protocol for a multicenter randomized clinical trial. Trials 2021 Dec 13;22(1):916 [FREE Full text] [doi: 10.1186/s13063-021-05809-1] [Medline: 34903251]
61. Kim S, Moon JG, Cho CH, Kim GJ. Effects of immersive virtual reality content type to mindfulness and physiological parameters. 2020 Presented at: VRST '20: 26th ACM Symposium on Virtual Reality Software and Technology; November 1-4, 2020; Virtual, Canada p. 1-9. [doi: 10.1145/3385956.3418942]
62. Keshavarz N, Hashemi Nosrat Abad T, Beyrami M, Mahmoud alilou M, Bakhshipour Roudsari A. Efficacy of virtual reality based worry exposure therapy on the anxiety severity and worry in generalized anxiety disorder. Adv Biosci Clin Med 2021 Apr 29;9(2):21. [doi: 10.7575/aiac.abcmed.v.9n.2p.21]
63. Niharika P, Reddy N, Srujana P, Srikanth K, Daneswari V, Geetha K. Effects of distraction using virtual reality technology on pain perception and anxiety levels in children during pulp therapy of primary molars. J Indian Soc Pedod Prev Dent 2018;36(4):364. [doi: 10.4103/jisppd.jisppd_1158_17]
64. Rahani V, Vard A, Najafi M. Claustrophobia game: design and development of a new virtual reality game for treatment of claustrophobia. J Med Signals Sens 2018;8(4):231. [doi: 10.4103/jmss.jmss_27_18]

65. Shimizu S, Jincho N, Kikuchi H. Influence of interactive questions on the sense of presence and anxiety in a virtual-reality job-interview simulation. 2019 Presented at: ICVARS '19: 3rd International Conference on Virtual and Augmented Reality Simulations; February 23-25, 2019; Perth, Australia. [doi: 10.1145/3332305.3332307]
66. Sigurvinsdottir R, Soring K, Kristinsdottir K, Halfdanarson SG, Johannsdottir KR, Vilhjalmsson HH, et al. Social anxiety, fear of negative evaluation, and distress in a virtual reality environment. Behav Change 2021 Feb 24;38(2):109-118. [doi: 10.1017/bec.2021.4]
67. Stupar-Rutenfrans S, Ketelaars LEH, van Gisbergen MS. Beat the fear of public speaking: mobile 360° video virtual reality exposure training in home environment reduces public speaking anxiety. Cyberpsychol Behav Soc Netw 2017 Oct;20(10):624-633. [doi: 10.1089/cyber.2017.0174] [Medline: 29039704]
68. Tadayon R, Gupta C, Crews D, McDaniel T. Do trait anxiety scores reveal information about our response to anxious situations? A psycho-physiological VR study. 2019 Presented at: HealthMedia '19: 4th International Workshop on Multimedia for Personal Health & Health Care; October 21, 2019; Nice, France. [doi: 10.1145/3347444.3356239]
69. Tielman ML, Neerincx MA, Bidarra R, Kybartas B, Brinkman W. A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories. J Med Syst 2017 Aug 11;41(8):125 [FREE Full text] [doi: 10.1007/s10916-017-0771-y] [Medline: 28699083]
70. Tomasi D, Ferris H, Booraem P, Enman L, Gates S, Reyns T. Olfactory virtual reality (OVR) for wellbeing and reduction of stress, anxiety and pain. J Med Res Health Sci 2021 Mar;4(3):1212-1221. [doi: 10.15520/jmrhs.v4i3.322]
71. Trahan MH, Morley RH, Nason EE, Rodrigues N, Huerta L, Metsis V. Virtual reality exposure simulation for student veteran social anxiety and PTSD: a case study. Clin Soc Work J 2021 Jan 19;49(2):220-230 [FREE Full text] [doi: 10.1007/s10615-020-00784-7] [Medline: 33487778]
72. van't Wout-Frank M, Shea MT, Larson VC, Greenberg BD, Philip NS. Combined transcranial direct current stimulation with virtual reality exposure for posttraumatic stress disorder: feasibility and pilot results. Brain Stimul 2019 Jan;12(1):41-43 [FREE Full text] [doi: 10.1016/j.brs.2018.09.011] [Medline: 30266416]
73. Voinescu A, Petrini K, Stanton Fraser D, Lazarovicz R, Papavă I, Fodor LA, et al. The effectiveness of a virtual reality attention task to predict depression and anxiety in comparison with current clinical measures. Virtual Reality 2021 Apr 17;27(1):119-140. [doi: 10.1007/s10055-021-00520-7]
74. Weerdmeester J, van Rooij MMJW, Maciejewski DF, Engels RCME, Granic I. A randomized controlled trial assessing the efficacy of a virtual reality biofeedback video game: anxiety outcomes and appraisal processes. Technol Mind Behav 2021 Aug 05;2(2):1-16. [doi: 10.1037/tmb0000028]
75. Xia Y, Ding Z, Gan Y. Gamification App Design Based on Augmented Reality Technique for Depression Rehabilitation. Cham: Springer; 2023:68-76.
76. Lindner P, Dagöö J, Hamilton W, Miloff A, Andersson G, Schill A, et al. Virtual reality exposure therapy for public speaking anxiety in routine care: a single-subject effectiveness trial. Cogn Behav Ther 2020 Sep 01;50(1):67-87 [FREE Full text] [doi: 10.1080/16506073.2020.1795240] [Medline: 35549980]
77. Wang Z, An J, Zhao Y, Wang Z, Yu Y, Jiang T, et al. The effect of restorative environment and presence based on virtual reality for anxiety and depression. In: Ahram TZ, Falcão CS, editors. Advances in Usability, User Experience, Wearable and Assistive Technology. Lecture Notes in Networks and Systems, vol 275. Cham: Springer; 2021.
78. Brás S, Soares SC, Cruz T, Magalhães T, Marques B, Dantas C, et al. The feasibility of an augment reality system to study the psychophysiological correlates of fear-related responses. Brain Behav 2018 Sep 22;8(9):e01084 [FREE Full text] [doi: 10.1002/brb3.1084] [Medline: 30136383]
79. Traister TA. Virtual reality simulation's influence on nursing students' anxiety and communication skills with anxious patients: a pilot study. Clin Simul Nurs 2023 Sep;82:101433. [doi: 10.1016/j.ecns.2023.101433]
80. Krupić D, Žuro B, Corr PJ. Anxiety and threat magnification in subjective and physiological responses of fear of heights induced by virtual reality. Pers Individ Differ 2021 Feb;169:109720. [doi: 10.1016/j.paid.2019.109720]
81. Grieger F, Klapperich H, Hassenzahl M. Trash it, punch it, burn it – using virtual reality to support coping with negative thoughts. 2021 Presented at: CHI2019: Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, UK p. 1-6. [doi: 10.1145/3411763.3451738]
82. De Asis KMR, Guillem EJP, Reyes FAM, Samonte MJ. Serenity: a stress-relieving virtual reality application based on Philippine environmental variables. 2020 Jun Presented at: ICFET 2020: 6th International Conference on Frontiers of Educational Technologies; June 5-8, 2020; Tokyo Japan p. 155-159. [doi: 10.1145/3404709.3404745]
83. Liszio S, Graf L, Basu O, Masuch M. Pengunaut trainer: a playful VR app to prepare children for MRI examinations: in-depth game design analysis. 2020 Presented at: 2020 Interaction Design and Children Conference, London, United Kingdom; June 21-24, 2020; London, UK p. 470-482. [doi: 10.1145/3392063.3394432]
84. Li YJ, Luo HI. Depression prevention by mutual empathy training: using virtual reality as a tool. 2021 Presented at: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops; March 27-April 1, 2021; Lisbon, Portugal p. 60-63. [doi: 10.1109/vrw52623.2021.00017]
85. Premkumar P, Heym N, Brown DJ, Battersby S, Sumich A, Huntington B, et al. The effectiveness of self-guided virtual-reality exposure therapy for public-speaking anxiety. Front Psychiatr 2021 Aug 19;12:694610 [FREE Full text] [doi: 10.3389/fpsyt.2021.694610] [Medline: 34489755]

86. Cieślik B, Juszko K, Kiper P, Szczepańska-Gieracha J. Immersive virtual reality as support for the mental health of elderly women: a randomized controlled trial. Virtual Real 2023 May 07;27(3):1-9 [FREE Full text] [doi: 10.1007/s10055-023-00797-w] [Medline: 37360811]
87. Berberyan A, Berberyan H, Alsina-Jurnet I. Virtual reality as anxiety management tool. Int J Cogn Res Sci, Eng Educ 2023 Dec 28;11(3):449-459. [doi: 10.23947/2334-8496-2023-11-3-449-459]
88. Mittmann G, Barnard A, Krammer I, Martins D, Dias J. LINA - a social augmented reality game around mental health, supporting real-world connection and sense of belonging for early adolescents. Proc ACM Hum-Comput Interact 2022 Oct 31;6(CHI PLAY):1-21. [doi: 10.1145/3549505]
89. Bellinger D, Wehrmann K, Rohde A, Schuppert M, Störk S, Flohr-Jost M, et al. The application of virtual reality exposure versus relaxation training in music performance anxiety: a randomized controlled study. BMC Psychiatry 2023 Aug 01;23(1):555 [FREE Full text] [doi: 10.1186/s12888-023-05040-z] [Medline: 37528410]
90. Turoń-Skrzypińska A, Tomska N, Mosiejczuk H, Rył A, Szylińska A, Marchelek-Myśliwiec M, et al. Impact of virtual reality exercises on anxiety and depression in hemodialysis. Sci Rep 2023 Aug 01;13(1):12435 [FREE Full text] [doi: 10.1038/s41598-023-39709-y] [Medline: 37528161]
91. El-Qirem FA, Malak MZ, Bani Salameh AK, Ali R, Alsswey A. Effect of virtual reality therapy on stress and anxiety symptoms, and physiological measures among University students: an experimental study in Jordan. Curr Psychol 2022 Apr 09;42(22):18997-19005. [doi: 10.1007/s12144-022-02939-w]
92. Kim M, Eom H, Kwon JH, Kyeong S, Kim J. Neural effects of a short-term virtual reality self-training program to reduce social anxiety. Psychol Med 2020 Sep 03;52(7):1296-1305. [doi: 10.1017/s0033291720003098]
93. Li Q, Sui H. Application of the perspective based on virtual reality technology to relieve anxiety in practice. 2022 Presented at: EITCE '22: 6th International Conference on Electronic Information Technology and Computer Engineering; October 21-23, 2022; Xiamen, China p. 870-874. [doi: 10.1145/3573428.3573585]
94. Bernaerts S, Bonroy B, Daems J, Sels R, Struyf D, Gies I, et al. Virtual reality for distraction and relaxation in a pediatric hospital setting: an interventional study with a mixed-methods design. Front Digit Health 2022 May 31;4:866119 [FREE Full text] [doi: 10.3389/fdgth.2022.866119] [Medline: 35712230]
95. Popa CO, Sava FA, Muresan S, Schenk A, Cojocaru CM, Muntean LM, et al. Standard CBT versus integrative and multimodal CBT assisted by virtual-reality for generalized anxiety disorder. Front Psychol 2022 Sep 28;13:1008981 [FREE Full text] [doi: 10.3389/fpsyg.2022.1008981] [Medline: 36248526]
96. Colombo D, Suso-Ribera C, Ortigosa-Beltrán I, Fernández-Álvarez J, García-Palacios A, Botella C. Behavioral activation through virtual reality for depression: a single case experimental design with multiple baselines. J Clin Med 2022 Feb 25;11(5):1262 [FREE Full text] [doi: 10.3390/jcm11051262] [Medline: 35268353]
97. Camacho-Conde JA, Legarra L, Bolinches VM, Cano P, Guasch M, Llanos-Torres M, et al. Assessment of attentional processes in patients with anxiety-depressive disorders using virtual reality. J Pers Med 2021 Dec 09;11(12):1341 [FREE Full text] [doi: 10.3390/jpm11121341] [Medline: 34945813]
98. Lundin J, Lundström A, Gulliksen J, Blendulf J, Ejeby K, Nyman H, et al. Using 360-degree videos for virtual reality exposure in CBT for panic disorder with agoraphobia: a feasibility study. Behav Cogn Psychother 2021 Nov 18;50(2):158-170. [doi: 10.1017/s1352465821000473]
99. Tan J, Kannis-Dymand L, Jones C. Examining the potential of VR program Tilt Brush in reducing anxiety. Virtual Real 2022 Nov 10;27(4):1-13 [FREE Full text] [doi: 10.1007/s10055-022-00711-w] [Medline: 36405877]
100. Modrego-Alarcón M, López-Del-Hoyo Y, García-Campayo J, Pérez-Aranda A, Navarro-Gil M, Beltrán-Ruiz M, et al. Efficacy of a mindfulness-based programme with and without virtual reality support to reduce stress in university students: a randomized controlled trial. Behav Res Ther 2021 Jul;142:103866. [doi: 10.1016/j.brat.2021.103866] [Medline: 33957506]
101. Arnfred B, Bang P, Hjorthøj C, Christensen CW, Stengaard Moeller K, Hvenegaard M, et al. Group cognitive behavioural therapy with virtual reality exposure versus group cognitive behavioural therapy with in vivo exposure for social anxiety disorder and agoraphobia: a protocol for a randomised clinical trial. BMJ Open 2022 Feb 02;12(2):e051147 [FREE Full text] [doi: 10.1136/bmjopen-2021-051147] [Medline: 35110313]
102. Binder FP, Pöhlchen D, Zwanzger P, Spoormaker VI. Facing your fear in immersive virtual reality: avoidance behavior in specific phobia. Front Behav Neurosci 2022 Apr 27;16:827673 [FREE Full text] [doi: 10.3389/fnbeh.2022.827673] [Medline: 35571283]
103. Hildebrand AS, Roesmann K, Planert J, Machulska A, Otto E, Klucken T. Self-guided virtual reality therapy for social anxiety disorder: a study protocol for a randomized controlled trial. Trials 2022 May 12;23(1):395 [FREE Full text] [doi: 10.1186/s13063-022-06320-x] [Medline: 35549980]
104. Harrison K, Potts E, King AC, Braun-Trocchio R. The effectiveness of virtual reality on anxiety and performance in female soccer players. Sports (Basel) 2021 Dec 13;9(12):167 [FREE Full text] [doi: 10.3390/sports9120167] [Medline: 34941805]
105. Toffolo MBJ, Fehribach JR, van Klaveren CPBJ, Cornelisz I, van Straten A, van Gelder J, et al. Automated app-based augmented reality cognitive behavioral therapy for spider phobia: study protocol for a randomized controlled trial. PLoS One 2022 Jul 13;17(7):e0271175 [FREE Full text] [doi: 10.1371/journal.pone.0271175] [Medline: 35830423]

106. Rubin M, Muller K, Hayhoe MM, Telch MJ. Attention guidance augmentation of virtual reality exposure therapy for social anxiety disorder: a pilot randomized controlled trial. Cogn Behav Ther 2022 Sep 06;51(5):371-387 [FREE Full text] [doi: 10.1080/16506073.2022.2053882] [Medline: 35383544]
107. Shin B, Oh J, Kim B, Kim HE, Kim H, Kim S, et al. Effectiveness of self-guided virtual reality-based cognitive behavioral therapy for panic disorder: randomized controlled trial. JMIR Ment Health 2021 Nov 22;8(11):e30590 [FREE Full text] [doi: 10.2196/30590] [Medline: 34813486]
108. Na H, Park S, Dong S. Mixed reality-based interaction between human and virtual cat for mental stress management. Sensors (Basel) 2022 Feb 03;22(3):1159 [FREE Full text] [doi: 10.3390/s22031159] [Medline: 35161904]
109. Estrella-Juarez F, Requena-Mullor M, Garcia-Gonzalez J, Lopez-Villen A, Alarcon-Rodriguez R. Effect of virtual reality and music therapy on the physiologic parameters of pregnant women and fetuses and on anxiety levels: a randomized controlled trial. J Midwifery Womens Health 2023 Jan 16;68(1):35-43. [doi: 10.1111/jmwh.13413] [Medline: 36383473]
110. Donker T, Fehribach J, van Klaveren C, Cornelisz I, Toffolo MBJ, van Straten A, et al. Automated mobile virtual reality cognitive behavior therapy for aviophobia in a natural setting: a randomized controlled trial. Psychol Med 2022 Nov 25;53(13):6232-6241. [doi: 10.1017/s0033291722003531]
111. Wei S, Freeman D, Rovira A. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. Sci Rep 2023 Jul 17;13(1):11517 [FREE Full text] [doi: 10.1038/s41598-023-38499-7] [Medline: 37460586]
112. Orr E, Arbel T, Levy M, Sela Y, Weissberger O, Liran O, et al. Virtual reality in the management of stress and anxiety disorders: a retrospective analysis of 61 people treated in the metaverse. Heliyon 2023 Jul;9(7):e17870 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e17870] [Medline: 37483756]
113. Zeka F, Clemmensen L, Arnfred BT, Nordentoft M, Glenthøj LB. Examination of gaze behaviour in social anxiety disorder using a virtual reality eye-tracking paradigm: protocol for a casecontrol study. BMJ Open 2023 Aug 24;13(8):e071927 [FREE Full text] [doi: 10.1136/bmjopen-2023-071927] [Medline: 37620268]
114. Bond J, Kenny A, Pinfold V, Couperthwaite L, gameChange Lived Experience Advisory Panel, Kabir T, et al. A safe place to learn: peer research qualitative investigation of gameChange virtual reality therapy. JMIR Serious Games 2023 Jan 16;11:e38065 [FREE Full text] [doi: 10.2196/38065] [Medline: 36645707]
115. Lacey C, Frampton C, Beaglehole B. oVRcome - self-guided virtual reality for specific phobias: a randomised controlled trial. Aust N Z J Psychiatr 2023 May 11;57(5):736-744. [doi: 10.1177/00048674221110779] [Medline: 35818296]
116. Levy AN, Nittas V, Wray TB. Patient perceptions of in vivo versus virtual reality exposures for the treatment of anxiety disorders: cross-sectional survey study. JMIR Form Res 2023 Oct 16;7:e47443 [FREE Full text] [doi: 10.2196/47443] [Medline: 37843884]
117. Pan J. Internet-based cognitive behavioral therapy and virtual reality exposure therapy for social anxiety disorder: protocol for a randomized controlled trial in Hong Kong. JMIR Res Protoc 2023 Sep 14;12:e48437 [FREE Full text] [doi: 10.2196/48437] [Medline: 37707944]
118. Cinalioglu K, Lavín P, Bein M, Lesage M, Gruber J, Se J, et al. Effects of virtual reality guided meditation in older adults: the protocol of a pilot randomized controlled trial. Front Psychol 2023 Jul 28;14:1083219 [FREE Full text] [doi: 10.3389/fpsyg.2023.1083219] [Medline: 37575420]
119. Panchal N, Saunders H, Rudowitz R, Cox C. The implications of COVID-19 for mental health and substance use. KFF. 2023 Mar 20. URL: https://tinyurl.com/2r44kx9b [accessed 2024-06-07]
120. World Health Organization. Mental health and COVID-19: early evidence of the pandemic's impact: scientific brief. World Health Organization. 2022 Mar 2. URL: https://iris.who.int/handle/10665/352189 [accessed 2024-06-07]
121. Pfefferbaum B, North CS. Mental health and the COVID-19 pandemic. N Engl J Med 2020 Aug 06;383(6):510-512. [doi: 10.1056/nejmp2008017]
122. Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 2020 Apr;147:103778. [doi: 10.1016/j.compedu.2019.103778]
123. Bell IH, Nicholas J, Alvarez-Jimenez M, Thompson A, Valmaggia L. Virtual reality as a clinical tool in mental health research and practice . Dialogues Clin Neurosci 2022 Apr 01;22(2):169-177. [doi: 10.31887/dcns.2020.22.2/lvalmaggia]
124. Baghaei N, Chitale V, Hlasnik A, Stemmet L, Liang H, Porter R. Virtual reality for supporting the treatment of depression and anxiety: scoping review. JMIR Ment Health 2021 Sep 23;8(9):e29681 [FREE Full text] [doi: 10.2196/29681] [Medline: 34554097]
125. Liu Z, Ren L, Xiao C, Zhang K, Demian P. Virtual reality aided therapy towards Health 4.0: a two-decade bibliometric analysis. Int J Environ Res Public Health 2022 Jan 28;19(3):1525 [FREE Full text] [doi: 10.3390/ijerph19031525] [Medline: 35162546]

Abbreviations

AR: augmented reality
CBT: cognitive behavioral therapy

HMD: head-mounted display
MR: mixed reality
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VR: virtual reality
WHO: World Health Organization
XR: extended reality

Edited by T Leung, G Eysenbach; submitted 31.03.22; peer-reviewed by N Titov, XZ Hu; comments to author 09.06.22; revised version received 05.09.22; accepted 24.03.24; published 24.07.24.

Please cite as:
Omisore OM, Odenigbo I, Orji J, Beltran AIH, Meier S, Baghaei N, Orji R
Extended Reality for Mental Health Evaluation: Scoping Review
JMIR Serious Games 2024;12:e38413
URL: https://games.jmir.org/2024/1/e38413
doi:10.2196/38413
PMID:

Review

Effects of Electronic Serious Games on Older Adults With Alzheimer's Disease and Mild Cognitive Impairment: Systematic Review With Meta-Analysis of Randomized Controlled Trials

Xinyi Zuo[1*], EdD; Yong Tang[1], PhD; Yifang Chen[2*], DPS; Zhimiao Zhou[3]

[1]Sociology Department, School of Government, Shenzhen University, Shenzhen, Guangdong, China

[2]Institution of Policy Studies, Lingnan University, Tuen Mun, Hong Kong SAR, China (Hong Kong)

[3]Shenzhen Senior High School, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Yong Tang, PhD
Sociology Department
School of Government
Shenzhen University
1066 Xueyuan Avenue, Nanshan District
Shenzhen, Guangdong, 518055
China
Phone: 86 13670191924
Email: tangyong@szu.edu.cn

Related Article:

This is a corrected version. See correction statement: https://games.jmir.org/2024/1/e65184

Abstract

Background: Serious games (SGs) are nonpharmacological interventions that are widely applied among older adults. To date, no evidence has been published regarding the effect of digital SGs on cognitive ability, daily behavioral capacity, or depression in older adults with Alzheimer's disease (AD) and mild cognitive impairment (MCI).

Objective: This study aimed to assess the effect of SGs on older adults with AD and MCI by summarizing and pooling the results of previous studies.

Methods: This meta-analysis examined the effectiveness of digital SGs in improving cognitive ability, enhancing daily behavioral capacity, and alleviating depression in older adults with AD and MCI. We searched the following databases up to December 31, 2023, to identify relevant high-quality randomized controlled trials (RCTs): PubMed, Embase, Web of Science, Scopus, and Cochrane Library. Stata 15.1 and Review Manager 5.3 were used to screen the 14 studies, extract data, code the data, and perform meta-analysis. Mean differences and standardized mean differences (SMDs) with 95% CIs were used to calculate continuous variables. The Cochrane risk-of-bias assessment tool was used to evaluate the risk of bias. Eligibility criteria were developed in accordance with the Population, Intervention, Comparison, Outcomes, and Study Design framework: (1) population (older adults with AD and MCI), (2) intervention (digital SG intervention), (3) comparison (digital SG intervention vs routine health care), (4) outcomes (cognitive ability, daily behavioral capacity, and depression), and (5) study or research design (RCT). Sensitivity analysis was performed, and a funnel plot was constructed.

Results: From January 2017 to December 2023, we enrolled 714 individuals across 14 RCTs, with 374 (52.4%) in the severe game group using digital SGs and 340 (47.6%) in the control group using traditional methods. The results of our meta-analysis indicated that using digital SGs in older adults with AD and MCI is more effective than traditional training methods in several key areas. Specifically, digital SG therapy significantly increased cognitive ability, as found in the Mini-Mental State Examination (SMD 2.11, 95% CI 1.42-2.80; $P<.001$) and the Montreal Cognitive Assessment (SMD 2.75, 95% CI 1.98-3.51; $P<.001$), significantly increased daily behavioral capacity (SMD 0.53, 95% CI 0.06-0.99; $P=.03$), and significantly reduced depression (SMD –2.08, 95% CI –2.94 to –1.22; $P<.001$) in older adults with AD and MCI. No publication bias was detected based on the results of Begg and Egger tests.

Conclusions: Digital SGs offer a viable and effective nonpharmacological approach for older adults with AD and MCI, yielding better results compared to traditional formats. However, caution is warranted in interpreting these findings due to limited RCTs, small sample sizes, and low-quality meta-analyzed evidence.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews: CRDCRD42023486090; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=486090

(JMIR Serious Games 2024;12:e55785) doi:10.2196/55785

KEYWORDS

digital serious games; cognitive ability; daily behavioral capacity; mental health; depression; older adults with AD and MCI; AD; Alzheimer's disease; MD; mild cognitive impairment; systematic review; meta-analysis

Introduction

Alzheimer's disease (AD) and mild cognitive impairment (MCI) are among the most common diseases affecting older adults. As the worldwide population is aging, the incidence of AD and MCI is gradually increasing, leading to major challenges for older adults. AD and MCI are being increasingly researched worldwide because of their adverse effects on cognitive ability, daily behavioral capacity, and depression among older adults. Currently, more than 55 million people worldwide have dementia [1]. Dementia is a progressive degenerative disease of the central nervous system that metastasizes through different clinical stages [2] and is characterized by abnormal behavior, declining cognitive function, and declining daily living ability [3]. Dementia affects multiple aspects of cognitive function, including complex attention, memory and learning, perceptual-motor skills, social cognition, and language [4]. AD is the most common form of dementia and may account for 60%-70% (38.77 million) of the 55 million cases of dementia [5]. MCI is an intermediate stage between dementia and normal aging [6]; clinicians and researchers consider the stage of MCI to be a "window" for potentially delaying the progression into dementia [7]. Moreover, the number of MCI cases among individuals above the age of 60 years is 5.82 million [8]. In 2019, the worldwide economic cost of dementia amounted to US $1.3 trillion, of which approximately 50% (US $0.65 trillion) involved care provided by informal caregivers (eg, family members and close friends), who provided an average of 5 hours of care and supervision per day [9]. China has the largest number of patients with dementia worldwide, which leads to a huge burden on the health care and public system [10]. Approximately 13.75 million (25%) of 55 million people with dementia are Chinese [11], and thus, this country ranks first in the world. The results of a recent national cross-sectional study showed that there are 16.07 million people aged 60 years and over in China who have dementia; 9.83 million of these individuals have AD [11]. Presently, there is no known cure for dementia; health care and medication can only slow the disease progression [12].

When older people get MCI and AD, they experience a range of painful experiences related to this particular stage of life, such as depression [7,13], decreased cognitive ability [7,14], and decreased daily living ability [15,16]. Sudden illness can cause patients to experience both physical and psychological pain [17]. A significant proportion of the population with MCI and AD reports higher levels of perceived [18]. Depression is characterized by threats that exceed an individual's coping abilities, thus resulting in a feeling of being overwhelmed [19]. A growing body of evidence suggests that several important dimensions of life are frequently sacrificed when individuals get MCI and AD, and cognitive ability, daily behavioral capacity and depression are often neglected. Therefore, digital serious games (SGs) for older adults with MCI and AD may be crucial in the early intervention and prevention of these mental and physical disorders [20]. According to a previous report, even in worse cases, mental and physical disorders are risk factors for MCI and AD [21]. Patients with MCI and AD are completely dependent on caregivers, have markedly impaired memory function, are unable to take care of themselves in daily life, exhibit incontinence, and may even be completely unable to speak and have hemiplegia [22]. MCI and AD even cause an increase in the risk of mortality [23], which significantly damages people's health and imposes a growing burden on individuals and the entire society. Therefore, it is highly emergent to develop nondrug therapies and use them in the field of social work for older adults. This may slow the progression of cognitive disability in individuals with MCI and AD to combat personal and socioeconomic problems, to address the alarming increase in the incidences of MCI and dementia worldwide, and to overcome the lack of effective drug treatments for these diseases.

There are many therapeutic methods for treating MCI and AD in older adults. Game therapy has been recommended as a nonpharmacological treatment method due to its limited side effects. According to recent meta-analyses [24-26], game therapy could increase cognitive ability [26] and daily performance ability [27,28] and alleviate depression [29,30] in older adults with MCI and AD. Game therapy refers to an intervention method in which games are used as a therapeutic medium. This therapy originated from a child psychoanalysis program developed by the famous psychologist Sigmund Freud. He observed that there is a relationship between early childhood development and children's games. In the discussion of children's control of emotions through games, Freud presented the rationale for game therapy [31]. After nearly a century of clinical development, the therapeutic function of games has been proven in practice [32,33]. SGs may take many forms, depending on their treatment, such as (1) cognitive training games, which are mainly used to improve or maintain users' cognitive abilities (eg, memory, executive function, and memory) [34-37]; (2) sports games or video games, in which part of the game needs to include exercise [38,39]; (3) biofeedback games, which use electronic sensors connected to the participant's body to receive information about their physical status (eg, an electrocardiograph sensor), with the aim of adjusting some bodily functions (eg, heart rate) [39-46]; and (4) cognitive behavioral therapy (CBT)–based games that provide computerized computer-based training. Due to the popularity of smart devices, handheld computers, personal computers, video game consoles, tablets, and smartphones, there is a wide variety of SGs [47,48]. Digital SGs, as defined in this paper, are SGs played through digital media. Digitally designed games help promote mental health [49], emotions [50], well-being [51], social function [51], and cognitive flexibility [52] compared to traditional cognitive training and exercise [53].

With respect to the development of game therapy theory, many studies have shown that game therapy is applicable for older adults [54]. The introduction of virtual reality technology has recently provided a new way to implement environmental enrichment in the context of clinical practice, which improves the physical health of patients with MCI and AD [55], treats chronic disease [56], increases cognitive ability [57] and daily behavioral capacity [50,52], alleviates depression [58], and increases self-esteem [59] and cohesion [60] by combining virtual immersion with cognitive stimulation. In conclusion, SG therapy is regarded as a promising treatment strategy for older adults with dementia and MCI in terms of increasing cognitive ability, improving daily behavioral capacity, and alleviating mental health issues. Game therapy is therefore increasingly used in gerontology, with the goal of providing psychosocial stabilization and support.

There are a few research gaps. First, there are inconsistent clinical results regarding the effects of SG therapy on cognitive performance, daily performance, and depression. According to Saragih et al [21] and Ning et al [60], SGs have been shown to improve cognitive function. However, Huang et al [61] and Kleschnitzki et al [62] reported that no significant changes are found in cognitive function or global status after virtual reality intervention. Wang et al [63] discovered that SGs have a positive impact on executive function. However, Abd-Alrazaq et al [52] found that there is no difference between adaptive SGs and nonadaptive SGs in improving executive function, although daily behavioral capacity is particularly related to executive function, such as scheduling appointments, making monthly payments, managing the household economy, shopping, or taking the bus. This means the effect of SGs in improving daily behavioral capacity is yet to be explored. Bai et al [50] compared SGs with intervention-free practices and found that game-based interventions lead to clinically and statistically significant improvements in depression severity. Bojan et al [64] discovered that SGs can cause a significant increase in positive emotions and a decrease in negative emotions [11,65]. However, according to Fitzgerald et al [66], whether SGs can improve depression in some cases is worth exploring. Therefore, there are inconsistent clinical results regarding the effects of SG therapy on cognitive performance, daily performance, and depression. A more comprehensive review is needed to systematically evaluate the effects of SGs on the physical and mental health of older adults with AD and MCI. Second, the previous literature only focused on SGs for a type of disease for older adults. Manca et al [67] observed the positive impact of SGs with humanoid robots on older adults with MCI. Pacheco et al [68] found that exergames improve balance and mobility in older adults to keep performing balance exercises. Yu and Chan [69] discovered video game training can effectively improve the cognition of older adults. This study investigated the effects of SGs on the cognitive ability, daily behavioral capacity, and depression of older adults with MCI and AD simultaneously. Third, it is important to investigate not only the effects of SGs but also the variables that may influence their effectiveness, including the duration of treatment and the target population. In this research, we used weeks and continental plates (within and outside Europe) to divide the subgroup analysis, which was not available in previous studies. In view of this, we conducted a systematic review and meta-analysis to comprehensively evaluate the effect of digital SGs on cognitive ability, daily behavioral capacity, and depression of older adults (≥60 years old) with MCI and AD.

Methods

Data Sources and Search Strategy

This review was conducted in strict accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [70]; also see Multimedia Appendices 1 and 2 [71]. We registered the study in PROSPERO (Prospective Register of Systematic Reviews; CRD42023486090). We used the PICOS (Population, Intervention, Comparison, Outcomes, and Study Design) framework [72] to structure study inclusion and exclusion criteria (Table 1). Our population of interest had at least 1 indicator indicating cognitive ability, behavioral ability, or emotion-related ability, and their age was ≥60 years. We sought studies describing interventions with digital SGs implemented among older adults with cognitive impairment in an experimental group (severe game group [SGG]), including exergames or video games, cognitive training games, computerized CBT-based games, and biofeedback games. There were no restrictions regarding the time of intervention. For comparison, we had a control group (CG) that received a different intervention, such as routine health care, traditional Tai Chi, or music. Our outcomes of interest included cognitive ability, daily behavioral capacity, and depression in older adults with cognitive impairment. We excluded studies that presented reviews, non–randomized controlled trials (RCTs), and papers for which the full text was unavailable. All included studies were published in Chinese or English. Studies in which the results were interpreted from the perspective of older adults with cognitive impairment were deemed eligible.

Table 1. Inclusion and exclusion criteria.

Element	Inclusion criteria	Exclusion criteria
Population	• At least 1 indicator representing cognitive ability, behavioral ability, or emotion-related ability • Age≥60 years	N/A[a]
Intervention	• Digital SGs[b] (eg, exergames, video games, cognitive training games, computerized CBT[c]-based games, biofeedback games)	N/A
Comparison	• Different interventions (eg, routine health care, traditional Tai Chi, music)	N/A
Outcomes	• Cognitive ability • Daily behavioral capacity • Depression	N/A
Study design	• RCT[d]	• Reviews • Non-RCTs • Papers with full text unavailable
Language	• Chinese or English papers	• Not Chinese or English papers

[a]N/A: not applicable.
[b]SG: serious game.
[c]CBT: cognitive behavioral therapy.
[d]RCT: randomized controlled trial.

Search Strategy

The following databases were searched to identify the relevant literature: PubMed, Embase, Web of Science, Scopus, and Cochrane Library. The following search terms were used for the search: ("Cognitive Ability" OR "Daily Behavior Ability" OR "Mental Health" OR "Depression") AND ("Serious Games Based" OR "Serious Games" OR "Web-Based or Mobile") AND ("Old Adults in Cognitive Impairment" OR "Old Adults" OR "Older People") AND ("A pilot study" OR "Randomized Controlled Trial" OR "RCT"). Detailed search strategies for each database are given in Multimedia Appendix 3. The keywords were used to search for papers published from January 1, 2017, to December 31, 2023. The snowball method was adopted to search the reference lists of the included studies. Additionally, the reference lists of the included studies were manually searched to identify eligible papers. Unpublished academic literature was not considered eligible. The reference lists and relevant reviews were then screened to identify any pertinent studies. To identify additional relevant publications, we also retrieved gray literature [73].

Data Extraction

Four researchers (authors XZ, YT, YC, and ZZ) screened all the literature for inclusion. After removing duplicates using EndNote X9, all studies were independently screened based on titles and abstracts. Next, we carefully read the full texts of the remaining papers in accordance with the inclusion and exclusion criteria. Finally, data were extracted from the included studies. The co-first author (XZ) used the modified version of the data extraction table in the *Systematic Review Manual of Cochrane Interventions* to extract data. A modified version of the intervention description and replication was using following the template for intervention description and replication (TIDieR) list and guidelines [74] to extract data of the intervention, design, and delivery features, as defined in the introduction. First, data were extracted by the co-first author (XZ). Next, the second author (ZZ) checked the accuracy of the process. Finally, the co-first author (YC), through thematic analysis [75], extracted the following data from the list of 14 studies [34-47]:

- Study characteristics: name of author, publication year, country, publication type, and study design
- Subject characteristics: sample size, participant mean age, participant age, Mini-Mental State Examination (MMSE) score, participant health conditions, and population group
- Intervention details: setting, intervention type, delivery mode design (how), content (ie, procedures, materials, process, activities [what]), and delivery details (ie, intervention delivery format, intervention duration, session length, session frequency, intensity)
- Study quality: Jadad score
- Main outcomes: reported results, conclusion, measurement, and follow-up

Risk-of-Bias Assessment and Quality Assessment

Using the Cochrane Collaboration risk-of-bias tool [76], we independently assessed the risk of bias. Additionally, we assessed the quality of the 14 studies using the Jadad scale [77]. The risk-of-bias tool assesses 7 domains: (1) generation of a random sequence, (2) allocation concealment, (3) subject and experimenter blinding, (4) outcome assessor blinding, (5) resulting data integration, (6) selective reports, and (7) other bias risk. Each study was categorized as having a low, high, or unclear risk of bias. We also performed Begg and Egger tests to evaluate the degree of publication bias [78].

Statistical Analysis

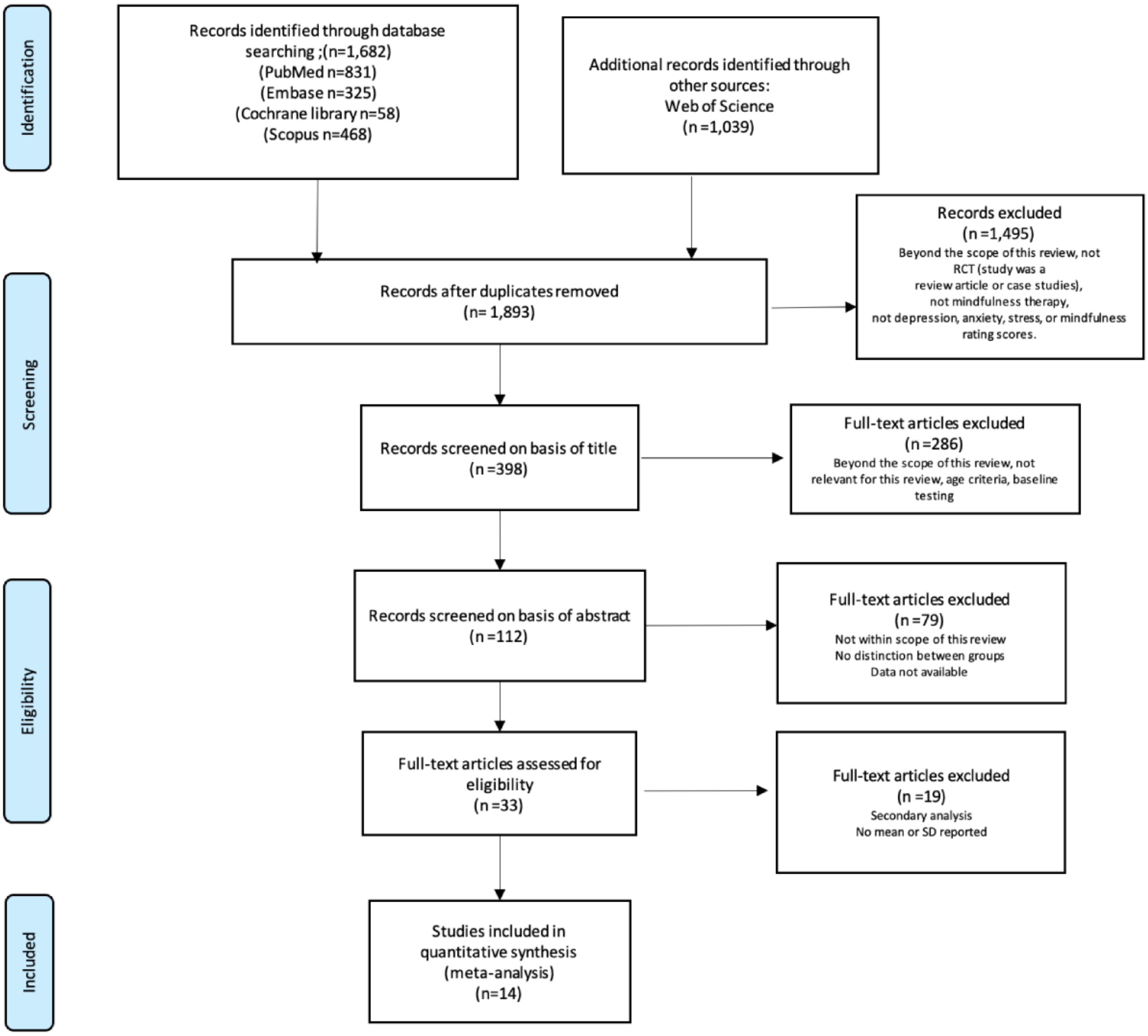
Stata 15.1 and Review Manager 5.3 were used to conduct data analysis. Forest plots were constructed to intuitively illustrate the results. In the included literature, the outcomes were measured as continuous variables, and the same indicator was assessed with different tools. These indicators were presented as standard mean differences (SMDs). α=.05 was considered statistically significant. Heterogeneity was assessed with I^2 statistics, which were classified as high (>75%), moderate (50%-75%), or low (<50%) [79]. In cases of high heterogeneity, sensitivity analysis was performed using the leave-1-out method to identify the sources of heterogeneity. In the analysis, numerous weeks were compared as subpoints to check the results. The Begg test [80] and Egger test [81] were used to assess publication bias. P<.05 indicated significant results. When a meta-analysis included at least 10 studies, publication bias was evaluated with a funnel plot [82,83]. In this study, 95% CI and the SMD were examined [84]. Overall, P<.05 indicated statistically significant effects. Subgroup analyses were performed based on the country of intervention (outside Europe or not) and the length of intervention (weeks).

Results

Study Selection

Figure 1 shows the study selection process. After searching 13 databases, 4260 relevant records were identified. For duplicate removal, we imported all the studies into EndNote X8 [85]. After removing 1893 (44.4%) duplicates and eliminating 2352 (55.2%) papers via a strict screening process, 14 (0.3%) RCTs [34-47] from before December 2023 involving 714 participants were finally included. Studies were excluded because they did not report SDs [26,86-88], were review papers [89,90], or were not within the scope of this meta-analysis [91-94]. All the included studies reported that digital SGs have a positive impact on cognitive ability, daily behavioral capacity, and depression. The primary outcomes of interest were an increase in cognitive ability scores (assessed with the MMSE and the MoCA), an increase in daily behavioral capacity indicators (assessed with the Barthel Index, the Katz Index, the Lawton and Brody Index, the Korean version of the Modified Barthel Index [K-MBI], and the Activities of Daily Living [ADL] scale), and a reduction in depression (assessed with the Goldberg Anxiety and Depression Scale [EADG], the 15-item Geriatric Depression Scale [EDG-15], and the Cornell Scale for Depression in Dementia [CSDD]).

Figure 1. Study selection flowchart based on PRISMA guidelines, 2020. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Study Characteristics

Table 2 and Multimedia Appendix 4 [34-47] show the overall characteristics of the included studies. All 14 studies were published before 2023. The sample sizes ranged from 17 to 112, and 714 older adults aged >60 years were enrolled in the included studies, including 340 (47.6%) subjects in the SGG and 374 (52.4%) subjects in the CG. All participants were older adults with cognitive impairment, but they had not been diagnosed with psychiatric disorders. All interventions were based on digital SGs, and their durations ranged from 4 weeks to 12 weeks. The interventions lasted from 45 minutes to 5 hours each week. Both group training and individual training methods were used [34-47]. All 14 studies [34-47] were subdivided into digital SG training methods (exergames or video games, cognitive training games, computerized CBT games, and biofeedback games).

Table 2. Characteristics and findings of the included studies (N=14)[a].

Study	Location	Sample size: SGG[b] and CG[c], n/N (%)	Age (years): SGG and CG, range (average)/mean (SD)	Disease	Time frame (weeks)	Reported results and conclusion
Yang et al [34]	Republic of Korea	SGG: 10/20 (50) CG: 10/20 (50)	SGG: 64-78 (71.1) CG: 61-78 (69.9)	Dementia	4	Result: Postintervention, the K-MMSE[d] score of the SGG was higher than that of the CG. Conclusion: Computer-based CBT[e] training might have beneficial effects on general cognitive function in the early stage of AD[f].
Kwan et al [35]	Hong Kong, China	SGG: 16/33 (48) CG: 17/33 (52)	SGG: (average): 70.5 CG: (average): 71	Dementia	12	Result: Postintervention, the MoCA[g] score of the SGG was higher than that of the CG. Conclusion: The mobile health (mHealth) intervention is feasible for implementation in older people with cognitive impairment and is effective at enhancing cognitive function.
Lim et al [36]	Republic of Korea	SGG: 12/24 (50) CG: 12/24 (50)	SGG: 69-81 (75.42) CG: 65-90 (73.33)	Dementia	4	Result: Postintervention, the K-MMSE, K-MoCA[h], and DQoL[i] scores of the SGG were higher than those of the CG. Conclusion: The research group showed significant improvement in cognitive function after training. Family SGs[j] are believed to help improve cognitive function.
Lee et al [37]	Republic of Korea	SGG: 10/20 (50) CG: 10/20 (50)	SGG: 65-83 (74.5) CG: 65-82 (74)	MCI, AD, dementia	3	Result: Postintervention, the K-MMSE and ADL[k] scores of the SGG were higher than those of the CG. Conclusion: Through this preliminary study, we verified that the newly developed computerized cognitive rehabilitation program is effective in improving cognitive function.
Yang [38]	Fujian, China	SGG: 44/88 (50) CG: 44/88 (50)	SGG: 62-83 (71.89) CG: 61-82 (71.63)	Dementia	12	Result: Postintervention, the MMSE[l] and ADL scores of the SGG were higher than those of the CG. Conclusion: Interactive games can improve the cognitive function and daily living ability of patients with senile dementia.
Van Santen et al [39]	Netherlands	SGG: 73/112 (65) CG: 39/112 (35)	SGG: 73-85 (79) CG: 72-86 (79)	Dementia	8	Result: Postintervention, the MMSE score of the SGG was higher than that of the CG. Conclusion: Sports games can improve the cognitive function of patients with senile dementia.
Zheng et al [40]	Zhejiang, China	SGG: 18/38 (47) CG: 20/38 (53)	SGG: 75-87 (81.74) CG: 78-89 (84.26)	Dementia	8	Result: Postintervention, compared to the CG, the MMSE score of the SGG was significantly higher and the CSDD[m] score was significantly lower. Conclusion: Somatic interactive game interventions can improve memory and language function and reduce depression in older patients with dementia, but the impact on cognitive function and ADL needs to be explored.
Wu et al [41]	Shandong, China	SGG: 47/94 (50) CG: 47/94 (50)	SGG: 61-80 (71.74) CG: 60-80 (72.65)	Dementia	12	Result: Postintervention, the MoCA score of the SGG was higher than that of the CG. Conclusion: Sensory stimulation combined with sensory interaction games can improve cognitive function.
Savulich et al [42]	United Kingdom	SGG: 21/42 (50) CG: 21/42 (50)	SGG: 67-82 (75.2) CG: 68-85 (76.9)	Dementia	4	Result: Postintervention, the MMSE score of the SGG was higher than that of the CG. Conclusion: Gamification maximizes engagement with cognitive training by increasing motivation and could complement pharmacological treatments for amnestic MCI[n] and mild AD.

Study	Location	Sample size: SGG[b] and CG[c], n/N (%)	Age (years): SGG and CG, range (average)/mean (SD)	Disease	Time frame (weeks)	Reported results and conclusion
Swinnen et al [43]	Belgium	SGG: 23/45 (51) CG: 22/45 (49)	SGG: 79-90 (84.7) CG: 78-91 (85.3)	Dementia	8	Result: Postintervention, compared to the CG, the MMSE, MoCA, and DQoL scores of the SG were higher and the CSDD score was significantly lower. Conclusion: Individually adapted exergame training improves lower extremity functioning, cognitive function, and depression in long-term-care facilities.
Liu et al [44]	Taipei, China	SGG: 16/33 (48) CG: 17/33 (52)	SGG: 68-80 (74.6) CG: 66-79 (73.4)	Dementia	12	Result: Postintervention, the MoCA score of the SGG was higher than that of the CG. Conclusion: Exergaming-based Tai Chi is comparable to traditional Tai Chi for enhancement of cognitive function.
Thapa et al [45]	Republic of Korea	SGG: 34/68 (50) CG: 34/68 (50)	SGG: 67-78 (72.6) CG: 67-78 (72.7)	MCI, dementia	8	Result: Postintervention, the MMSE score of the SGG was higher than that of the CG. Conclusion: Encouraging patients to perform virtual reality– and game-based training may be beneficial to prevent cognitive decline.
Oliveira et al [46]	Portugal	SGG: 10/17 (59) CG: 7/17 (41)	SGG: 77-88 (82.6) CG: 77-90 (84.14)	Dementia	5	Result: Postintervention, the MMSE and ADL scores of the SGG were higher than those of the CG. Conclusion: This approach is effective for neurocognitive stimulation in older adults with dementia, contributing to maintaining cognitive function in AD.
Jahouh et al [47]	Spain	SGG: 40/80 (50) CG: 40/80 (50)	SGG: 85.05 (8.63) CG: 83.25 (8.78)	MCI	8	Result: Postintervention, the MMSE and ADL scores of the SGG were significantly higher and the EGD-15[o] and EADG[p] scores were significantly lower than those of the CG. Conclusion: The Wii video console has a positive influence on older people by increasing cognitive ability, enhancing the level of ADL, and improving the psychological status.

[a]Publication type: journal paper; study design: randomized controlled trial (RCT); study population: older adults; setting: clinical.
[b]SGG: serious game group.
[c]CG: control group.
[d]K-MMSE: Korean version of the Mini-Mental State Examination.
[e]CBT: cognitive behavioral therapy.
[f]AD: Alzheimer's disease.
[g]MoCA: Montreal Cognitive Assessment.
[h]K-MoCA: Korean version of the Montreal Cognitive Assessment.
[i]DQoL: Dementia Quality of Life.
[j]SG: serious game.
[k]ADL: activities of daily living.
[l]MMSE: Mini-Mental State Examination.
[m]CSDD: Cornell Scale for Depression in Dementia.
[n]MCI: mild cognitive impairment.
[o]EGD-15: 15-item Geriatric Depression Scale.
[p]EADG: Goldberg Anxiety and Depression Scale.

Risk of Bias and Quality Assessment

Figures 2 and 3 show the results of the risk-of-bias assessment. All 14 papers [34-47] described the randomization methods in detail. The blinding method was detailed in 7 (50%) studies: 4 (57%) of these studies were double-blind trials [40,44,45,47], while 3 (43%) were single-blind trials [36,37,46]. The average Jadad score across all included studies was 4.7, indicating fair-to-mild quality.

Figure 2. Risk-of-bias graph.

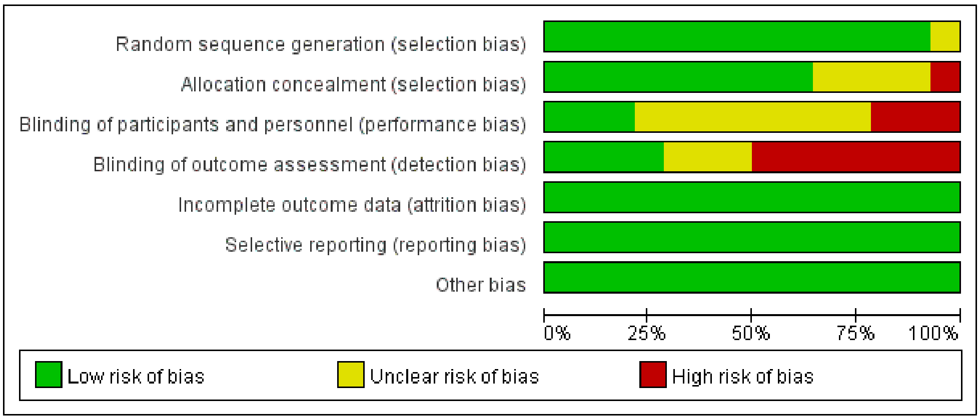
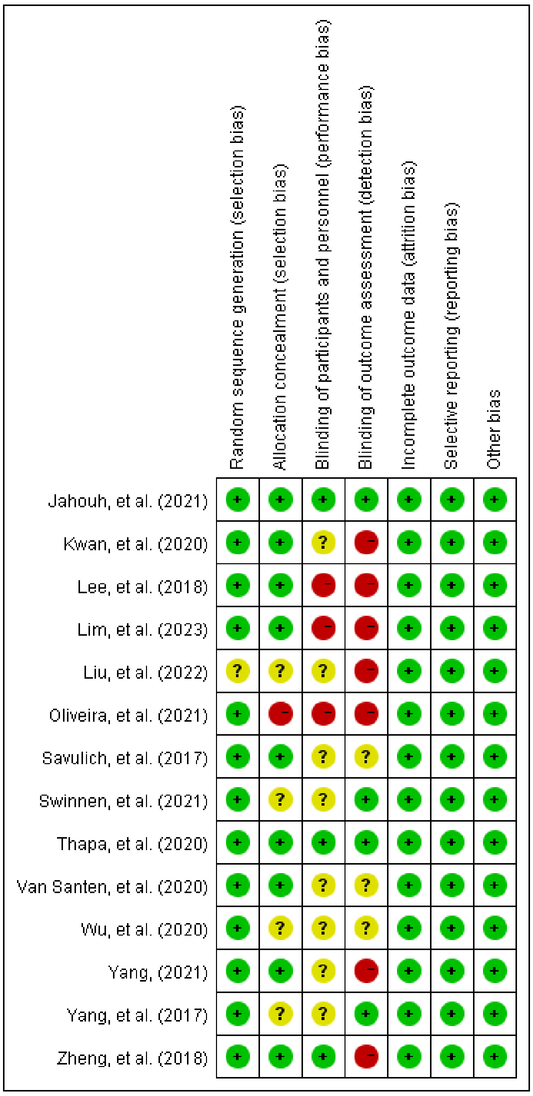


Figure 3. Risk-of-bias summary of the 14 studies included.



Meta-Analyses

Cognitive Ability Indicators

Of the 14 studies, 10 (71%) [34,36-40,42,45-47] involving 409 (57.3%) older adults (n=272, 66.5%, in the SGG and n=237, 33.5%, in the CG) assessed the effects of SGs on older adults' cognitive ability based on MMSE scores. Because different evaluation tools were used, the SMD was considered the pooled effect size measure. The pooled results revealed a low degree of heterogeneity among the studies (P=.23, I^2=23%), necessitating the use of a random effects model for the meta-analysis. Figure 4 demonstrates that cognitive ability scores in digital SGs significantly increased between the 2 groups under investigation (SMD 2.11, 95% CI 1.42-2.80; P<.001). Of the 14 studies, 5 (36%) [35,36,41,43,44] involving 229 (32.1%) older adults (n=114, 49.8%, in the SGG and n=115, 50.2%, in the CG) assessed the effects of SGs on older adults' cognitive ability based on MoCA scores. Because different evaluation tools were used, the SMD was considered the pooled effect size measure. The pooled results revealed a low degree of heterogeneity among the studies (P=.33, I^2=13%), necessitating the use of a random effects model for the meta-analysis. Figure 5 demonstrates that cognitive ability scores in digital SGs significantly increased between the 2 groups under investigation (SMD 2.75, 95% CI 1.98-3.51; P<.001).

Figure 4. Forest plot for MMSE. MMSE: Mini-Mental State Examination.

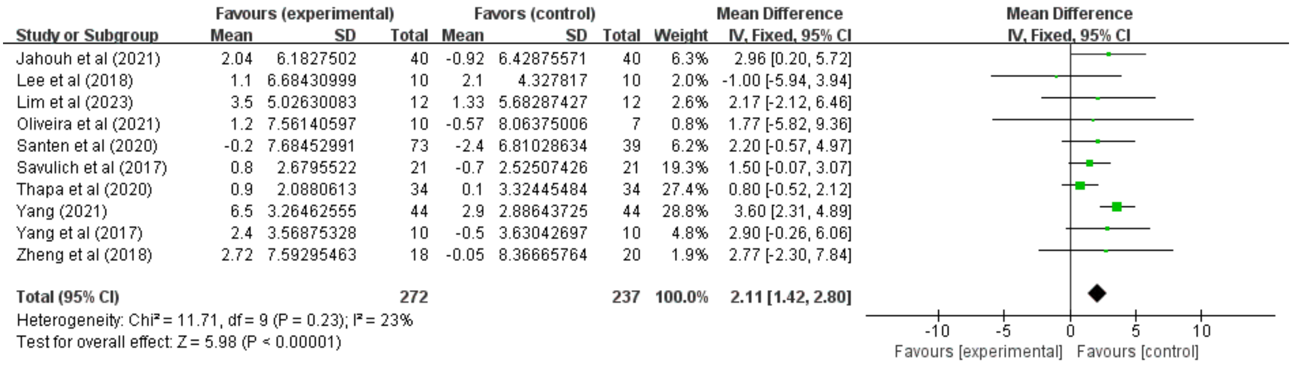
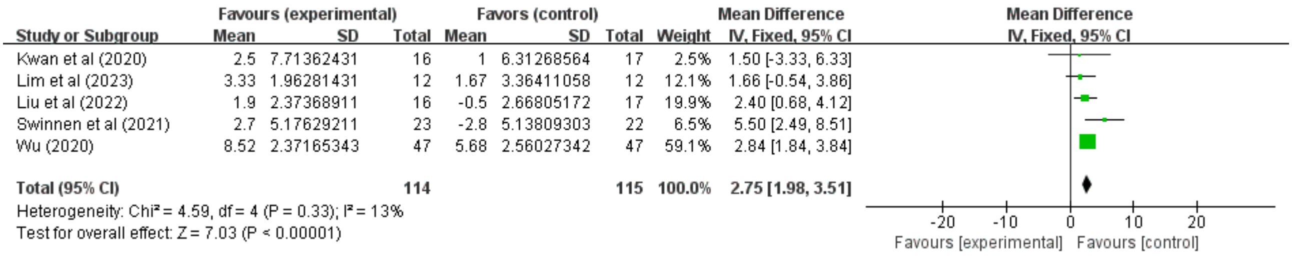


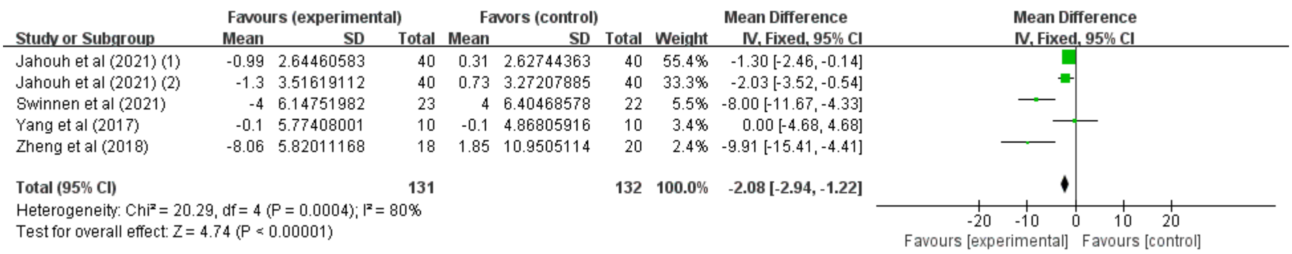
Figure 5. Forest plot for MoCA. MoCA: Montreal Cognitive Assessment.



Daily Behavioral Capacity Indicators

Of the 14 studies, 6 (43%) [37,38,40,43,46,47] involving 448 (62.7%) older adults (n=225, 50.2%, in the SGG and n=223, 49.8%, in the CG) assessed the effects of SGs on older adults' daily behavioral capacity based on the Barthel Index, the Katz Index, the Lawton and Brody Index, the K-MBI, and the ADL scale. Because different evaluation tools were used, the SMD was considered the pooled effect size measure. The pooled results revealed significant heterogeneity among the studies (P=.03, I^2=91%), necessitating the use of a random effects model for the meta-analysis. Figure 6 demonstrates that daily behavioral capacity scores in digital SGs significantly increased between the 2 groups under investigation (SMD 0.53, 95% CI 0.06-0.99; P=.03).

Figure 6. Forest plot for daily behavioral capacity indicators.

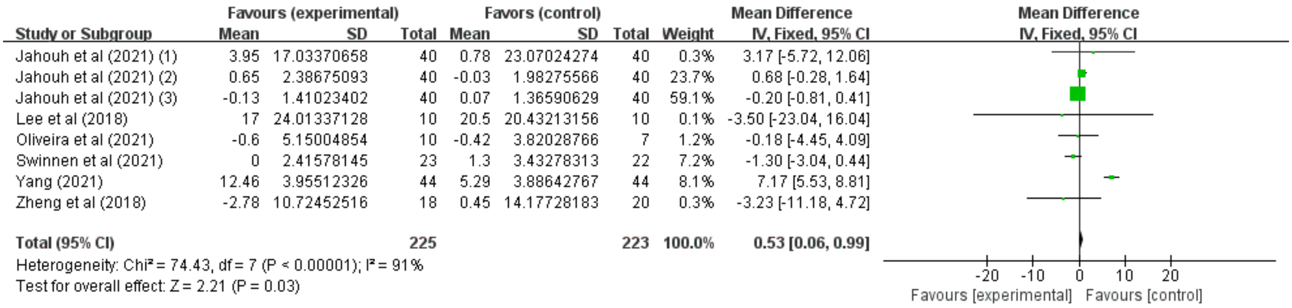


Depression Indicators

Of the 14 studies, 4 (29%) [34,40,43,47] involving 292 (40.9%) older adults (n=145, 49.7%, in the SGG and n=147, 50.3%, in the CG) assessed the effects of SGs on older adults' depression based on the EADG, the EDG-15, the Geriatric Depression Scale (GDS), and the CSDD. Because different evaluation tools were used, the SMD was considered the pooled effect size measure. The pooled results revealed a low degree of heterogeneity among the studies (P<.001, I^2=80%), necessitating the use of a random effects model for the meta-analysis. Figure 7 demonstrates that cognitive ability scores in digital SGs significantly increased between the 2 groups under investigation (SMD –2.08, 95% CI –2.94 to 1.22; P<.001).

In conclusion, the included studies reported scores on the MMSE, MoCA, the Barthel Index, the Katz Index, the Lawton and Brody Index, the K-MBI, the ADL scale, the EADG, the EDG-15, and the CSDD. A total of 3 outcomes were evaluated: cognitive ability, daily behavioral capacity, and depression. The SGG had significantly greater total MMSE scores than the CG (SMD 2.11, 95% CI 1.42-2.80; P=.23, I^2=23%), as shown in Figure 4. The SGG had significantly greater total MoCA scores than the CG (SMD 2.75, 95% CI 1.98-3.51; P=.33, I^2=13%), as shown in Figure 5. In addition, a reduction in daily behavioral capacity indicators was observed based on the Barthel Index, the Katz Index, the Lawton and Brody Index, the K-MBI, and the ADL scale. The SG led to increased daily behavioral capacity scores in the SGG (SMD 0.53, 95% CI 0.06-0.99; P=.03, I^2=91%), as shown in Figure 6, compared to the CG. The SGG also showed a decrease in depression based on the EADG, the EDG-15, and the CSDD (SMD –2.08, 95% CI –2.94 to –1.25; P<.001, I^2=80%), as shown in Figure 7.

Figure 7. Forest plot for depression indicators.



Subgroup Analyses

Subgroup analyses of daily behavioral capacity and depression scores were performed based on the country of intervention (outside or within Europe) and the duration of intervention (weeks). For daily behavioral capacity, significant differences were found in the SMD between the 2 subgroups: within Europe (P=.85) [43,46,47] and outside Europe (P<.001) [37,38,40]. SGs had a significant effect on the outcome in the outside-Europe group (SMD 6.68, 95% CI 5.08-8.28; P<.001), as shown in Figure 8. For depression, significant differences were found in the SMD between the 2 subgroups: within Europe (P<.001) [43,47] and outside Europe (P=.02) [34,40]. SGs had a significant effect on the outcome in the within-Europe group (SMD –2.08, 95% CI –2.94 to –1.22; P<.001), as shown in Figure 9. Of the 6 (43%) studies, 4 (67%) [38,40,46,47] reported a pooled effect of an intervention duration of ≥8 weeks (SMD 0.67, 95% CI 0.19-1.16; P=.007). For the remaining 2 (33%) studies [37,43], the pooled effect within the intervention period was as follows: SMD −1.32, 95% CI −3.05 to −0.42; P=.14. The effects of SGs on daily behavioral capacity during the 2 periods of intervention were significantly different from those in the CG (P<.001), as shown in Figure 10. For depression, no significant differences were found in the SMD between the 2 groups after >8 weeks (P=.005) [40,47] or <8 weeks (P=.008) [34,43], as shown in Figure 11. Among the 4 (29%) studies on depression indicators [34,40,43,47], 2 (50%) [40,47] had a combined effect in an intervention duration of ≥8 weeks (SMD −2.57, 95% CI −4.00 to −1.13; P=.005). For 2 (50%) studies [34,43] in which the intervention duration was <8 weeks, the combined effect within the intervention period was as follows: SMD −4.95, 95% CI −7.84 to −2.07; P=.008. An intervention period of 8 weeks or more had a significant effect on reducing cognitive impairment in older adults (Figure 11). Compared to with the CG, the effects of SGs on depression during the 2 periods of intervention were significantly different, and a significant difference between the 2 groups was found (P<.001), as shown in Figure 11.

Figure 8. Forest graph showing subgroup analysis of daily behavioral capacity based on the country of the intervention.

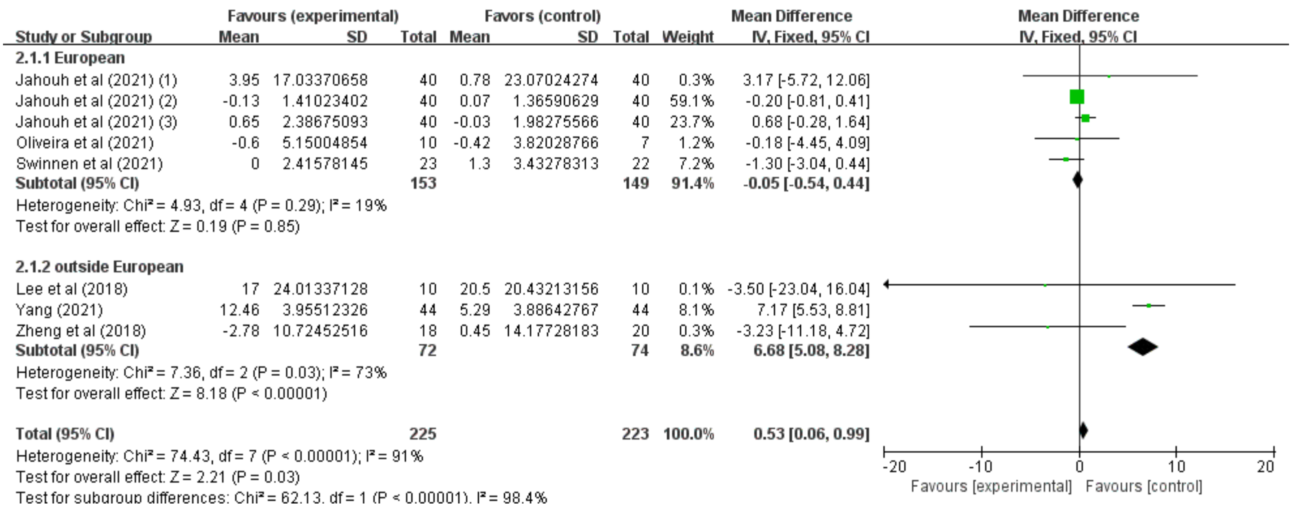


Figure 9. Forest graph showing subgroup analysis of depression based on the country of the intervention. CSDD: Cornell Scale for Depression in Dementia; GDS: Geriatric Depression Scale.

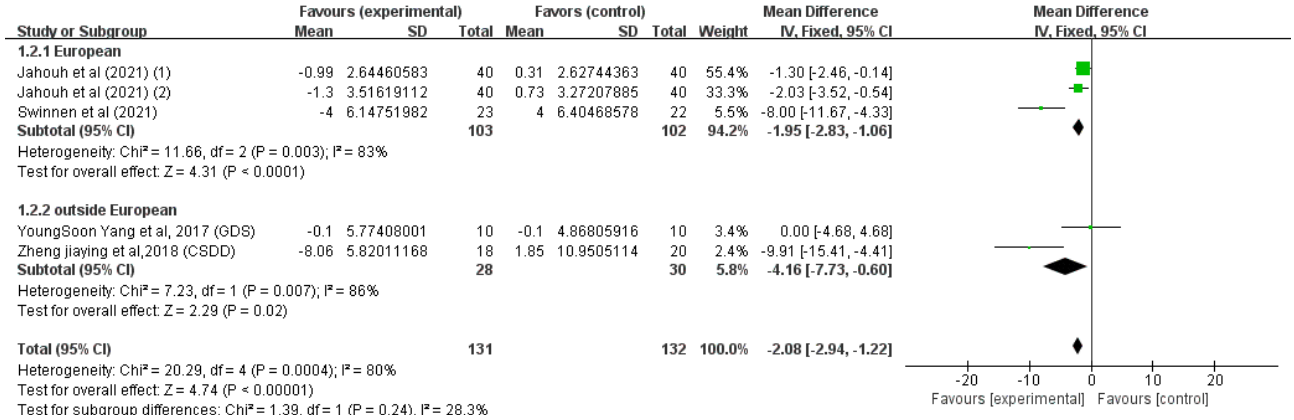


Figure 10. Forest graph showing subgroup analysis of daily behavioral capacity based on the duration of the intervention.

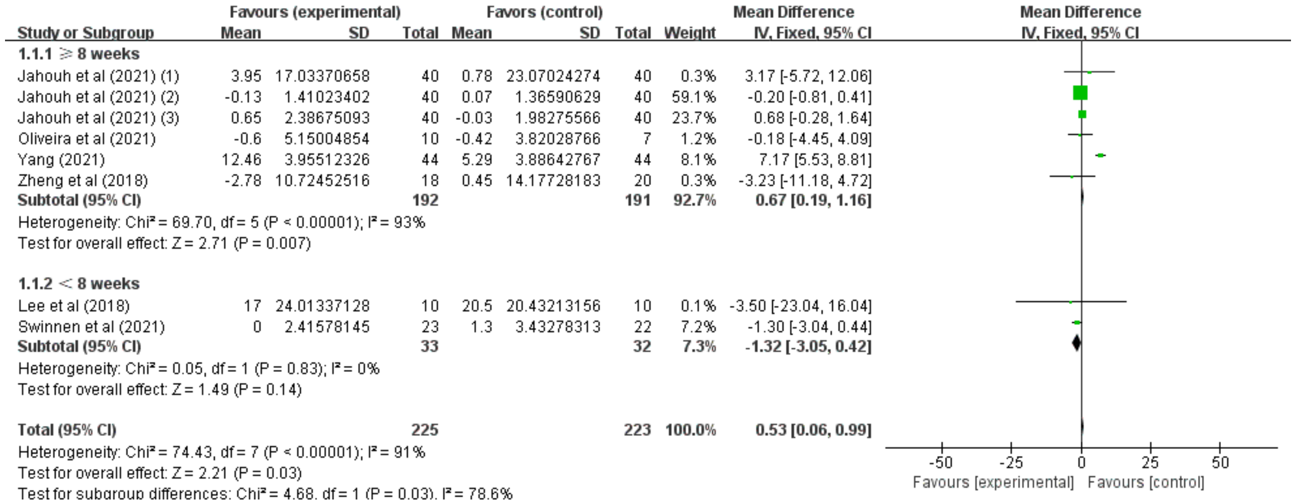
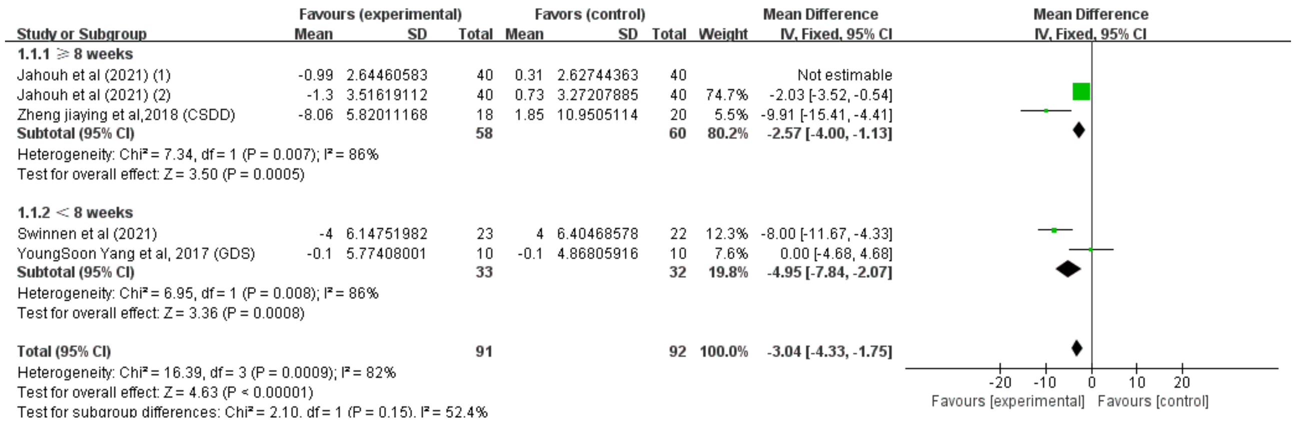


Figure 11. Forest graph showing subgroup analysis of depression based on the duration of the intervention. GDS: Geriatric Depression Scale.



Sensitivity Analysis

In the sensitivity analysis (Table 3), we excluded the study by Yang et al [38] and observed a significant change in heterogeneity from 91% to 0%. We hypothesized that the outcome of daily behavioral capacity indicators may have been the possible source of heterogeneity in this study. Additional possible reasons included the inclusion of patients, specific treatment modalities, and inconsistencies in clinical indicators between domestic and foreign countries. The patients included and specific treatment modalities, as well as the clinical indicators applied, in this study were as follows: Yang et al [38], Fujian (China), age 61-83 years; older adults with cognitive impairment, participated in somatosensory interactive games, and completed the MMSE and the ADL scale. All these different factors were possible sources of heterogeneity.

Table 3. Sensitivity analysis for daily behavioral capacity indicators after excluding a few studies.

Study excluded	Tool used	Heterogeneity analyses		
		χ^2 (*df*)	P value	I^2 (%)
Jahouh et al [47]	Barthel Index	74.09 (6)	<.001	92
Jahouh et al [47]	Katz Index	61.17 (6)	<.001	90
Jahouh et al [47]	Lawton and Brody Index	74.02 (6)	<.001	92
Lee et al [37]	K-MBI[a]	74.01 (6)	<.001	92
Oliveira et al [46]	IADL[b]	74.07 (6)	<.001	92
Swinnen et al [43]	ADL[c]	69.67 (6)	<.001	91
Yang et al [34]	ADL	5.18 (6)	.52	0
Zheng et al [40]	ADL	73.31 (6)	<.001	92

[a]K-MBI: Korean version of the Modified Barthel Index.

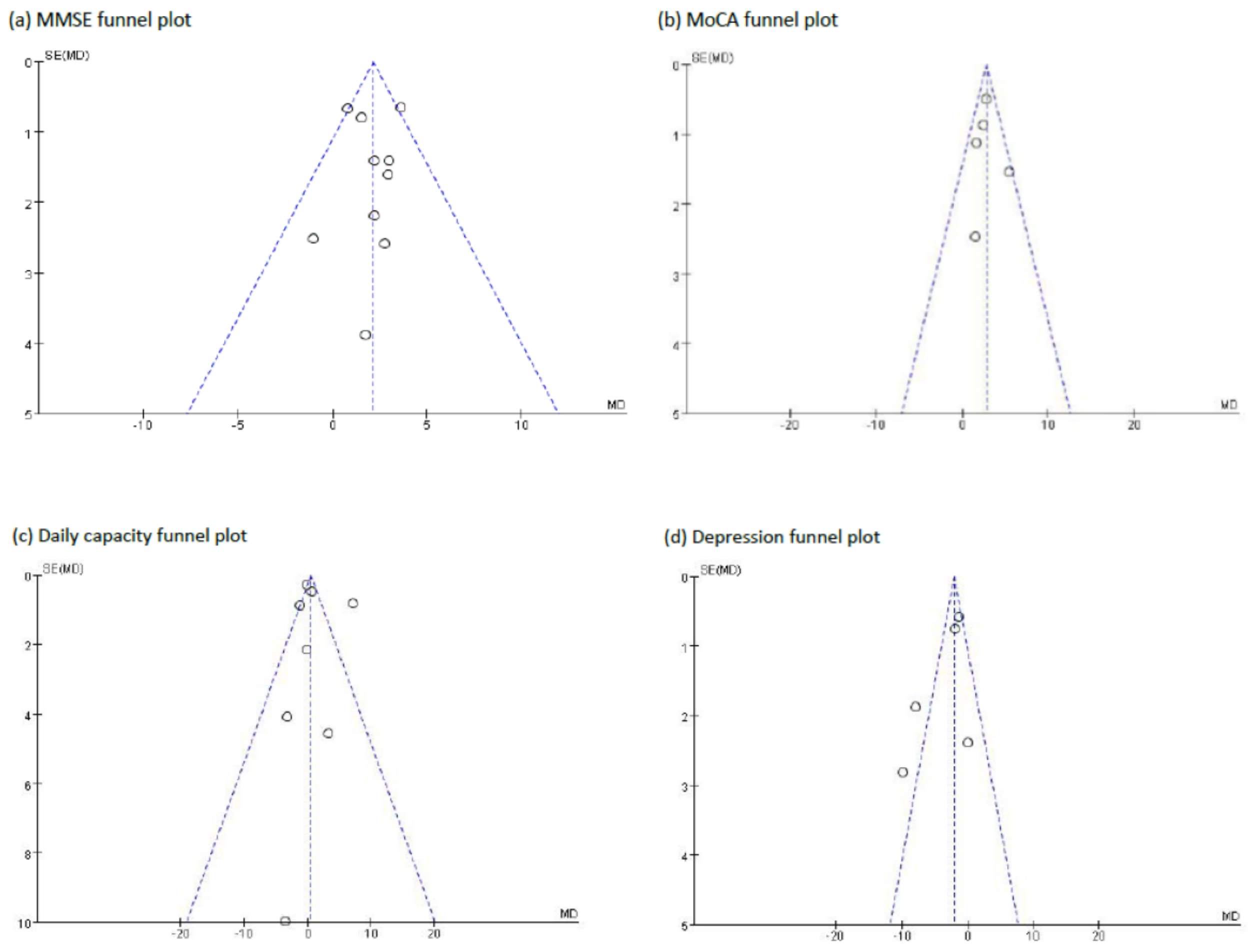
[b]IADL: Instrumental Activities of Daily Living.

[c]ADL: activities of daily living.

Publication Bias

We found that the funnel plots for the MMSE (Figure 12a), MoCA (Figure 12b), daily behavioral capacity indicators (Figure 12c), and depression indicators (Figure 12d) were all symmetrical, indicating the absence of publication bias. P>.05 indicated the absence of obvious publication bias. The Begg test (P=.03) and Egger regression (P=.026) also indicated a lack of publication bias.

Figure 12. Funnel plots for (a) MMSE, (b) MoCA, (c) daily behavioral capacity indicators, and (d) depression indicators. MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment.



Discussion

Principal Findings

This study systematically evaluated the efficacy of training older adults with cognitive impairment using digital SGs, drawing upon data from 14 studies with a total of 714 participants. Our findings revealed that digital SGs significantly alleviate depression and yield large improvements in cognitive ability and daily behavioral capacity. This suggests that SGs offer a highly effective alternative for older adults with cognitive impairment. Due to statistical heterogeneity, these results should be interpreted with caution. In this study, sensitivity analysis was used to explore the major sources of statistical heterogeneity to determine the impact of a single study on overall risk. The type of intervention (various digital SGs), sample size (range 17-112), intervention duration (range 4-12 weeks), weekly intervention time (range 45-300 minutes/week), type of CG (traditional Taijiquan group, routine health care group, music group), cultural background, measurement tools, or other confounding factors may have caused heterogeneity. To the best of our knowledge, this is the first meta-analysis and systematic review evaluating the effectiveness of the 3 outcome indicators of serious gameplay in older adults with cognitive impairment. In this study, SGs were found to significantly increase cognitive ability and daily behavioral capacity and decrease depression.

Previous studies have explored the effects of digital SGs on cognitive ability [57], daily behavioral capacity [95], and depressive emotion problems [96] in patients with MCI or AD. However, previous studies have reported inconsistent findings regarding the effects of digital SG therapy on cognitive function [21,60-62], daily behavioral capacity [52,63], and depression [21,50,64-66] in older adults with MCI or AD. This systematic review included the recent literature (2017-2023) and examined 3 indicators of the effects of SGs on cognitive ability, daily behavioral capacity, and depression among older adults with MCI or AD. Consistent with previous findings, our results indicated positive effects of SGs on MMSE [97] and CSDD [50,98] scores and negative effects on depression on EDG-15, EADG, and GDS scores. In addition, there was a significant increase in cognitive ability (measured with MoCA) and daily behavioral capacity (measured with the ADL scale, the K-MBI, the Instrumental Activities of Daily Living (IADL) scale, the Katz Index, the Barthel Index, and the Lawton and Brody Index). This systematic review is the first to simultaneously test the effects of SGs on the 3 indicators of cognitive ability, daily behavioral capacity, and depression. This systematic review also divided the subgroup analysis of weeks and continental plates, which was not available in previous systematic reviews.

SGs are defined as games that are not for entertainment purposes and are mainly used for prevention, education, diagnosis, screening, and rehabilitation [99,100]. The word first appeared in 1968 when Abt [101] named his new book on SGs. According to Clark's [101] proposed treatment, games can exist in a variety of forms, including (1) exergames or video games that require physical exercise as part of playing the games [38,39]; (2) cognitive training games, which aim to maintain or improve users' cognitive abilities (eg, executive function, memory, learning) [34-37]; (3) computerized CBT games, which are video games that provide CBT to the users; and (4) biofeedback games, which are video games that use electrical sensors attached to the player to receive information about the player's body state (eg, electrocardiogram sensors) and seek to influence some of the player's body functions (eg, heart rate). Special care units (SCUs), also known as dementia care units, began to appear in the United States in the 1980s for older adults with dementia [102]. With an increase in the number of older adults with dementia and the diversification of the demand for dementia care in the market, these SCUs provide not only services to older adults in institutions but also family counseling, care support, and community therapy services to older adults with MCI [103]. At the policy level, the United States also launched the National Plan to Address Alzheimer's Disease in 2012, with the goal of preventing and effectively treating AD by 2025 [104]. The latest SG therapy has also had a positive impact on easing the stress of dementia among older adults in US society [66]. However, the attention to older patients with MCI and AD in the West is much earlier than in China. Model-based projection data put China's older adults with dementia at 6.73-15.7 million in 2025; the number of older adults with dementia will increase to 10.018-23.3754 million in 2035 [105]. China's Circular on the Promotion of Dementia Prevention and Treatment Action (2023-2025) [106] formally focuses on groups such as older adults with MCI and AD at the national policy level. Ren et al [11] found that the 2015 annual treatment cost of patients with AD in China was US $167.74 billion, with ever-rising treatment costs expected to reach US $1.8 trillion by 2050. Older care facilities, social work stations, and counseling services in China may use SGs to carry out nonpharmacological treatments for patients with MCI and AD. This will provide both physical and psychological health improvements to older adults and at the same time reduce the financial burden on China. In traditional Chinese society, positive family and social relationships contribute to physical and mental health [107]. Due to the "Xiao Dao" culture, this also constitutes a unique social security system for older adults [108,109]. However, rapid industrialization and urbanization have changed these traditional arrangements [110]. Today, older adults are less likely to live with and interact with younger generations, which has resulted in older adults being isolated [111]. This has a direct impact on access to social care and financial security and even affects the physical and mental health of older adults [112]. SGs can provide opportunities for older adults to interact with others [113] and reduce the negative emotional impact on them [114]. In addition, the sports elements included in SGs [115] can exercise older adults' cognitive ability [67] and daily behavioral capacity [116] and may even promote intergenerational relationships [117]. Our findings also show that SGs can have a positive impact on cognitive ability and daily behavioral capacity and decrease depression in older adults. With this in mind, we suggest that nursing home caregivers, social workers, and psychologists consider using SGs in their work and collaborate with SG professionals to bring benefits to older adults with MCI or AD. This will slow down the cognitive loss in older adults with MCI or AD to a certain extent.

Strengths and Limitations

Compared with previous studies [90,118,119], this review is the only one to specifically evaluate the effects of SGs on older adults and related types of games. Because the study was conducted in strict accordance with PRISMA guidelines, it was a high-quality, transparent review. Only RCTs were included; this is the most rigorous method of studying causal relationships [120,121]. Therefore, the conclusions of this review are more credible. In this review, there was minimal publication bias. We searched 12 worldwide databases and included research from the United Kingdom, China, Belgium, the Netherlands, Portugal, Spain, and Korea. We also performed forward and backward checks of reference lists, and an all-around search strategy was applied. We searched a database of gray literature without restricting the search to a specific comparator, population, setting, or country, which made the samples more representative. Since all the processes (ie, data extraction, research selection, evidence quality assessment, and bias risk assessment) were independently conducted by 4 reviewers, there was no risk of bias in this review. In addition, because we used the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) method to evaluate the quality of evidence, this review will enable readers to draw more accurate conclusions. Where possible, our statistics integrate data, which will increase not only the power of the research but also estimates of the effects of digital SG therapy on cognitive ability, daily behavioral capacity, and depression. Moreover, due to the large sample size, we performed a subgroup analysis and a sensitivity analysis. In addition, this review used pre- and postintervention outcome data from each group to evaluate the effectiveness of each meta-analysis. All previous studies have reported the SD and average overall cognitive changes pre- and postintervention in each group. In addition, in these studies, there were significant differences in overall cognition between groups at baseline.

The limitations of this study should also be considered. First, due to the design of the systematic review, a causal relationship with a clear structure could not be inferred. Second, the outcome indicators of anxiety (EADG scores) and quality of life (Dementia Quality of Life (DQoL)) were also related to the mental health of older adults with cognitive impairment. However, with only 1 study reporting these findings, a comparison could not be made, so we did not include these indicators. Third, because of the limitations of the designs of the included studies, the randomization and blinding methods used were seldom described in detail. Only 11 studies described the randomization method. In contrast, randomization was only mentioned in other studies, but no explanation of the method used was provided. Blinding was only implemented in 7 studies. Fourth, since many potentially eligible studies were pilot RCTs or quasiexperiments published before 2017 or in languages other than English, we excluded them. As a result, some related research may have been missed. The reason for excluding these studies is that the internal validity of pilot RCTs and quasiexperiments is lower than that of RCTs [122,123]. In addition, we were also unable to translate all the non-English research literature. In our previous review, we found some studies published before 2017. However, serious progress has been made in the field of gaming in the past 7 years, so we did not include RCTs published before 2017. Fifth, this review focused on the short-term effects of SGs. This meta-analysis of pre- and postintervention data did not include follow-up data. We found that follow-up data were reported in only 4 studies, and there was no consistency between the follow-up periods. Therefore, it was impossible to evaluate the long-term effects of SGs on the cognitive ability, mental health, and daily behavioral capacity of older adults. The quality of the evidence in our meta-analysis was average, which may have affected the internal validity of the findings. Sixth, in this review, interest intervention was limited to SGs used as therapeutic interventions on digital platforms, so it would not have been possible to comment on and check the effectiveness of SGs and nondigital SGs used for other purposes, such as monitoring, screening, or diagnosis.

Research and Practical Implications

Since the review concentrated on the effectiveness of SGs in improving cognitive ability, daily behavioral capacity, and depression among older adults with cognitive impairment, researchers should evaluate the effectiveness of SGs by performing further reviews of their positive effects on cognitive abilities (eg, processing speed, executive function, learning, memory, speed), daily behavioral capacity (eg, executive function, capacity for reaction, upper limb activity), and mental health (eg, anxiety, stress, sleep quality, quality of life) among people in different age groups. We included some studies from middle-income countries, so the generalizability of the findings of this review may be limited to these countries. As the cultural, social, and economic conditions of middle-income countries are different, additional relevant research should be carried out. To date, only a few studies have assessed the effectiveness of SGs in improving cognitive ability. Therefore, additional research is needed to help draw clearer conclusions about the effectiveness of SGs. Due to the lack of follow-up data, the long-term effects of SGs were not evaluated in this review. To assess the long-term cognitive effects of SGs in a worldwide context, researchers should follow up with participants. A few studies did not report the SD or average global cognitive changes in each group before and after the intervention. In the future, relevant studies need to report this information to evaluate the effect more accurately. Because most of the existing studies have problems reporting results and choosing the randomization process, 8 studies had a low risk of overall bias. Therefore, when developing and reporting on RCTs, researchers should use recommended tools (eg, RoB-2 [124]) and follow recommended guidelines to avoid the aforementioned biases. Many studies have examined the worldwide impact of digital SGs on the cognition of healthy older adults [57], but this review revealed that only 14 studies have investigated the effects of digital SGs on older adults with cognitive impairment. We hope that researchers can perform additional research in this area to fill this gap.

According to the results of this review, SGs, especially digital SGs, exhibit greater effectiveness than a lack of interventions. These findings have meaningful implications for clinical practice since the use of SGs can improve cognitive ability and daily behavioral capacity and decrease depression in older adults with cognitive impairment. However, these findings should be interpreted with caution because the quality of the evidence is low in most meta-analyses. Moreover, most of the studies included in this review had some problems in terms of overall bias. Except for 7 studies [34,39,41-43,45,47], in all the other included studies, heterogeneity was high and the total effect was not accurate enough. Therefore, before more strong evidence is obtained, psychiatrists and psychologists should consider SGs as a complement to existing interventions rather than a direct substitute. In addition, the number of older persons worldwide will grow exponentially in the next few years, so we can try to use SGs to lighten the burden on the health care system. We can use SGs to improve the physical, psychological, social, and motor functions of older persons with cognitive impairment, thereby improving their quality of life [54,125]. Currently, many old people live alone and have limited social interactions, which can often lead to increased morbidity and mortality [126]. Older adults can play SGs in a relatively cozy environment (ie, at home), thus promoting their connections with friends and family [51]. In all the studies, the platforms for SGs were mobile devices (tablets and smartphones). Mobile devices are more popular than game consoles, cheaper than computers, and easier to access; thus, they are particularly popular. As of 2021, there were more than 7.1 billion mobile device users and 15 billion mobile devices worldwide [127]. Game and app developers can work together to develop SGs aimed at improving the mental health, daily behavior capacity, and cognitive ability of older people. These SGs can be operated and played using mobile devices. A survey of a small number of studies conducted in middle-income countries revealed that these countries seem to prefer SGs more than high-income countries do [128]. Middle-income countries lack many psychological professionals (currently 1 for every 10 million people). Therefore, more SGs should be developed in middle-income countries to improve the cognitive ability, mental health, and daily behavioral capacity of older persons in those countries.

Conclusion

Overall, SGs, particularly digital SGs, may be associated with increased cognitive function and daily behavioral capacity, as well as decreased depression, among older adults with cognitive impairment. However, we are still unable to draw definitive conclusions on the safety and effectiveness of SGs in improving the cognition, behavioral capacity, and mental health of older people with cognitive impairment. This is due to the poor quality of existing meta-analysis evidence, which is mainly caused by concerns about the bias of most studies, the accuracy of the total effect, and the high degree of evidence heterogeneity. Therefore, until more reliable evidence is obtained, psychiatrists, psychologists, patients, and social workers should all regard SGs as a complement to existing interventions rather than a direct substitute. Before SGs are used, their effectiveness and safety should be further reviewed and evaluated, mainly in terms of improving daily behavioral capacity (eg, reaction ability, executive function, upper limb activity) of people with or without cognitive impairment in different age groups, cognitive ability (eg, processing speed, executive function, learning, memory), and mental health aspects (eg, stress, anxiety, sleep quality, quality of life). Further research is needed to assess the safety of SGs, the effects of sports games, and their long-term effects. In the future, further research on methods for increasing the motivation of participants, reducing the rate of missing data, and maintaining the effects of SGs is needed. Second, due to study limitations, it was impossible to perform subgroup analyses for intervention types. Consequently, in-depth and stratified discussions and comparisons of different intervention types were performed. In the coming years, it will be necessary for researchers to perform higher-quality studies with larger samples to verify the effectiveness of SG for older adults with cognitive impairment.

Acknowledgments

We confirm the submission of this manuscript, the key message of the manuscript, and the uniqueness of the study. We express our gratitude to the reviewers for their valuable remarks. We also state that the paper authorship, research, and publication have received financial support from the National Social Science Foundation in China (19BSH161) and the Shenzhen University-Lingnan University Joint Research Programme (2023004).

Data Availability

In this study, the original contributions have been included in the supplementary materials, and further inquiries can be made directly from the corresponding author. No artificial intelligence was used in any portion of the manuscript writing.

All the claims made by this paper represent only those of the authors and do not necessarily represent the claims of the affiliated organizations or reviewers, editors, and publishers. The publisher does not guarantee or endorse any product that may be evaluated by this paper or any claim that may be made by its manufacturer.

Authors' Contributions

XZ and YT managed conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing—original draft, and writing—review and editing. YC was responsible for conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualization, writing—original draft, and writing—review. ZZ was responsible for data curation, formal analysis, methodology, resources, and writing—original draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 abstract checklist.

[PDF File (Adobe PDF File), 131 KB - games_v12i1e55785_app1.pdf]

Multimedia Appendix 2

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 checklist.

[PDF File (Adobe PDF File), 145 KB - games_v12i1e55785_app2.pdf]

Multimedia Appendix 3

Search strategy.

[DOCX File , 14 KB - games_v12i1e55785_app3.docx]

Multimedia Appendix 4

Intervention characteristics.

[DOCX File , 25 KB - games_v12i1e55785_app4.docx]

References

1. Gaidai O, Yakimov V, Balakrishna R. Dementia death rates prediction. BMC Psychiatry 2023 Sep 22;23(1):691 [FREE Full text] [doi: 10.1186/s12888-023-05172-2] [Medline: 37736716]
2. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, International Psychogeriatric Association Expert Conference on mild cognitive impairment. Mild cognitive impairment. Lancet 2006 Apr 15;367(9518):1262-1270. [doi: 10.1016/S0140-6736(06)68542-5] [Medline: 16631882]
3. Garre-Olmo J. Epidemiology of Alzheimer's disease and other dementias. Rev Neurol 2018 Jun 1;66(11):377-386. [doi: 10.33588/rn.6611.2017519]
4. Sachdev PS, Blacker D, Blazer DG, Ganguli M, Jeste DV, Paulsen JS, et al. Classifying neurocognitive disorders: the DSM-5 approach. Nat Rev Neurol 2014 Nov 30;10(11):634-642. [doi: 10.1038/nrneurol.2014.181] [Medline: 25266297]
5. Nandi A, Counts N, Chen S, Seligman B, Tortorice D, Vigo D, et al. Global and regional projections of the economic burden of Alzheimer's disease and related dementias from 2019 to 2050: a value of statistical life approach. EClinicalMedicine 2022 Sep;51:101580 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101580] [Medline: 35898316]
6. Festari C, Massa F, Cotta Ramusino M, Gandolfo F, Nicolosi V, Orini S, et al. European consensus for the diagnosis of MCI and mild dementia: preparatory phase. Alzheimers Dement 2023 May 16;19(5):1729-1741 [FREE Full text] [doi: 10.1002/alz.12798] [Medline: 36209379]
7. Anderson ND. State of the science on mild cognitive impairment (MCI). CNS Spectr 2019 Jan 17;24(1):78-87. [doi: 10.1017/s1092852918001347]
8. Jia L, Quan M, Fu Y, Zhao T, Li Y, Wei C, et al. Dementia in China: epidemiology, clinical management, and research advances. Lancet Neurol 2020 Jan;19(1):81-92. [doi: 10.1016/s1474-4422(19)30290-x]
9. Wimo A, Seeher K, Cataldi R, Cyhlarova E, Dielemann JL, Frisell O, et al. The worldwide costs of dementia in 2019. Alzheimers Dement 2023 Jul 08;19(7):2865-2873 [FREE Full text] [doi: 10.1002/alz.12901] [Medline: 36617519]
10. Jia L, Du Y, Chu L, Zhang Z, Li F, Lyu D, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671. [doi: 10.1016/s2468-2667(20)30185-7]
11. Ren R, Qi J, Lin S, Liu X, Yin P, Wang Z, et al. The China Alzheimer report 2022. Gen Psychiatr 2022 Mar 11;35(1):e100751 [FREE Full text] [doi: 10.1136/gpsych-2022-100751] [Medline: 35372787]
12. Lobbia A, Carbone E, Faggian S, Gardini S, Piras F, Spector A, et al. The efficacy of cognitive stimulation therapy (CST) for people with mild-to-moderate dementia. Eur Psychol 2019 Jul;24(3):257-277. [doi: 10.1027/1016-9040/a000342]
13. Rubin R. Exploring the relationship between depression and dementia. JAMA 2018 Sep 11;320(10):961-962. [doi: 10.1001/jama.2018.11154] [Medline: 30140930]
14. Arvanitakis Z, Bennett DA. What is dementia? JAMA 2019 Nov 05;322(17):1728 [FREE Full text] [doi: 10.1001/jama.2019.11653] [Medline: 31688886]
15. Cipriani G, Danti S, Picchi L, Nuti A, Fiorino MD. Daily functioning and dementia. Dement Neuropsychol 2020 Jun;14(2):93-102 [FREE Full text] [doi: 10.1590/1980-57642020dn14-020001] [Medline: 32595877]
16. Cloutier S, Chertkow H, Kergoat M, Gélinas I, Gauthier S, Belleville S. Trajectories of decline on instrumental activities of daily living prior to dementia in persons with mild cognitive impairment. Int J Geriatr Psychiatry 2021 Feb 17;36(2):314-323 [FREE Full text] [doi: 10.1002/gps.5426] [Medline: 32892375]

17. Xu L, Hsiao H, Denq W, Chi I. Training needs for dementia care in China from the perspectives of mental health providers: who, what, and how. Int Psychogeriatr 2017 Dec 10;30(7):929-940. [doi: 10.1017/s1041610217002654]
18. Linnemann C, Lang UE. Pathways connecting late-life depression and dementia. Front Pharmacol 2020 Mar 13;11:279 [FREE Full text] [doi: 10.3389/fphar.2020.00279] [Medline: 32231570]
19. Paykel ES. Basic concepts of depression. Dialogues Clin Neurosci 2008;10(3):279-289 [FREE Full text] [doi: 10.31887/DCNS.2008.10.3/espaykel] [Medline: 18979941]
20. Asad J, Kousar S, Mehmood N. Dementia-related serious games: a comparative study. Univ Sindh J Inf Commun Technol 2019;3(4):171-177.
21. Saragih ID, Everard G, Lee B. A systematic review and meta-analysis of randomized controlled trials on the effect of serious games on people with dementia. Ageing Res Rev 2022 Dec;82:101740. [doi: 10.1016/j.arr.2022.101740] [Medline: 36191740]
22. Leung DKY, Chan WC, Spector A, Wong GHY. Prevalence of depression, anxiety, and apathy symptoms across dementia stages: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2021 Sep;36(9):1330-1344. [doi: 10.1002/gps.5556] [Medline: 33905138]
23. Bränsvik V, Granvik E, Minthon L, Nordström P, Nägga K. Mortality in patients with behavioural and psychological symptoms of dementia: a registry-based study. Aging Ment Health 2021 Jun 18;25(6):1101-1109 [FREE Full text] [doi: 10.1080/13607863.2020.1727848] [Medline: 32067466]
24. Lau SJ, Agius H. A framework and immersive serious game for mild cognitive impairment. Multimed Tools Appl 2021 May 31;80(20):31183-31237. [doi: 10.1007/s11042-021-11042-4]
25. Contreras-Somoza LM, Irazoki E, Toribio-Guzmán JM, de la Torre-Díez I, Diaz-Baquero AA, Parra-Vidales E, et al. Usability and user experience of cognitive intervention technologies for elderly people with MCI or dementia: a systematic review. Front Psychol 2021 Apr 22;12:636116 [FREE Full text] [doi: 10.3389/fpsyg.2021.636116] [Medline: 33967901]
26. Amjad I, Toor H, Niazi IK, Pervaiz S, Jochumsen M, Shafique M, et al. Xbox 360 Kinect cognitive games improve slowness, complexity of EEG, and cognitive functions in subjects with mild cognitive impairment: a randomized control trial. Games Health J 2019 Apr;8(2):144-152. [doi: 10.1089/g4h.2018.0029] [Medline: 30239213]
27. Ng Y, Ma F, Ho FK, Ip P, Fu K. Effectiveness of virtual and augmented reality-enhanced exercise on physical activity, psychological outcomes, and physical performance: a systematic review and meta-analysis of randomized controlled trials. Comput Hum Behav 2019 Oct;99:278-291. [doi: 10.1016/j.chb.2019.05.026]
28. Christogianni A, Bojan K, Mukaetova-Ladinska EB, Sriramm VT, Murth G, Kumarpilla G. Attention and motor skill improvements in mild cognitive impairment patients using COSMA application. Int J Serious Games 2022 Nov 08;9(4):25-42. [doi: 10.17083/ijsg.v9i4.517]
29. Montana JI, Matamala-Gomez M, Maisto M, Mavrodiev PA, Cavalera CM, Diana B, et al. The benefits of emotion regulation interventions in virtual reality for the improvement of wellbeing in adults and older adults: a systematic review. J Clin Med 2020 Feb 12;9(2):500 [FREE Full text] [doi: 10.3390/jcm9020500] [Medline: 32059514]
30. Xue B, Xiao A, Luo X, Li R. The effect of a game training intervention on cognitive functioning and depression symptoms in the elderly with mild cognitive impairment: a randomized controlled trial. Int J Methods Psychiatr Res 2021 Dec 22;30(4):e1887 [FREE Full text] [doi: 10.1002/mpr.1887] [Medline: 34291527]
31. George W. Children's Play: Grow in the Game. Beijing, China: China Light Industry Press; 2008.
32. Dewhirst A, Laugharne R, Shankar R. Therapeutic use of serious games in mental health: scoping review. BJPsych Open 2022 Feb 02;8(2):e37 [FREE Full text] [doi: 10.1192/bjo.2022.4] [Medline: 35105418]
33. Sriwatanathamma P, Sirivesmas V, Simatrang S, Bhowmik NH. Gamifying cognitive behavioral therapy techniques on smartphones for Bangkok's millennials with depressive symptoms: interdisciplinary game development. JMIR Serious Games 2023 May 12;11:e41638 [FREE Full text] [doi: 10.2196/41638] [Medline: 37171845]
34. Yang Y, Kwak YT. Improvement of cognitive function after computer-based cognitive training in early stage of Alzheimer's dementia. Dement Neurocogn Disord 2017 Mar;16(1):7-11 [FREE Full text] [doi: 10.12779/dnd.2017.16.1.7] [Medline: 30906364]
35. Kwan RY, Lee D, Lee PH, Tse M, Cheung DS, Thiamwong L, et al. Effects of an mHealth brisk walking intervention on increasing physical activity in older people with cognitive frailty: pilot randomized controlled trial. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e16596 [FREE Full text] [doi: 10.2196/16596] [Medline: 32735218]
36. Lim EH, Kim D, Won Y, Park S, Seo J, Ko M, et al. Effects of home based serious game training (Brain Talk™) in the elderly with mild cognitive impairment: randomized, a single-blind, controlled trial. Brain Neurorehabil 2023 Mar;16(1):e4 [FREE Full text] [doi: 10.12786/bn.2023.16.e4] [Medline: 37033004]
37. Lee GJ, Bang HJ, Lee KM, Kong HH, Seo HS, Oh M, et al. A comparison of the effects between 2 computerized cognitive training programs, Bettercog and COMCOG, on elderly patients with MCI and mild dementia: a single-blind randomized controlled study. Medicine (Baltimore) 2018 Nov;97(45):e13007 [FREE Full text] [doi: 10.1097/MD.0000000000013007] [Medline: 30407291]
38. Yang F. The effect of somatosensory interactive games on patients with Alzheimer's disease. Chin Foreign Med Res 2021;19(14):126-129 [FREE Full text] [doi: 10.14033/j.cnki.cfmr.2021.14.042]

39. van Santen J, Dröes RM, Twisk JWR, Blanson Henkemans OA, van Straten A, Meiland FJM. Effects of exergaming on cognitive and social functioning of people with dementia: a randomized controlled trial. J Am Med Dir Assoc 2020 Dec;21(12):1958-1967.e5 [FREE Full text] [doi: 10.1016/j.jamda.2020.04.018] [Medline: 32651132]
40. Zheng J, Chen X. The effect of Kinect-based somatosensory game on patients with dementia. J Nurs Sci 2018 May;33(9):5. [doi: 10.3870/j.issn.1001-4152.2018.09.005]
41. Wu Y, Liu H, Sun D. The application value of sensory stimulation combined with somatosensory interactive games in the care of elderly patients with dementia. Int Med Health Guide 2021;27(14):2189-2192. [doi: 10.3760/cma.j.issn.1007-1245.2021.14.038]
42. Savulich G, Piercy T, Fox C, Suckling J, Rowe JB, O'Brien JT, et al. Cognitive training using a novel memory game on an iPad in patients with amnestic mild cognitive impairment (aMCI). Int J Neuropsychopharmacol 2017 Aug 01;20(8):624-633 [FREE Full text] [doi: 10.1093/ijnp/pyx040] [Medline: 28898959]
43. Swinnen N, Vandenbulcke M, de Bruin ED, Akkerman R, Stubbs B, Firth J, et al. The efficacy of exergaming in people with major neurocognitive disorder residing in long-term care facilities: a pilot randomized controlled trial. Alzheimers Res Ther 2021 Mar 30;13(1):70 [FREE Full text] [doi: 10.1186/s13195-021-00806-7] [Medline: 33785077]
44. Liu C, Cheng F, Wei M, Liao Y. Effects of exergaming-based Tai Chi on cognitive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2022 Mar 15;14:761053 [FREE Full text] [doi: 10.3389/fnagi.2022.761053] [Medline: 35370622]
45. Thapa N, Park HJ, Yang J, Son H, Jang M, Lee J, et al. The effect of a virtual reality-based intervention program on cognition in older adults with mild cognitive impairment: a randomized control trial. J Clin Med 2020 Apr 29;9(5):1283 [FREE Full text] [doi: 10.3390/jcm9051283] [Medline: 32365533]
46. Oliveira J, Gamito P, Souto T, Conde R, Ferreira M, Corotnean T, et al. Virtual reality-based cognitive stimulation on people with mild to moderate dementia due to Alzheimer's disease: a pilot randomized controlled trial. Int J Environ Res Public Health 2021 May 16;18(10):5290 [FREE Full text] [doi: 10.3390/ijerph18105290] [Medline: 34065698]
47. Jahouh M, González-Bernal JJ, González-Santos J, Fernández-Lázaro D, Soto-Cámara R, Mielgo-Ayuso J. Impact of an intervention with Wii video games on the autonomy of activities of daily living and psychological-cognitive components in the institutionalized elderly. Int J Environ Res Public Health 2021 Feb 07;18(4):1570 [FREE Full text] [doi: 10.3390/ijerph18041570] [Medline: 33562249]
48. Abd-Alrazaq A, Abuelezz I, AlSaad R, Al-Jafar E, Ahmed A, Aziz S, et al. Serious games for learning among older adults with cognitive impairment: systematic review and meta-analysis. J Med Internet Res 2023 Apr 12;25:e43607 [FREE Full text] [doi: 10.2196/43607] [Medline: 37043277]
49. Dewhirst A, Laugharne R, Shankar R. Therapeutic use of serious games in mental health: scoping review. BJPsych Open 2022 Feb 02;8(2):e37 [FREE Full text] [doi: 10.1192/bjo.2022.4] [Medline: 35105418]
50. Bai J, Tang J, Xie Q. Does children’s marriage matter for parents' mental health？ Evidence from China. J Asian Econom 2022 Feb;78:101426. [doi: 10.1016/j.asieco.2021.101426]
51. Papoutsi C, Drigas A, Skianis C. Serious games for emotional intelligence’s skills development for inner balance and quality of life: a literature review. Retos 2022 Sep 28;46:199-208. [doi: 10.47197/retos.v46.91866]
52. Abd-Alrazaq A, Alhuwail D, Ahmed A, Househ M. Effectiveness of serious games for improving executive functions among older adults with cognitive impairment: systematic review and meta-analysis. JMIR Serious Games 2022 Jul 25;10(3):e36123 [FREE Full text] [doi: 10.2196/36123] [Medline: 35877166]
53. Jones CM, Scholes L, Johnson D, Katsikitis M, Carras MC. Gaming well: links between videogames and flourishing mental health. Front Psychol 2014 Mar;5:260 [FREE Full text] [doi: 10.3389/fpsyg.2014.00260] [Medline: 24744743]
54. Muñoz J, Montoya M, Boger J. From exergames to immersive virtual reality systems: Serious games for supporting older adults. In: Smart Home Technologies and Services for Geriatric Rehabilitation. New York: Academic Press; 2021.
55. Wang Y, Wang Z, Liu G, Wang Z, Wang Q, Yan Y, et al. Application of serious games in health care: scoping review and bibliometric analysis. Front Public Health 2022 Jun 10;10:896974 [FREE Full text] [doi: 10.3389/fpubh.2022.896974] [Medline: 35757619]
56. Tori AA, Tori R, Nunes FDLDS. Serious game design in health education: a systematic review. IEEE Trans Learning Technol 2022 Dec 1;15(6):827-846. [doi: 10.1109/tlt.2022.3200583]
57. Abd-Alrazaq A, Alajlani M, Alhuwail D, Toro CT, Giannicchi A, Ahmed A, et al. The effectiveness and safety of serious games for improving cognitive abilities among elderly people with cognitive impairment: systematic review and meta-analysis. JMIR Serious Games 2022 Mar 10;10(1):e34592 [FREE Full text] [doi: 10.2196/34592] [Medline: 35266877]
58. Gonzaga A, Lamas M. Money management: serious games for the elderly. In: Handbook of Research on Promoting Economic and Social Development Through Serious Games. Pennsylvania: IGI Global; 2022.
59. De Vidts K. Generic futures and serious games: lessons from 2040 for social innovation in Flanders, Belgium. Journal of Futures Studies. 2023. URL: https://jfsdigital.org/articles-and-essays/2023-2/generic-futures-and-serious-games-lessons-from-2040-for-social-innovation-in-flanders-belgium/ [accessed 2024-07-11]
60. Ning H, Li R, Ye X, Zhang Y, Liu L. A review on serious games for dementia care in ageing societies. IEEE J Transl Eng Health Med 2020;8:1400411 [FREE Full text] [doi: 10.1109/JTEHM.2020.2998055] [Medline: 32537264]

61. Huang L, Yang Y. The long-term effects of immersive virtual reality reminiscence in people with dementia: longitudinal observational study. JMIR Serious Games 2022 Jul 25;10(3):e36720 [FREE Full text] [doi: 10.2196/36720] [Medline: 35877169]
62. Kleschnitzki JM, Beyer L, Beyer R, Großmann I. The effectiveness of a serious game (MemoreBox) for cognitive functioning among seniors in care facilities: field study. JMIR Serious Games 2022 Apr 01;10(2):e33169 [FREE Full text] [doi: 10.2196/33169] [Medline: 35172959]
63. Wang P, Fang Y, Qi J, Li H. FISHERMAN: a serious game for executive function assessment of older adults. Assessment 2023 Jul 28;30(5):1499-1513. [doi: 10.1177/10731911221105648] [Medline: 35762827]
64. Bojan K, Stavropoulos TG, Lazarou I, Nikolopoulos S, Kompatsiaris I, Tsolaki M, et al. The effects of playing the COSMA cognitive games in dementia. Int J Serious Games 2021 Mar 09;8(1):45-58. [doi: 10.17083/ijsg.v8i1.412]
65. James S, Conrad M, Sengupta S. Development of an innovative telerehabilitation system to improve handgrip strength. Int J Telerehabil 2022 Dec 13;14(2):e6497 [FREE Full text] [doi: 10.5195/ijt.2022.6497] [Medline: 38026566]
66. Fitzgerald M, Ratcliffe G. Serious serious games, gamification, and serious mental illness: a scoping review. Psychiatr Serv 2020 Feb 01;71(2):170-183. [doi: 10.1176/appi.ps.201800567] [Medline: 31640521]
67. Manca M, Paternò F, Santoro C, Zedda E, Braschi C, Franco R, et al. The impact of serious games with humanoid robots on mild cognitive impairment older adults. Int J Hum-Comput Stud 2021 Jan;145:102509 [FREE Full text] [doi: 10.1016/j.ijhcs.2020.102509]
68. Pacheco TBF, de Medeiros CSP, de Oliveira VHB, Vieira ER, de Cavalcanti FAC. Effectiveness of exergames for improving mobility and balance in older adults: a systematic review and meta-analysis. Syst Rev 2020 Jul 18;9(1):163 [FREE Full text] [doi: 10.1186/s13643-020-01421-7] [Medline: 32682439]
69. Yu RWL, Chan AHS. Meta-analysis of the effects of game types and devices on older adults-video game interaction: implications for video game training on cognition. Appl Ergon 2021 Oct;96:103477. [doi: 10.1016/j.apergo.2021.103477] [Medline: 34107433]
70. Tugwell P, Tovey D. PRISMA 2020. J Clin Epidemiol 2021 Jun;134:A5-A6. [doi: 10.1016/j.jclinepi.2021.04.008] [Medline: 34016443]
71. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
72. Amir-Behghadami M, Janati A. Population, Intervention, Comparison, Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. Emerg Med J 2020 Jun 05;37(6):387-387. [doi: 10.1136/emermed-2020-209567] [Medline: 32253195]
73. Grey literature database. OPENGREY. URL: https://opengrey.eu/ [accessed 2024-07-10]
74. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348(mar07 3):g1687-g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]
75. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Vol. 2. Washington, DC: American Psychological Association; 2012.
76. Higgins JPT, Savović J, Page M, Elbers R, Sterne J. Assessing risk of bias in a randomized trial. In: Cochrane Handbook for Systematic Reviews of Interventions. London, UK: Cochrane; 2019:205-228.
77. Mohsina S, Gurushankari B, Niranjan R, Sureshkumar S, Sreenath G, Kate V. Assessment of the quality of randomized controlled trials in surgery using Jadad score: where do we stand? J Postgrad Med. J Postgrad Med 2022;68(4):207-212. [doi: 10.4103/jpgm.jpgm_104_21]
78. Harbord RM, Harris RJ, Sterne JAC. Updated tests for small-study effects in meta-analyses. Stata J 2009 Aug 01;9(2):197-210. [doi: 10.1177/1536867x0900900202]
79. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]
80. Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. J R Stat Soc A (Stat Soc) 1988;151(3):419. [doi: 10.2307/2982993]
81. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
82. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000 Jun 24;56(2):455-463. [doi: 10.1111/j.0006-341x.2000.00455.x] [Medline: 10877304]
83. Sterne J, Becker B, Egger M. The funnel plot. In: Rothstein HR, Sutton AJ, Borenstein M, editors. Publication Bias in Meta‐Analysis: Prevention, Assessment and Adjustments. Hoboken, NJ: John Wiley & Sons; 2005:73-98.
84. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019 Oct 03;10(10):ED000142 [FREE Full text] [doi: 10.1002/14651858.ED000142] [Medline: 31643080]

85. Rathvon D. EndNote X8--citation manager--what's new? UT Southwestern Medical Center. 2017. URL: https://utswmed-ir.tdl.org/items/04c4fae7-b82c-4a72-935b-cc7637d66eb2 [accessed 2024-07-11]
86. Padala KP, Padala PR, Lensing SY, Dennis RA, Bopp MM, Roberson PK, et al. Home-based exercise program improves balance and fear of falling in community-dwelling older adults with mild Alzheimer's disease: a pilot study. J Alzheimers Dis 2017;59(2):565-574. [doi: 10.3233/JAD-170120] [Medline: 28655135]
87. Rica RL, Shimojo GL, Gomes MC, Alonso AC, Pitta RM, Santa-Rosa FA, et al. Effects of a Kinect-based physical training program on body composition, functional fitness and depression in institutionalized older adults. Geriatr Gerontol Int 2020 Mar;20(3):195-200. [doi: 10.1111/ggi.13857] [Medline: 31923924]
88. Robert P, Albrengues C, Fabre R, Derreumaux A, Pancrazi MP, Luporsi I, et al. Efficacy of serious exergames in improving neuropsychiatric symptoms in neurocognitive disorders: results of the X-TORP cluster randomized trial. Alzheimers Dement (N Y) 2021 May 11;7(1):e12149 [FREE Full text] [doi: 10.1002/trc2.12149] [Medline: 34013018]
89. Min A, Min H, Kim S. Effectiveness of serious games in nurse education: a systematic review. Nurse Educ Today 2022 Jan;108:105178. [doi: 10.1016/j.nedt.2021.105178] [Medline: 34717098]
90. Sharifzadeh N, Kharrazi H, Nazari E, Tabesh H, Edalati Khodabandeh M, Heidari S, et al. Health education serious games targeting health care providers, patients, and public health users: scoping review. JMIR Serious Games 2020 Mar 05;8(1):e13459 [FREE Full text] [doi: 10.2196/13459] [Medline: 32134391]
91. Laver K, Liu E, Clemson L, Davies O, Gray L, Gitlin LN, et al. Does telehealth delivery of a dyadic dementia care program provide a noninferior alternative to face-to-face delivery of the same program? A randomized, controlled trial. Am J Geriatr Psychiatry 2020 Jun;28(6):673-682 [FREE Full text] [doi: 10.1016/j.jagp.2020.02.009] [Medline: 32234275]
92. Adomavičienė A, Daunoravičienė K, Kubilius R, Varžaitytė L, Raistenskis J. Influence of new technologies on post-stroke rehabilitation: a comparison of Armeo Spring to the Kinect system. Medicina (Kaunas) 2019 Apr 09;55(4):98 [FREE Full text] [doi: 10.3390/medicina55040098] [Medline: 30970655]
93. Tollár J, Nagy F, Hortobágyi T. Vastly different exercise programs similarly improve parkinsonian symptoms: a randomized clinical trial. Gerontology 2019;65(2):120-127. [doi: 10.1159/000493127] [Medline: 30368495]
94. Butler O, Herr K, Willmund G, Gallinat J, Kühn S, Zimmermann P. Trauma, treatment and Tetris: video gaming increases hippocampal volume in male patients with combat-related posttraumatic stress disorder. J Psychiatry Neurosci 2020 Jul 01;45(4):279-287 [FREE Full text] [doi: 10.1503/jpn.190027] [Medline: 32293830]
95. Abd-Alrazaq A, Alhuwail D, Al-Jafar E, Ahmed A, Shuweihdi F, Reagu SM, et al. The effectiveness of serious games in improving memory among older adults with cognitive impairment: systematic review and meta-analysis. JMIR Serious Games 2022 Aug 09;10(3):e35202 [FREE Full text] [doi: 10.2196/35202] [Medline: 35943792]
96. Kim Y, Hong S, Choi M. Effects of serious games on depression in older adults: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Sep 06;24(9):e37753 [FREE Full text] [doi: 10.2196/37753] [Medline: 36066964]
97. Tong T, Chignell M, Tierney MC, Lee J. A serious game for clinical assessment of cognitive status: validation study. JMIR Serious Games 2016 May 27;4(1):e7 [FREE Full text] [doi: 10.2196/games.5006] [Medline: 27234145]
98. Zheng J, Yu P, Chen X. An evaluation of the effects of active game play on cognition, quality of life and depression for older people with dementia. Clin Gerontol 2022 Oct 19;45(4):1034-1043. [doi: 10.1080/07317115.2021.1980170] [Medline: 34666621]
99. Adlakha S, Chhabra D, Shukla P. Effectiveness of gamification for the rehabilitation of neurodegenerative disorders. Chaos, Solitons Fractals 2020 Nov;140:110192. [doi: 10.1016/j.chaos.2020.110192]
100. Abd-Alrazaq A, Abuelezz I, Hassan A, AlSammarraie A, Alhuwail D, Irshaidat S, et al. Artificial intelligence-driven serious games in health care: scoping review. JMIR Serious Games 2022 Nov 29;10(4):e39840 [FREE Full text] [doi: 10.2196/39840] [Medline: 36445731]
101. Abt CC. Serious games. Am Behav Sci 2016 Jul 27;14(1):129-129. [doi: 10.1177/000276427001400113]
102. Park-Lee E, Sengupta M, Harris-Kojetin LD. Dementia special care units in residential care communities: United States, 2010. NCHS Data Brief 2013 Nov(134):1-8 [FREE Full text] [Medline: 24314070]
103. Chaudhury H, Murray K, Seetharaman K. Special care units. In: Encyclopedia of Gerontology and Population Aging. Cham: Springer; 2022:4722-4728.
104. Alzheimer's Association National Plan Care Support Milestone Workgroup, Borson S, Boustani MA, Buckwalter KC, Burgio LD, Chodosh J, et al. Report on milestones for care and support under the U.S. National Plan to Address Alzheimer's Disease. Alzheimers Dement 2016 Mar;12(3):334-369 [FREE Full text] [doi: 10.1016/j.jalz.2016.01.005] [Medline: 26868060]
105. Li F, Qin W, Zhu M, Jia J. Model-based projection of dementia prevalence in China and worldwide: 2020–2050. J Alzheimers Dis 2021 Aug 17;82(4):1823-1831. [doi: 10.3233/jad-210493]
106. Notice of the General Office of the National Health Commission on launching the Alzheimer's Disease Prevention and Treatment Promotion Action (2023-2025). General Office of the National Health Commission. 2023 May 26. URL: https://www.gov.cn/zhengce/zhengceku/202306/content_6886277.htm [accessed 2024-07-11]

107. Tang D, Lin Z, Chen F. Moving beyond living arrangements: the role of family and friendship ties in promoting mental health for urban and rural older adults in China. Aging Ment Health 2020 Sep;24(9):1523-1532. [doi: 10.1080/13607863.2019.1602589] [Medline: 30977378]
108. Bedford O, Yeh KH. History of Chinese indigenous psychology. In: Oxford Research Encyclopedia of Psychology. Oxford, UK: Oxford University Press; 2020.
109. Huang CS. Confucianism and Sacred Space: The Confucius Temple from Imperial China to Today. New York: Columbia University Press; 2021.
110. Zhu J, Guo Y. Social justice in spatial change: transition from autonomous rural development to integrated urbanization in China. Cities 2022 Mar;122:103539 [FREE Full text] [doi: 10.1016/j.cities.2021.103539]
111. Lu N, Spencer M, Sun Q, Lou VWQ. Family social capital and life satisfaction among older adults living alone in urban China: the moderating role of functional health. Aging Ment Health 2021 Apr 03;25(4):695-702. [doi: 10.1080/13607863.2019.1709155] [Medline: 31899943]
112. Ma X, Piao X, Oshio T. Impact of social participation on health among middle-aged and elderly adults: evidence from longitudinal survey data in China. BMC Public Health 2020 Apr 15;20(1):502 [FREE Full text] [doi: 10.1186/s12889-020-08650-4] [Medline: 32295560]
113. Fonseca X, Slingerland G, Lukosch S, Brazier F. Designing for meaningful social interaction in digital serious games. Entertainm Comput 2021 Jan;36:100385 [FREE Full text] [doi: 10.1016/j.entcom.2020.100385]
114. Jacobs RS. Serious games: play for change. In: The Video Game Debate 2. London, UK: Routledge; 2020:19-40.
115. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
116. van der Lubbe LM, Gerritsen C, Klein MCA, Hindriks KV. Empowering vulnerable target groups with serious games and gamification. Entertainm Comput 2021 May;38:100402 [FREE Full text] [doi: 10.1016/j.entcom.2020.100402]
117. Reis L, Mercer K, Boger J. Technologies for fostering intergenerational connectivity and relationships: scoping review and emergent concepts. Technol Soc 2021 Feb;64:101494 [FREE Full text] [doi: 10.1016/j.techsoc.2020.101494]
118. Zayeni D, Raynaud J, Revet A. Therapeutic and preventive use of video games in child and adolescent psychiatry: a systematic review. Front Psychiatry 2020;11:36 [FREE Full text] [doi: 10.3389/fpsyt.2020.00036] [Medline: 32116851]
119. David O, Costescu C, Cardos R, Mogoaşe C. How effective are serious games for promoting mental health and health behavioral change in children and adolescents? A systematic review and meta-analysis. Child Youth Care Forum 2020 Jul 31;49(6):817-838 [FREE Full text] [doi: 10.1007/s10566-020-09566-1]
120. Curley L, Lin J. Randomised controlled trials and pharmacy practice research. Pharm Pract Res Methods 2020:203-218. [doi: 10.1007/978-981-15-2993-1_10]
121. Mellis CM. How to choose your study design. J Paediatr Child Health 2020 Jul;56(7):1018-1022. [doi: 10.1111/jpc.14929] [Medline: 32479703]
122. Bärnighausen T, Tugwell P, Røttingen JA, Shemilt I, Rockers P, Geldsetzer P, et al. Quasi-experimental study designs series-paper 4: uses and value. J Clin Epidemiol 2017 Sep;89:21-29 [FREE Full text] [doi: 10.1016/j.jclinepi.2017.03.012] [Medline: 28365303]
123. Flannelly KJ, Flannelly LT, Jankowski KRB. Threats to the internal validity of experimental and quasi-experimental research in healthcare. J Health Care Chaplain 2018 Jan 24;24(3):107-130. [doi: 10.1080/08854726.2017.1421019] [Medline: 29364793]
124. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
125. Zhu K, Zhang Q, He B, Huang M, Lin R, Li H. Immersive virtual reality-based cognitive intervention for the improvement of cognitive function, depression, and perceived stress in older adults with mild cognitive impairment and mild dementia: pilot pre-post study. JMIR Serious Games 2022 Feb 21;10(1):e32117 [FREE Full text] [doi: 10.2196/32117] [Medline: 35188466]
126. Sen K, Prybutok G, Prybutok V. The use of digital technology for social wellbeing reduces social isolation in older adults: a systematic review. SSM Popul Health 2022 Mar;17:101020 [FREE Full text] [doi: 10.1016/j.ssmph.2021.101020] [Medline: 35024424]
127. Mobile statistics report, 2021-2025. The Radicati Group. URL: https://www.radicati.com/wp/wp-content/uploads/2021/Mobile_Statistics_Report,_2021-2025_Executive_Summary.pdf [accessed 2024-07-11]
128. Wainberg ML, Scorza P, Shultz JM, Helpman L, Mootz JJ, Johnson KA, et al. Challenges and opportunities in global mental health: a research-to-practice perspective. Curr Psychiatry Rep 2017 May;19(5):28 [FREE Full text] [doi: 10.1007/s11920-017-0780-z] [Medline: 28425023]

Abbreviations

AD: Alzheimer's disease
ADL: activities of daily living

CBT: cognitive behavioral therapy
CG: control group
CSDD: Cornell Scale for Depression in Dementia
DQoL: Dementia Quality of Life
EADG: Goldberg Anxiety and Depression Scale
EDG-15: 15-item Geriatric Depression Scale
GDS: Geriatric Depression Scale
IADL: Instrumental Activities of Daily Living
K-MBI: Korean version of the Modified Barthel Index
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
PICOS: Population, Intervention, Comparison, Outcomes, and Study Design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SCU: special care unit
SG: serious game
SGG: severe game group
SMD: standardized mean difference

Edited by T Leung, A Mavragani; submitted 24.12.23; peer-reviewed by X Zhao; comments to author 09.02.24; revised version received 19.02.24; accepted 26.06.24; published 31.07.24.

Please cite as:
Zuo X, Tang Y, Chen Y, Zhou Z
Effects of Electronic Serious Games on Older Adults With Alzheimer's Disease and Mild Cognitive Impairment: Systematic Review With Meta-Analysis of Randomized Controlled Trials
JMIR Serious Games 2024;12:e55785
URL: https://games.jmir.org/2024/1/e55785
doi:10.2196/55785
PMID:

Original Paper

Immersive Virtual Reality Use in Medical Intensive Care: Mixed Methods Feasibility Study

Brian W Locke[1,2], MD, MSc; Te-yi Tsai[3], MSc; C Mahony Reategui-Rivera[3], MD; Aileen S Gabriel[3], BS; Aref Smiley[3], PhD; Joseph Finkelstein[3], MD, PhD

[1]Division of Respiratory, Critical Care, and Occupational Pulmonary Medicine, Department of Internal Medicine, University of Utah, Salt Lake City, UT, United States

[2]Department of Pulmonary and Critical Care, Intermountain Medical Center, Murray, UT, United States

[3]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Brian W Locke, MD, MSc
Division of Respiratory, Critical Care, and Occupational Pulmonary Medicine
Department of Internal Medicine
University of Utah
30 N Mario Capecchi Dr
Salt Lake City, UT, 84112
United States
Phone: 1 801 581 7606
Email: brian.locke@imail.org

Abstract

Background: Immersive virtual reality (VR) is a promising therapy to improve the experience of patients with critical illness and may help avoid postdischarge functional impairments. However, the determinants of interest and usability may vary locally and reports of uptake in the literature are variable.

Objective: The aim of this mixed methods feasibility study was to assess the acceptability and potential utility of immersive VR in critically ill patients at a single institution.

Methods: Adults without delirium who were admitted to 1 of 2 intensive care units were offered the opportunity to participate in 5-15 minutes of immersive VR delivered by a VR headset. Patient vital signs, heart rate variability, mood, and pain were assessed before and after the VR experience. Pre-post comparisons were performed using paired 2-sided *t* tests. A semistructured interview was administered after the VR experience. Patient descriptions of the experience, issues, and potential uses were summarized with thematic analysis.

Results: Of the 35 patients offered the chance to participate, 20 (57%) agreed to partake in the immersive VR experience, with no difference in participation rate by age. Improvements were observed in overall mood (mean difference 1.8 points, 95% CI 0.6-3.0; *P*=.002), anxiety (difference of 1.7 points, 95% CI 0.8-2.7; *P*=.001), and pain (difference of 1.3 points, 95% CI 0.5-2.1; *P*=.003) assessed on 1-10 scales. The heart rate changed by a mean of –1.1 (95% CI –0.3 to –1.9; *P*=.008) beats per minute (bpm) from a baseline of 86.1 (SD 11.8) bpm and heart rate variability, assessed by the stress index (SI), changed by a mean of –5.0 (95% CI –1.5 to –8.5; *P*=.004) seconds^{-2} from a baseline SI of 40.0 (SD 23) seconds^{-2}. Patients commented on the potential for the therapy to address pain, lessen anxiety, and facilitate calmness. Technical challenges were minimal and there were no adverse effects observed.

Conclusions: Patient acceptance of immersive VR was high in a mostly medical intensive care population with little prior VR experience. Patients commented on the potential of immersive VR to ameliorate cognitive and emotional symptoms. Investigators can consider integrating minimally modified commercial VR headsets into the existing intensive care unit workflow to further assess VR's efficacy for a variety of endpoints.

(JMIR Serious Games 2024;12:e62842) doi:10.2196/62842

KEYWORDS

immersive virtual reality; intensive care unit; distraction therapy; virtual reality; mixed methods; feasibility study; semistructured interview; therapy; therapist; critical illness; critically ill; adult; patient acceptance; user experience; games for health; serious games; gamification

Introduction

Patients with critical illness experience many noxious sensations, stress, and restricted mobility while being treated in intensive care units (ICUs). There is a burgeoning evidence base demonstrating a loss in mental, emotional, and physical functioning after critical illness. Approaches to improve the experience of critical illness and functional outcomes after hospitalization are needed.

Immersive virtual reality (VR) has been proposed as a promising tool to address these issues [1,2]. Immersive VR often involves the use of a headset to project the viewer into an interactive artificial environment that elicits the feeling of embodiment in the artificial environment [3]. Preliminary work assessing the efficacy of immersive VR for physical and cognitive mobilization [4-8], sleep [9], distraction from pain [10,11], and mood [12,13] has been performed, with many further trials ongoing.

A major challenge to applying the results from prior studies of VR is that there are numerous permutations of how, when, and for whom immersive VR might be used. For example, the equipment [14], particular VR experience [14,15], and clinical purpose [3] may all vary. Potential barriers that might influence whether VR is accepted could be specific to the setting, providers, and patients for which the VR is being used [16]. Accordingly, widely variable uptake has been reported in prior studies [6,17,18]. Thus, it is difficult to infer, from the current literature, how acceptable VR might be in a particular situation.

In this study, we evaluated the feasibility of an immersive VR experience for patients admitted to one of 2 ICUs at a single institution. We hypothesized that patients with critical illness in the ICU would be interested in experiencing immersive VR; patients and staff would encounter minimal barriers to its use; and that VR usage would be associated with improvements in qualitative and quantitative accounts of mood, anxiety, and well-being.

Methods

Study Design and Setting

We performed a prospective, mixed methods, nonrandomized feasibility study of patients with critical illness using immersive VR headsets. The study was conducted in 2 ICUs (a 25-bed medical ICU and a 16-bed cancer-specialty ICU that cares for both medical and surgical critically ill patients with cancer) at a single institution in Salt Lake City, Utah, United States. Reporting follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidance for observational research (Multimedia Appendix 1)[19].

Ethical Considerations

The study was approved by the University of Utah Institutional Review Board (00170975). Patients were individually consented with a waiver of the requirement for written documentation of consent. Patients were not offered compensation for participation. Deidentified data were recorded in case report forms that were later digitized for analysis.

Recruitment

Patients who were potentially eligible for study inclusion were identified by attending physicians after daily rounds on days when study staff were available for enrollment. To be included, patients needed to be 18 years or older, admitted to the ICU, free from delirium [20] (as assessed by their providers, nurses, and able to pass an additional attention screen), and able to consent on their own behalf. Exclusion criteria included severe visual or auditory impairments (eg, legal blindness or deafness), isolation precautions for infection, recent condition that could be potentially exacerbated by VR (eg, seizure, uncontrolled nausea, traumatic brain injury, history of psychosis, or admission for a mental health crisis), or other craniofacial injury prohibiting headset use. Additionally, patients under current use of an orofacial mask to deliver positive airway pressure ventilation were excluded as the mask precluded headset use; however, intubated patients, patients receiving high-flow nasal cannula, and patients receiving oxygen via a regular face mask were eligible to participate.

After patients were identified as potential candidates, nursing staff were approached to identify any conflicting patient-care tasks (eg, physical therapy or travel for diagnostic testing or procedures). Patients were then approached about whether they were interested in trying an experimental immersive VR therapy. Patient demographics, reasons for admission, and comorbidities were assessed by chart review of clinical notes. All patients who were identified as potential research participants by their attending physicians were screened for inclusion and approached if eligible.

Experimental Protocol

Participating patients were presented with 3 visual analog scales (1-10) asking them to rate their current overall well-being, level of anxiety, and level of pain. Preintervention heart rate signals were recorded using a BIOPAC MP160 system for 5 minutes prior to VR initiation (Figure 1).

Figure 1. Experimental setup, as demonstrated by study personnel. Two laptops were used to ensure the adequacy of monitoring, project the real-time virtual reality experience of the patient to troubleshoot any technical issues or adverse effects, and to transcribe interview responses after the experience. The Meta Quest headset was chosen out of several possibilities owing to the lack of fabric components to facilitate cleaning after use.

Patients were offered 1 of 3 commercially available VR experiences delivered by a Meta Quest Pro headset (Meta): an urban travel experience (YouTube VR; Google LLC), a nature experience (Nature Treks VR; GreenerGames), or a synthetic landscape experience (TRIPP). In all 3 scenarios, the experiences involved passive exploration of the environment and did not involve use of the hand controllers (Figure 2). Patients planned to use the VR headset for at least 5 minutes, with an option to continue for up to 15 minutes if desired. Physiologic recording was continued for 5 minutes after VR use was complete.

After completing the VR experience, 5 minutes of postintervention vital signs recording and visual analog scale assessments of well-being, anxiety, and pain were administered. These scales were modeled on the Visual Analog Mood Scale [21], Numeric Visual Analog Anxiety Scale [22], and Verbal Numerical Rating Scale for pain [23], respectively (Multimedia Appendix 2).

Lastly, a semistructured qualitative interview was conducted to elicit feedback. A moderator guide (Multimedia Appendix 3) consisting of open-ended questions aimed at understanding the patient's overall experience using the VR headset was administered by a trained researcher. The responses for each participant were transcribed concurrently by 2 researchers in individual Microsoft Word documents. The participant responses were deidentified to maintain patient confidentiality.

Figure 2. Screenshots from the virtual reality experiences. Patients chose between one of the following experiences: YouTube VR (Google LLC; top panel), Nature Treks VR (GreenerGames; bottom left panel), or TRIPP (TRIPP Inc; bottom right panel).

Data Analysis

The qualitative data from the semistructured interviews were analyzed using thematic analysis, a 6-phase method for organizing, identifying, and summarizing patterns and themes [24]. In the first phase, known as data familiarization, researchers carefully reviewed, extracted, and organized the textual data in Microsoft Excel. The second phase involved generating initial codes for the data using a grounded theory approach, which explores participants' attitudes, beliefs, norms, and processes to develop hypotheses from the data rather than testing preexisting hypotheses [25]. In the third phase, the assigned codes were aggregated to uncover underlying patterns, themes, and subthemes. Phases 4 and 5 focused on reviewing, refining, and defining these themes, subthemes, and codes. Finally, in phase 6, the results of the thematic analysis were summarized and reported.

A sample size of 20 participating patients was targeted in accordance with rule-of-thumb guidance for a pilot study emphasizing qualitative assessment and protocol feasibility [26]. Descriptive statistics of participants and nonparticipants in VR were used to describe the population of interest. For physiologic signals, the first 2 minutes of pre-VR recording were compared to the first 2 minutes after the VR experience. Kubios automatic beat detection software was used to preprocess the heart rate data [27]. Heart rate variability was characterized by the square root of the Baevsky stress index (SI) [28]. Pre-post comparisons of vital sign data and mood assessments were performed using paired 2-sided *t* tests. Statistical analyses were performed using Stata version 18 (StataCorp) and the code is openly available on GitHub [29].

Results

Patients were enrolled between November 8, 2023, and February 6, 2024. A total of 35 patients were identified as potential candidates and approached, 20 (57%) of whom agreed to participate in the VR experience (Figure 3). Comorbidities and reasons for ICU admission mirrored those of the general ICU population (Table 1).

Characteristics of participants and nonparticipants are listed in Table 2. The age of participants (mean 61, SD 17 years) did not differ significantly (*P*=.33) from that of nonparticipants (mean 54, SD 22 years); however, nonparticipants were significantly more likely to have previously used VR (4/15 vs 1/20; *P*=.005).

Among the 20 patients consenting to participate, 19 completed at least 5 minutes of the VR experience (1 did not begin VR due to uncontrolled pain, but completed the pre-VR baseline data collection) and 18 completed all study assessments (1 had competing care needs prior to the interview). Participants used VR for a mean of 10 (SD 3) minutes. Among the 19 participants who completed the experience, 10 (53%) chose the travel experience, 5 (26%) chose the nature experience, and 4 (21%) chose the synthetic experience. No cybersickness or other adverse events occurred.

Among the 17 patients with valid heart rate data (1 patient had an excess artifact), the mean heart rate prior to initiation of the VR experience was 86.1 (SD 11.8) beats per minute (bpm), which decreased by 1.1 (95% CI 0.3-1.9; *P*=.008) bpm by the end of the experience. Mean heart rate variability, based on the SI, was 40 (SD 23) seconds^{-2} at baseline and decreased by 5.0 (95% CI 1.5 to 8.5; *P*=.008) seconds^{-2}.

At baseline, participating patients reported moderate overall well-being (mean 6.5, SD 2.1 on a 1-10 visual analog scale with 10 being the best; n=20), anxiety (mean 4.0, SD 2.8 with 1 indicating no anxiety; n=20), and pain (mean 3.9, SD 2.8 with 1 indicating no pain; n=20). Overall mood improved by a mean of 1.8 points (95% CI 0.65-3.0; *P*=.002; n=19) from baseline, anxiety decreased by 1.7 points (95% CI 0.8-2.7; *P*=.001; n=19), and pain decreased by 1.3 points (95% CI 0.53-2.1; *P*=.003; n=19) after use of immersive VR (Figure 4).

Figure 3. Enrollment flow diagram. VR: virtual reality.

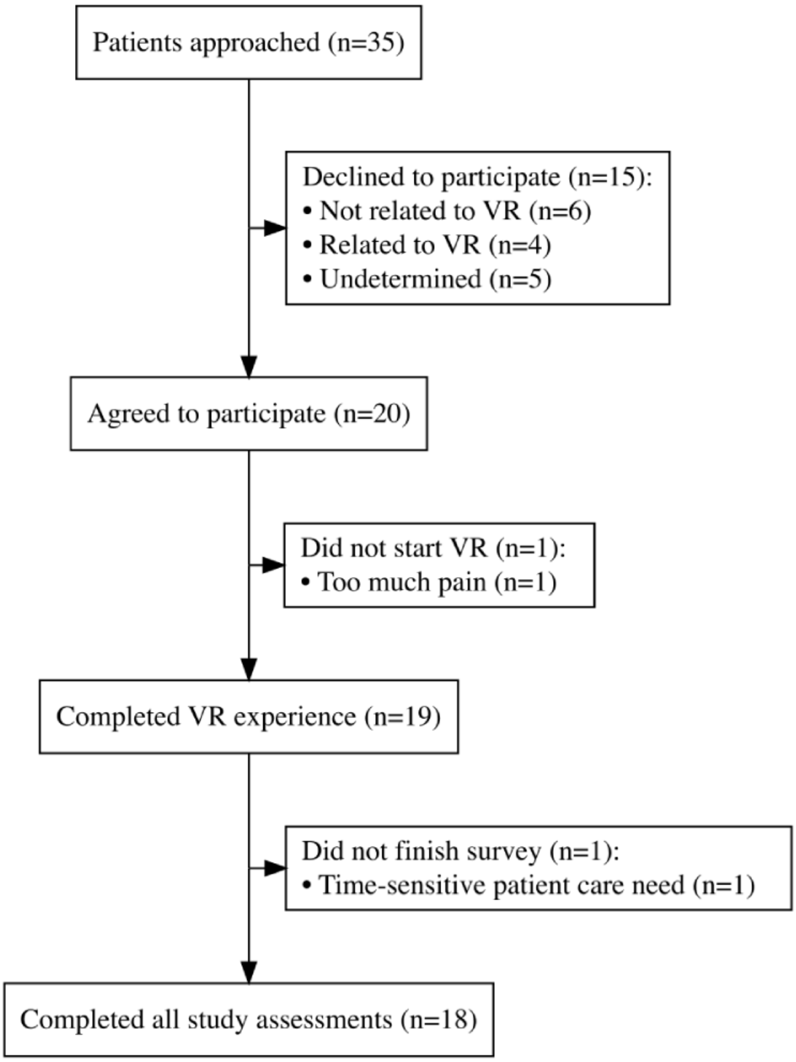


Table 1. Characteristics of patients approached (N=35).

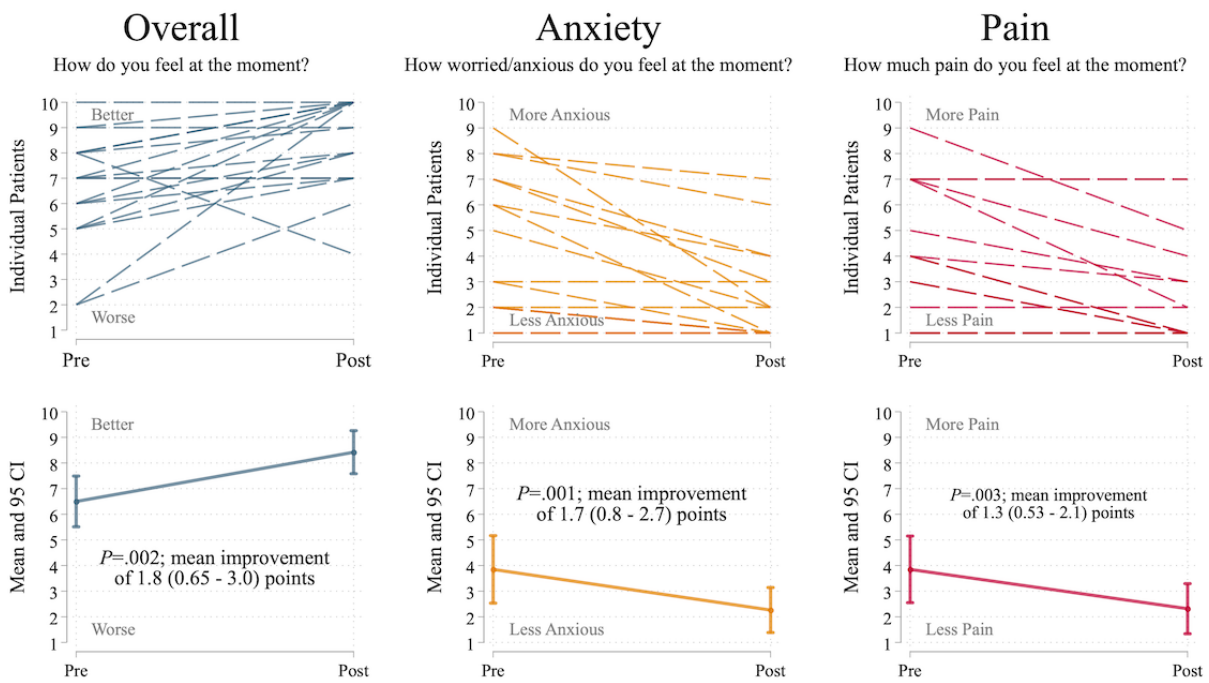
Characteristics	Value
Age (years), mean (SD)	58 (19)
Female, n (%)	11 (31)
Comorbidities, n (%)	
Type 2 diabetes	12 (34)
Atrial fibrillation	6 (17)
Chronic obstructive pulmonary disease	3 (9)
Congestive heart failure	7 (20)
Obstructive sleep apnea	2 (6)
Chronic kidney disease	3 (9)
Deep vein thrombosis	2 (6)
Cirrhosis	5 (14)
Cancer (any)	5 (14)
Common reasons for ICU [a] admission, n (%)	
Respiratory failure	10 (29)
Glucose/electrolytes	4 (11)
Thromboembolism	3 (9)
Gastrointestinal bleed	3 (9)
Heart failure	2 (6)
Sepsis	2 (6)

[a]ICU: intensive care unit.

Table 2. Characteristics of patients who declined or agreed to participate in the virtual reality (VR) experience.

Characteristics	Declined (n=15)	Agreed (n=20)	*P* value
Age (years), mean (SD)	54 (22)	61 (17)	.33
Female, n (%)	6 (40)	5 (25)	.34
Race/ethnicity, n (%)			.40
Asian	0 (0)	1 (5)	
Black	0 (0)	1 (5)	
Native Hawaiian/Pacific Islander	1 (9)	0 (0)	
White	10 (91)	18 (90)	
Hispanic	3 (27)	3 (15)	.41
Hearing impairment, n (%)	1 (11)	0 (0)	.13
Wears eyeglasses, n (%)	4 (44)	9 (45)	.98
Respiratory support, n (%)			.32
Face mask or variant	1 (9)	0 (0)	
High-flow nasal cannula	1 (9)	6 (30)	
Nasal cannula	5 (45)	9 (45)	
None	4 (36)	5 (25)	
Prior VR use, n (%)	4 (57)	1 (6)	.005

Figure 4. Overall mood, anxiety, and pain scores before and after use of immersive virtual reality. All scores were assessed using visual analog scales (see Multimedia Appendix 2) on a 1-10 scale, with 10 being best for overall mood and 1 being best for anxiety and pain. The paired t test was used for comparisons and 95% CIs of the mean change are presented in parentheses.



Five themes were commonly mentioned in the post-VR interviews (Table 3). Use of the VR headset to view relaxing scenery content was met with approval by all patients. Patients had a variety of positive responses with use of the VR headset, which was described as "good," "easy," "enjoyable," "comfortable," and "pleasant." These responses indicate overall acceptance and satisfaction.

Patients also identified potential benefits of VR in alleviating symptoms of anxiety and depression. Many reported a reduction in feelings of anxiety, nervousness, and isolation, suggesting a positive impact on mental well-being. Several patients highlighted that the VR intervention helped distract them from their current pain symptoms. The immersive nature and engaging VR content enabled patients to briefly "forget" about their pain and discomfort. The VR environment was described by patients as "calming," "relaxing," and "meditative." The opportunity to escape the ICU environment and immerse themselves in nature or travel scenery was highly valued, contributing to a sense of relaxation and calm.

Lastly, most patients reported having no problems using the VR headset. A few patients reported experiencing slight discomfort with the headset weight and difficulty seeing the side visuals. No patients reported motion sickness or claustrophobia.

Table 3. Themes and representative quotes from interviews after immersive virtual reality (VR) use among patients in the intensive care unit.

Theme	Representative quote
Acceptance of the VR system	"It was comforting, it was easy, from start to finish it's so calm, it felt so fast. I don't mind wearing it longer. it was my first time using it, I was amazed." [P18]
Improvement in mental health symptoms	"I think will help someone like me, facing what I face, sitting here for all days, waiting for my surgery, being anxious, not having anything to do, and not being allowed to eat or drink. I really appreciate you guys showing up. It really helps me." [P20]
Distraction from pain symptoms	"I forgot about the pain, some of it is there. I feel relaxed and comfortable, it took me out of my mind and I'm able to focus on the virtual world, I like what I see, it's so beautiful." [P09]
Feelings of relaxation, calmness	"Mostly the calmness, being calm helps you deal with your physical issues better, and the visual experience has a lot of benefits, it helped me forget that I have clog in my lung." [P16]
Problems using the VR headset	"No, don't have any problem with it, it could be lighter with the headset. The side part it's a bit blurry, but the other it's good. It will be good to wear my glasses." [P12]

Discussion

Principal Findings

In this feasibility study of commercially available immersive VR use among critically ill patients, we found that most patients did not have prior experience with VR technology but had high levels of interest and acceptance of the technology. Participants commented on VR's potential to alleviate cognitive and emotional symptoms. Changes in heart rate variability were consistent with increased relaxation. Only minor technical challenges and no severe adverse effects were noted.

Comparison to Prior Work

Prior work has explored the potential for immersive VR for several ICU use cases [1]. However, this previous work reported widely variable rates of uptake and hinted at a "digital divide," where older patients may be less interested in new technology [30]. In contrast, we found high levels of participation among patients of all ages. Thus, VR might be considered for study even in settings that frequently care for older adults. We found that patients who had previously used VR were less likely to participate in our study. However, this might indicate that the novelty of the experience drove participation in this feasibility study. Interest may improve in prior users if VR were offered as a validated treatment. Only a minority of patients who did not participate gave reasons related to the VR itself.

After use, several participants commented on the potential of the therapy to address anxiety and foster relaxation or calmness. This was corroborated by the improvement in vital sign correlates of relaxation, such as improved heart rate variability. Some [6], but not all [18], prior studies of VR have shown consistent changes in vital signs.

Participants also highlighted the potential of VR to distract from pain, which is consistent with current guidance from the Society of Critical Care Medicine that recommends consideration of "cybertherapy [VR]" for this purpose [31]. The high usability scores and low rate of technical challenges with "off-the-shelf" commercially available options suggests that extensive customization is not a prerequisite for use. Furthermore, we did not observe claustrophobia, nausea, or "cybersickness" in any patients. Cybersickness may be less common with more modern VR headset technology that minimizes latency [32] and discordance between virtual and actual head positioning [33], which could explain why this was not encountered in our study. In contrast to prior work suggesting that nature scenes may maximize relaxation [15], travel was the most frequent VR experience choice among our participants. The potential for VR to enable "escape" from the ICU was also frequently mentioned in qualitative interviews.

Strengths and Limitations

Strengths of this study include the assessment of both quantitative and qualitative dimensions of the VR experience, which both support the feasibility and potential utility of using VR in this group. In addition, there were few exclusion criteria and thus the sample of patients is expected to broadly represent characteristics of critically ill patients without delirium. Lastly, we demonstrated and reported an off-the-shelf and replicable experimental setup that investigators can use as a starting point for future studies.

However, several limitations also deserve mention. First, nonrepresentative sampling of research participants may have occurred due to only enrolling patients during select times when study staff were available. Despite exhaustive screening of patients who were identified as potential candidates during these times, we cannot fully describe what characteristics may have influenced which patients were judged to be potential candidates for the study by attending physicians. Second, there was no control group. Time trends and the influence of conversing with study staff may also contribute to changes in symptoms and vital signs, although we attempted to stabilize physiologic trends with a 5-minute run-in period and structured interview guides were used to focus the conversation. Third, prior work suggested that VR's effectiveness correlates with the degree of immersion [34], which we did not directly assess. However, we chose not to attempt to eliminate distractions to better emulate the conditions and degree of immersion expected in actual use. Fourth, our quantitative outcome scales are slightly modified from previously validated work, and thus the reported effect sizes should not be compared to other interventions or established minimally important differences. Lastly, the uptake of VR may differ when it is offered as a treatment for specific conditions as opposed to offering the opportunity to help assess feasibility, particularly among patients who have previously used the technology.

Future Directions

This work has several important implications for the study of VR in the ICU. We used commercially available technology and found that the sessions were acceptable to patients. This suggests that customization of either software or hardware is not necessarily required for some VR use cases. Furthermore, we found high interest among critically ill patients. Lastly, we encountered minimal disruptions to patient care or study protocols, with 95% of the patients completing the experience and 90% completing all study assessments. This suggests that protocols to study the impact of VR can be integrated into usual ICU care with little impingement on clinical workflows. Evaluation of the usability in patients excluded from the current work, such as those with mild delirium or more severe emotional symptoms, could help establish the potential of studying immersive VR in additional high-risk populations.

Conclusions

We found that a relatively short session of off-the-shelf immersive VR is acceptable to critically ill patients; resulted in improved pain, anxiety, and overall mood scores; and did not result in side effects or present major technical challenges. This work suggests that studies on the effect of VR on patient-relevant outcomes in the context of critical illness are feasible. Investigators can consider VR study protocols that do not involve substantial technology customization or large changes to patient care workflows.

Acknowledgments

This research was supported by the National Institutes of Health under the Ruth L. Kirschstein National Research Service Award (5T32HL105321) to BWL and a National Institutes of Health grant (R33HL143317) to JF.

Data Availability

Data can be provided upon Institutional Review Board–approved request to Joseph.Finkelstein@utah.edu.

Authors' Contributions

BWL contributed to study conceptualization, formal analysis, investigation, writing the original manuscript draft, manuscript preparation, and data visualization. TT and CMRR contributed to methodology, software, investigation, and data curation. ASG performed the formal analysis. AS contributed to data curation and formal analysis. JF contributed to study conceptualization, methodology, resources, manuscript review and editing, supervision, and project administration. All authors reviewed and approved the final manuscript.

Conflicts of Interest

BWL receives grant funding from the American Thoracic Society and holds financial stake in Mountain Biometrics, Inc, which focuses on applications of machine learning on physiologic sensor data unrelated to the current project. All other authors declare no conflicts of interest.

Multimedia Appendix 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist for reporting of observational research.
[DOCX File , 33 KB - games_v12i1e62842_app1.docx]

Multimedia Appendix 2
Visual analog scales used for assessing mood, anxiety, and pain before and after the virtual reality experience.
[PDF File (Adobe PDF File), 1457 KB - games_v12i1e62842_app2.pdf]

Multimedia Appendix 3
Interview guide for the qualitative assessment after the virtual reality experience.
[PDF File (Adobe PDF File), 98 KB - games_v12i1e62842_app3.pdf]

References

1. Hill JE, Twamley J, Breed H, Kenyon R, Casey R, Zhang J, et al. Scoping review of the use of virtual reality in intensive care units. Nurs Crit Care 2022 Nov;27(6):756-771. [doi: 10.1111/nicc.12732] [Medline: 34783134]
2. Vlake JH, Van Bommel J, Wils E, Korevaar TIM, Bienvenu OJ, Klijn E, et al. Virtual reality to improve sequelae of the postintensive care syndrome: a multicenter, randomized controlled feasibility study. Crit Care Explor 2021 Oct;3(9):e0538 [FREE Full text] [doi: 10.1097/CCE.0000000000000538] [Medline: 34549192]
3. Bruno RR, Wolff G, Wernly B, Masyuk M, Piayda K, Leaver S, et al. Virtual and augmented reality in critical care medicine: the patient's, clinician's, and researcher's perspective. Crit Care 2022 Oct 25;26(1):326 [FREE Full text] [doi: 10.1186/s13054-022-04202-x] [Medline: 36284350]
4. Navarra-Ventura G, Gomà G, de Haro C, Jodar M, Sarlabous L, Hernando D, et al. Virtual reality-based early neurocognitive stimulation in critically ill patients: a pilot randomized clinical trial. J Pers Med 2021 Nov 29;11(12):1260 [FREE Full text] [doi: 10.3390/jpm11121260] [Medline: 34945732]
5. Gerber SM, Jeitziner M, Knobel SEJ, Mosimann UP, Müri RM, Jakob SM, et al. Perception and performance on a virtual reality cognitive stimulation for use in the intensive care unit: a non-randomized trial in critically ill patients. Front Med 2019 Dec 10;6:287 [FREE Full text] [doi: 10.3389/fmed.2019.00287] [Medline: 31921867]
6. Turon M, Fernandez-Gonzalo S, Jodar M, Gomà G, Montanya J, Hernando D, et al. Feasibility and safety of virtual-reality-based early neurocognitive stimulation in critically ill patients. Ann Intensive Care 2017 Dec 2;7(1):81 [FREE Full text] [doi: 10.1186/s13613-017-0303-4] [Medline: 28770543]
7. Suvajdzic M, Bihorac A, Rashidi P, Rashidi P, Ruppert MM, Williams S, et al. Developing a patient-centered virtual reality healthcare system to prevent the onset of delirium in ICU patients. 2019 Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); August 5-7, 2019; Kyoto, Japan. [doi: 10.1109/segah.2019.8882442]
8. Kho ME, Damluji A, Zanni JM, Needham DM. Feasibility and observed safety of interactive video games for physical rehabilitation in the intensive care unit: a case series. J Crit Care 2012 Apr;27(2):219. [doi: 10.1016/j.jcrc.2011.08.017] [Medline: 21944880]

9. Lee SY, Kang J. Effect of virtual reality meditation on sleep quality of intensive care unit patients: a randomised controlled trial. Intensive Crit Care Nurs 2020 Aug;59:102849. [doi: 10.1016/j.iccn.2020.102849] [Medline: 32241625]
10. Faber AW, Patterson DR, Bremer M. Repeated use of immersive virtual reality therapy to control pain during wound dressing changes in pediatric and adult burn patients. J Burn Care Res 2013;34(5):563-568 [FREE Full text] [doi: 10.1097/BCR.0b013e3182777904] [Medline: 23970314]
11. Spiegel B, Fuller G, Lopez M, Dupuy T, Noah B, Howard A, et al. Virtual reality for management of pain in hospitalized patients: a randomized comparative effectiveness trial. PLoS One 2019;14(8):e0219115 [FREE Full text] [doi: 10.1371/journal.pone.0219115] [Medline: 31412029]
12. Haley A, Wacker D. Cinematic virtual reality for anxiety management in mechanically ventilated patients: a feasibility and pilot study. Acute Crit Care 2022 May;37(2):230-236 [FREE Full text] [doi: 10.4266/acc.2021.00843] [Medline: 35172527]
13. Rodrigues IM, Lima AG, Santos AED, Santos ACA, Nascimento LSD, Serra MVCL, et al. A single session of virtual reality improved tiredness, shortness of breath, anxiety, depression and well-being in hospitalized individuals with COVID-19: a randomized clinical trial. J Pers Med 2022 May 19;12(5):829 [FREE Full text] [doi: 10.3390/jpm12050829] [Medline: 35629250]
14. Merliot-Gailhoustet L, Raimbert C, Garnier O, Carr J, De Jong A, Molinari N, et al. Discomfort improvement for critically ill patients using electronic relaxation devices: results of the cross-over randomized controlled trial E-CHOISIR (Electronic-CHOIce of a System for Intensive care Relaxation). Crit Care 2022 Sep 03;26(1):263 [FREE Full text] [doi: 10.1186/s13054-022-04136-4] [Medline: 36057612]
15. Gerber SM, Jeitziner M, Sänger SD, Knobel SEJ, Marchal-Crespo L, Müri RM, et al. Comparing the relaxing effects of different virtual reality environments in the intensive care unit: observational study. JMIR Perioper Med 2019 Nov 15;2(2):e15579 [FREE Full text] [doi: 10.2196/15579] [Medline: 33393906]
16. Twamley J, Hamer O, Hill J, Kenyon R, Twamley H, Casey R, et al. Exploring the perceptions of former ICU patients and clinical staff on barriers and facilitators to the implementation of virtual reality exposure therapy: a qualitative study. Nurs Crit Care 2024 Mar;29(2):313-324. [doi: 10.1111/nicc.12868] [Medline: 36458458]
17. Jawed YT, Golovyan D, Lopez D, Khan SH, Wang S, Freund C, et al. Feasibility of a virtual reality intervention in the intensive care unit. Heart Lung 2021;50(6):748-753 [FREE Full text] [doi: 10.1016/j.hrtlng.2021.05.007] [Medline: 34217986]
18. Ong T, Ruppert M, Akbar M, Rashidi P, Ozrazgat-Baslanti T, Bihorac A, et al. Improving the intensive care patient experience with virtual reality-a feasibility study. Crit Care Explor 2020 Jun;2(6):e0122 [FREE Full text] [doi: 10.1097/CCE.0000000000000122] [Medline: 32695991]
19. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007 Oct 16;147(8):573-577 [FREE Full text] [doi: 10.7326/0003-4819-147-8-200710160-00010] [Medline: 17938396]
20. Kellmeyer P, Biller-Andorno N, Meynen G. Ethical tensions of virtual reality treatment in vulnerable patients. Nat Med 2019 Aug 29;25(8):1185-1188. [doi: 10.1038/s41591-019-0543-y] [Medline: 31359003]
21. Luria RE. The validity and reliability of the visual analogue mood scale. J Psychiatr Res 1975 May;12(1):51-57. [doi: 10.1016/0022-3956(75)90020-5] [Medline: 806683]
22. Davey HM, Barratt AL, Butow PN, Deeks JJ. A one-item question with a Likert or visual analog scale adequately measured current anxiety. J Clin Epidemiol 2007 Apr;60(4):356-360. [doi: 10.1016/j.jclinepi.2006.07.015] [Medline: 17346609]
23. Bijur PE, Latimer CT, Gallagher EJ. Validation of a verbally administered numerical rating scale of acute pain for use in the emergency department. Acad Emerg Med 2003 Apr;10(4):390-392 [FREE Full text] [doi: 10.1111/j.1553-2712.2003.tb01355.x] [Medline: 12670856]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Glaser B, Strauss A. Discovery of Grounded Theory: Strategies for Qualitative Research. New York: Routledge; 2017.
26. Birkett MA, Day SJ. Internal pilot studies for estimating sample size. Stat Med 1994 Oct 15;13(23-24):2455-2463. [doi: 10.1002/sim.4780132309] [Medline: 7701146]
27. Tarvainen MP, Niskanen J, Lipponen JA, Ranta-Aho PO, Karjalainen PA. Kubios HRV--heart rate variability analysis software. Comput Methods Programs Biomed 2014 Jan;113(1):210-220. [doi: 10.1016/j.cmpb.2013.07.024] [Medline: 24054542]
28. Baevsky, Chernikova AG. Heart rate variability analysis: physiological foundations and main methods. Cardiometry 2017 May(10):66-76. [doi: 10.12710/cardiometry.2017.10.6676]
29. VR_ICU statistical code. GitHub. URL: https://github.com/reblocke/vr_icu [accessed 2024-07-26]
30. Mosadeghi S, Reid MW, Martinez B, Rosen BT, Spiegel BMR. Feasibility of an immersive virtual reality intervention for hospitalized patients: an observational cohort study. JMIR Ment Health 2016 Jun 27;3(2):e28 [FREE Full text] [doi: 10.2196/mental.5801] [Medline: 27349654]

31. Devlin J, Skrobik Y, Gélinas C, Needham D, Slooter A, Pandharipande P, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. Crit Care Med 2018 Sep;46(9):e825-e873. [doi: 10.1097/CCM.0000000000003299] [Medline: 30113379]
32. Stauffert J, Niebling F, Latoschik M. Latency and cybersickness: impact, causes, and measures. A review. Front Virtual Real 2020 Nov 26;1:582204. [doi: 10.3389/frvir.2020.582204]
33. Palmisano S, Allison RS, Kim J. Cybersickness in head-mounted displays is caused by differences in the user's virtual and physical head pose. Front Virtual Real 2020 Nov 12;1:587698. [doi: 10.3389/frvir.2020.587698]
34. Phelan I, Furness P, Fehily O, Thompson A, Babiker N, Lamb M, et al. A mixed-methods investigation into the acceptability, usability, and perceived effectiveness of active and passive virtual reality scenarios in managing pain under experimental conditions. J Burn Care Res 2019 Jan 01;40(1):85-90. [doi: 10.1093/jbcr/iry052] [Medline: 30247616]

Abbreviations

bpm: beats per minute
ICU: intensive care unit
SI: stress index
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VR: virtual reality

Edited by T Leung; submitted 02.06.24; peer-reviewed by F Shaya, R Moskovitch; comments to author 12.07.24; revised version received 19.07.24; accepted 22.07.24; published 09.08.24.

Please cite as:
Locke BW, Tsai TY, Reategui-Rivera CM, Gabriel AS, Smiley A, Finkelstein J
Immersive Virtual Reality Use in Medical Intensive Care: Mixed Methods Feasibility Study
JMIR Serious Games 2024;12:e62842
URL: https://games.jmir.org/2024/1/e62842
doi:10.2196/62842
PMID:39046869

Original Paper

An Evidence-Based Serious Game App for Public Education on Antibiotic Use and Resistance: Randomized Controlled Trial

Zhilian Huang[1], MPH, PhD; Jing Teng Ow[1], BPsych; Wern Ee Tang[2], MMed, MBBS; Angela Chow[1,3,4], MBBS, MS, PhD

[1]Department of Preventive and Population Medicine, Tan Tock Seng Hospital, Singapore, Singapore
[2]Clinical Research Unit, National Healthcare Group Polyclinics, Singapore, Singapore
[3]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[4]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Corresponding Author:
Zhilian Huang, MPH, PhD
Department of Preventive and Population Medicine
Tan Tock Seng Hospital
11 Jalan Tan Tock Seng
Singapore, 308322
Singapore
Phone: 65 6256 6011
Email: zhilian_huang@ttsh.com.sg

Abstract

Background: The misuse and overuse of antibiotics accelerate the development of antimicrobial resistance (AMR). Serious games, any form of games that serve a greater purpose other than entertainment, could augment public education above ongoing health promotion efforts. Hence, we developed an evidence-based educational serious game app—SteWARdS Antibiotic Defence—to educate players on good antibiotic use practices and AMR through a game quest comprising 3 minigames and interaction with the nonplayer characters.

Objective: We aimed to evaluate the effectiveness of the SteWARdS Antibiotic Defence app in improving the knowledge of, attitude toward, and perceptions (KAP) of appropriate antibiotic use and AMR among the public in Singapore.

Methods: We conducted a 2-arm parallel randomized controlled trial, recruiting visitors aged 18-65 years from 2 polyclinics in Singapore. Intervention group participants had to download the SteWARdS Antibiotic Defence app (available only in English and on the Android platform) on their smartphones and complete the quest in the app. Participants took half a day to 2 weeks to complete the quest. The control group received no intervention. Knowledge questions on antibiotic use and AMR (11 binary questions) were self-administered at baseline, immediately after the intervention, and 6-10 weeks post intervention, while attitudes and perception questions (14 three-point Likert-scale questions) were self-administered at baseline and 6-10 weeks post intervention. We also collected participants' feedback on app usage.

Results: Participants (n=348; intervention: n=142, control: n=206) had a mean age of 36.9 years. Intervention group participants showed a statistically significant improvement in mean knowledge score (effect size: 0.58 [95% CI 0.28-0.87]) compared with controls after accounting for age, educational level, and exposure to advertisements on antibiotics and AMR. Intervention participants also showed a statistically significant improvement in mean attitude-perception scores (effect size: 0.98 (95% CI 0.44-1.52)) after adjusting for marital status and race. A majority of participants agreed that the "SteWARdS Antibiotic Defence" app improved their awareness on antibiotic use (135/142, 95.1%) and AMR (136/142, 95.8%). About 73.9% (105/142) of the participants agreed that the app is easy to use, 70.4% (100/142) agreed that the app was enjoyable, and 85.2% (121/142) would recommend the app to others.

Conclusions: Our educational serious game app improves participants' KAP on appropriate antibiotic use and AMR. Public education apps should be engaging, educational, easy to use, and have an attractive user interface. Future research should assess the effectiveness of interventions in facilitating long-term knowledge retention and long-lasting behavioral change.

Trial Registration: ClinicalTrials.gov NCT05445414; https://clinicaltrials.gov/ct2/show/NCT05445414

International Registered Report Identifier (IRRID): RR2-10.2196/45833

(*JMIR Serious Games 2024;12:e59848*) doi:10.2196/59848

KEYWORDS

serious game application; randomized controlled trial; antimicrobial resistance; antibiotic use; public education; mobile phone

Introduction

The World Health Organization declared antimicrobial resistance (AMR) as one of the top 10 global public health threats in 2019 [1]. Antibiotic-resistant bacteria can cause human infections that are harder to treat, leading to higher medical costs, decreased work productivity, and increased mortality [2-4]. By 2050, AMR is projected to cause 10 million annual deaths and up to US $100.2 trillion in economic losses worldwide if nothing is done to slow its progression [5]. The widespread misuse and overuse of antibiotics, often driven by the public's lack of knowledge regarding appropriate antibiotic use and AMR, contribute to AMR progression [6,7].

The need to educate the public on appropriate antibiotic use and AMR is apparent, as better knowledge of antibiotic use was found to be associated with favorable antibiotic attitudes, which lowered the odds of expecting and receiving an antibiotic prescription [8]. Despite yearly public outreach [9] and multiple national educational campaigns on antibiotics and AMR [10], misconceptions about antibiotics—such as believing that antibiotics treat viral infections—continue to persist among the Singaporean population [7,8]. Traditional efforts to educate the public on antibiotic use and AMR have shown ambiguous effectiveness [11,12]. Hence, the lackluster effectiveness of existing efforts to prevent antibiotic misuse and overuse provides an impetus to explore novel public education methods beyond traditional modalities (eg, posters and pamphlets) [13].

Serious games have emerged as a promising tool for health education and health promotion across various fields recently [14]. This learning modality encompasses online and offline tools that use gamification—the use of game-playing elements—to provide an enhanced learning experience. Studies have shown promise of serious games in improving the short-term health knowledge, attitudes, and beliefs of young people, but the studies were too heterogeneous to prove the efficacy of this learning modality [15,16]. There were fewer serious games studies with older adults, but a web-based serious game co-designed with members of the public to raise pancreatic cancer awareness found a statistically significant improvement of pancreatic cancer awareness [17].

Serious games have also been used to promote good antibiotic use behaviors [16,18]. Preliminary evidence of serious games interventions on antibiotic use and AMR has shown promise in improving the knowledge of appropriate antibiotic use [16,19]. However, these studies were primarily focused on children and students. Adults have greater access to antibiotics and are more likely to pass down antibiotic use behaviors to the younger generation [13]. Hence, it is essential to educate young adults on appropriate antibiotic use. Existing studies also limited participants to a controlled environment with restricted playing time and may not fully capture real-world behavior. Furthermore, many studies overlooked the importance of long-term knowledge retention. The paucity of high-quality randomized controlled trials (RCTs) in the real-world setting highlights a pertinent gap in the evidence base of serious games for public education.

Singapore offers an excellent environment for evaluating serious game applications (apps) given its high smartphone penetration rate (92%) [20] and emphasis on digital transformation [21]. Hence, we developed an evidence-based serious game app—"SteWARdS Antibiotic Defence"—to deliver adult education on appropriate antibiotic use and AMR and aimed to evaluate its effectiveness in improving knowledge of, attitude toward, and perceptions (KAP) of antibiotic use and AMR in Singapore. We also sought user feedback on the "SteWARdS Antibiotic Defence" app.

Methods

Study Design

We conducted a parallel 2-arm RCT with a 6- to 10-week follow-up period. Participants in the intervention group had to download and complete the quest in the "SteWARdS Antibiotic Defence" app on their smartphones, while controls received no intervention.

Study Setting

We recruited patients and their caregivers from 2 government-funded primary care clinics (ie, polyclinics) in Singapore. Recruitment occurred from January to March 2023, and follow-up was completed in early June 2023. The 23 polyclinics in Singapore manage 20% of Singapore's primary health care needs by providing subsidized services such as medical treatment for acute and chronic conditions, vaccinations, and health education [22]. Participants seen at the polyclinics mainly have acute conditions that are managed outside the health care setting. Hence, while our study recruitment and baseline data collection occurred in the policlinics, the study intervention and follow-ups were administered outside the health care setting.

Eligibility Criteria

Participants were 18-65 years of age, had access to an Android smartphone, and were English literate. We excluded those with visual or cognitive impairment, not proficient with smartphone apps, and those unable to install Android mobile apps from the study.

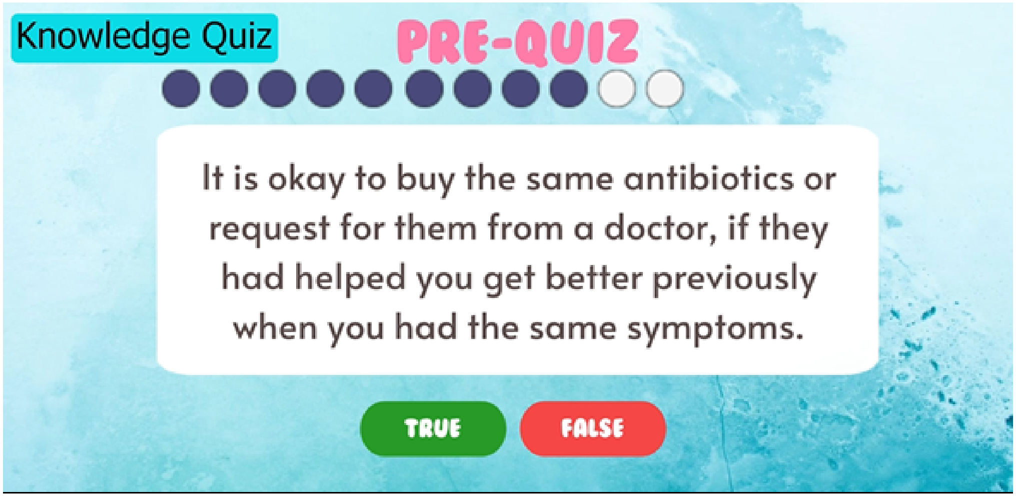
Intervention

App Development

The "SteWARdS Antibiotic Defence" app is an evidence-based serious game mobile app developed by the study team, in collaboration with Temasek Polytechnic (a public tertiary institution in Singapore), to educate the public on appropriate antibiotic use and AMR. Players learn about appropriate antibiotic use and effective methods to recover from uncomplicated upper respiratory tract infections by interacting with the nonplayer characters (NPCs) releasing bite-sized information and playing the minigames in the app. We derived the educational content in the app from a rigorous review of knowledge gaps [7,8,23], antibiotic guidance [24], and inputs from health care professionals. The in-app messages were also adapted to the local context to ensure their relevance to players [25]. A few knowledge questions in the postgame and postintervention surveys were reverse scored to reduce to learning effect on knowledge gains. The bite-sized messages and knowledge questions can be found in Tables S1 and S2 in Multimedia Appendix 1.

The pre- and postgame knowledge questions were embedded at the beginning and the end of the quest to reinforce knowledge retention (Figure 1). Completing the 11 baseline knowledge questions is a prerequisite to proceed to the minigames while completing the postgame knowledge questions is a prerequisite to fulfill the requirements of the intervention. Players are allowed to play the minigames repeatedly once they pass the level, but they can complete the pre-post survey only once. There is an in-app WhatsApp helpline for participants to contact the study team should they encounter technical difficulties. The feedback form link is also embedded in the app and appears after the completion of the postknowledge survey. Study team members (data collectors) provided intervention group participants with a onetime unique code to access the app during recruitment. The unique code allowed us to track the pre-post knowledge survey responses, time spent on the app, and interactions with the NPCs anonymously. We did not analyze the time spent on the app due to glitches that affected data quality.

Figure 1. An example of a knowledge question embedded in the app. Participants are expected to complete 11 preintervention knowledge questions, followed by the game quest and 11 post–intervention knowledge questions to complete the intervention.



Game Components

There are 3 "worlds"—the "Supermarket," "Park," and "Train station"—that lead to the minigames in the app. Each "world" comprises 2 NPCs and a minigame. The supermarket is linked to the Tower Defence game; park is linked to the Match3 game; and train station is linked to the Endless Runner game (Figures 2-5).

Figure 2. Interaction with a nonplayer character (NPC) at the park. Bite-sized messages on good antibiotic use practices were released through interactions with the NPCs. The rightmost NPC leads to the Tower Defence game.

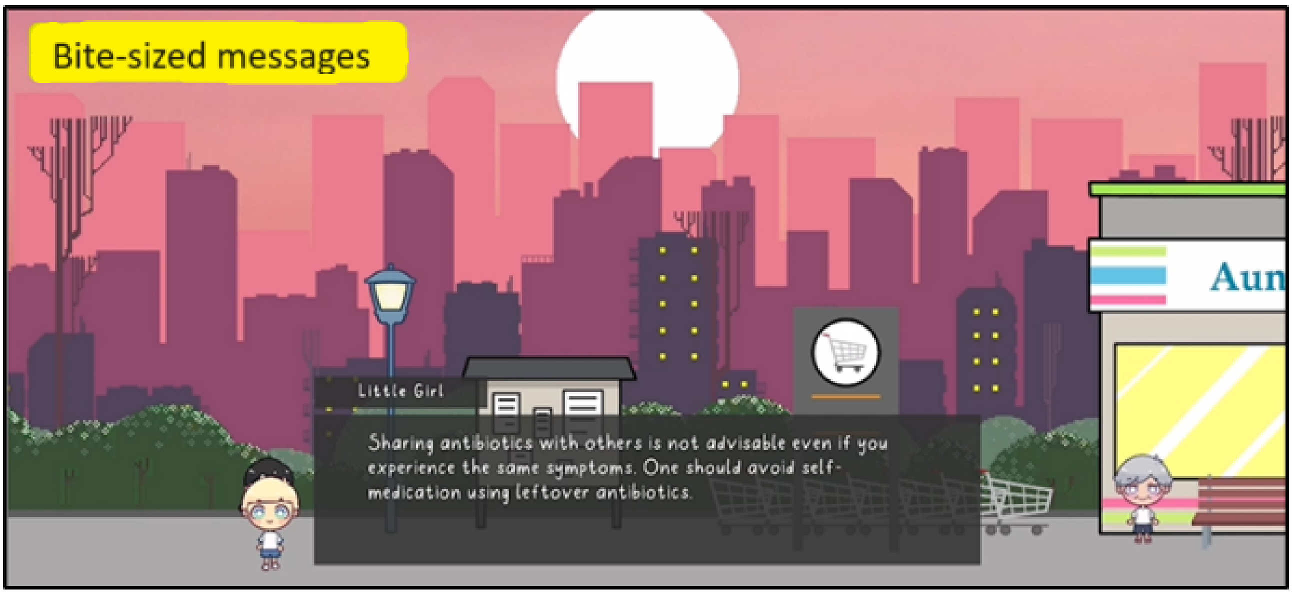


Figure 3. The Tower Defence game consists of 9 levels. The top figure shows the game interface with C-Stin turrets and the super bug (level 9). The subsequent figures show the messages released at the beginning of level 9 to teach players about the concept of antimicrobial resistance. Players have to choose the appropriate turrets and strategically place them at appropriate locations to effectively target their enemies (bacteria or viruses).

Figure 4. The Endless Runner game can be accessed via the train station (picture on top). The game reinforces the message that antibiotics are effective only against bacteria and not viruses. Players collect coins while running and use the antibiotic gun or immunity gun to shoot at bacteria or virus obstacles, respectively. Coins collected from the Endless Runner game can be used to purchase turrets in the Tower Defence game.

Figure 5. The Match3 game can be accessed via the park (picture on top). The game educates players on good health habits by focusing on sleep hygiene, nutrition, immunization, and physical exercise. Players are required to match 3 or more similar tiles within a specific time limit to successfully clear each of the 10 levels. Bite-sized messages are released before the start of each game to remind players of maintaining a healthy lifestyle.

There are 9 levels in the Tower Defence game to teach players about antibiotics and AMR. The first 2 levels teach players that antibiotics are effective against bacteria and the next 2 introduce the concept that antibiotics are bacteria-specific. Levels 5 and 6 educate players that antibiotics are not effective against viruses, levels 7 and 8 are reinforcements of previously taught concepts, and the last level educates players on antibiotic resistance by introducing a superbug that requires an expensive antibiotic (turret) to cure (kill the superbug). Players have to earn coins by passing Tower Defence levels and playing the Match3 and Endless Runner games.

The Endless Runner game is accessible via the train station. The game reinforces the message that antibiotics are effective only against bacteria and not viruses. Players collect coins while running to purchase turrets in the Tower Defence game, and use the antibiotic gun to shoot away bacteria obstacles and the immunity gun to shoot the virus obstacles.

The Match3 game is accessible via the park. The game educates players on good health habits by focusing on sleep hygiene, nutrition, immunization, and exercise. Players are required to match 3 or more similar tiles within a specific time limit to successfully clear each of the 10 levels.

Further details of the design and principles of the "SteWARdS Antibiotic Defence" app can be found in our published protocol [25]. Individual participants were followed up according to recruitment schedule from the point of recruitment. Participants took between half a day to 2 weeks to complete the intervention.

Outcomes

The primary outcome of interest was the change in KAP scores of antibiotic use and AMR. The secondary outcome was the level of user satisfaction for the "SteWARdS Antibiotic Defence" app. Primary outcomes were assessed for the intervention and control groups at baseline, immediately post intervention (intervention group only), and 6-10 weeks post intervention (from baseline for controls) via self-administered questionnaires. Knowledge was measured via 11 True/False questions, while attitudes and perceptions were measured via 14 questions on a 3-point Likert scale. Details of the KAP questionnaire are available in our published protocol [25]. For the intervention group, satisfaction with the "SteWARdS Antibiotic Defence" app was measured with 4 questions on a 5-point Likert scale and the reason(s) for recommending or not recommending the app to others, and suggestions for improvement were obtained via open-ended questions upon quest completion.

Sample Size

Assuming an effect size of 0.334 (95% CI 0.260-0.407) based on other serious games studies [18], we estimated a minimum sample size of 142 per group to detect a significant change in the knowledge score at a power of 80% and an a level of .05. We initially planned to recruit 200 participants per arm to account for a 30% attrition rate. We later increased the recruitment numbers to 240 per arm to account for a 40% attrition rate as we observed higher-than-expected dropout rates.

Study Procedures

Recruitment occurred on weekdays during the polyclinic's opening hours. Staff members of the polyclinics were trained to recruit participants in their respective polyclinics. The data collectors would walk through every level of the polyclinic systematically to recruit participants daily. We also displayed digital recruitment posters around the polyclinics, so that interested visitors could contact the study team. However, only 2 participants contacted the study team via digital posters.

Data collectors approached the patient or the visitor nearest to them who seemed to fall in the eligible age category (18-65 years of age) to explain the study while systematically walking around the polyclinic. Next, the data collectors explained the study to the persons they approached and informed them of the possibility of being randomized into either the "survey group" (control) or the "app and survey" (intervention) group. Informed consent was obtained from interested participants before asking participants to pick a card to determine the group assignment. Participants knew which group they were assigned to as the intervention group participants were asked to download the app after their demographic data was collected. Baseline data collection occurred after randomization because the pre-post knowledge questions were built into the app. We collected participants' KAP responses and demographic data at baseline and reimbursed them a small fee for completing the baseline surveys. All participants received the same amount of reimbursement at the baseline. Reimbursement was tiered to a higher value for the follow-up surveys to motivate participants toward task completion.

Baseline Data Collection

The baseline data collection process differed slightly between the intervention and control group participants. Intervention group participants had to complete the Attitudes-Perception demographic questionnaire on an iPad and install the "SteWARdS Antibiotic Defence" app on their smartphones. The data collector then provided the participant with a unique code to gain access to the app. Upon logging into the app, participants had to complete 11 baseline knowledge questions, followed by a walk-through of the basic features of the app. Controls completed the KAP and demographic questionnaire on an iPad.

Control Group Follow-Up

A study team member (ie, data collector) sends participants a detailed WhatsApp message 6 weeks after their recruitment date to inform them to complete the postintervention KAP questionnaire. The study team member sent a reminder if participants did not complete the questionnaire 2 weeks after the first WhatsApp message.

Intervention Group Follow-Up

Participants had to complete the "SteWARdS Antibiotic Defence" game quest and satisfaction survey and inform the study team via WhatsApp upon intervention completion. Six weeks after participants complete the intervention, a study team member sends a detailed WhatsApp message to inform them to complete the postintervention KAP questionnaire. A study team member reminded participants who did not complete the quest 2 weeks after recruitment and 2 weeks after sending the first KAP WhatsApp message.

Withdrawals and Loss to Follow-Up

Participants who became uncontactable during the study period were classified as loss to follow-up cases while those who indicated that they wished to be withdrawn from the study were considered as withdrawn cases. Withdrawn cases were asked to uninstall the app and were not further contacted.

Randomization

Eligible participants were 1:1 block randomized into the intervention or control group. Each block comprised 2 intervention and 2 control allocations. Participants determined their allocation by picking cards in the block, which was reset after all cards were drawn.

Blinding

Participants were not blinded to the intervention as the data collectors had to explain the "SteWARdS Antibiotic Defence" app to participants assigned to the intervention group. Apart from the data collectors, all other study team members (including the data analyst) were blinded to the group assignment.

Data Analyses

We initially planned to take the intention-to-treat approach in analyzing the study data but eventually took the per-protocol approach, as our app usage data showed that participants who did not complete the game quest had too few interactions with the app to derive any benefit from the intervention. We standardized the total Knowledge and Attitude-Perception scores to 100% to illustrate the pre-post changes in scores. Correct answers are coded as 1 while wrong answers are coded as 0. Similarly, the "correct" attitudes are coded as 1 while neutral and "incorrect" attitudes are coded as 0.

Chi-square tests and Mann-Whitney *U* tests were used for univariate baseline comparisons. Weighted least squares regressions (weighted on preintervention scores) were used to assess the effect of the intervention on knowledge and KAP scores. The best regression model was chosen based on the lowest Akaike Information Criterion. All statistical assumptions were checked to ensure the accuracy of analyses. STATA (version 15; StataCorp LLC) was used for all statistical analyses.

Participant feedback was analyzed qualitatively using thematic analysis. OJT telephoned participants to elaborate on any ambiguous feedback and categorized the feedback into themes. ZH reviewed and discussed the categorization with OJT until a consensus was reached.

Ethical Considerations

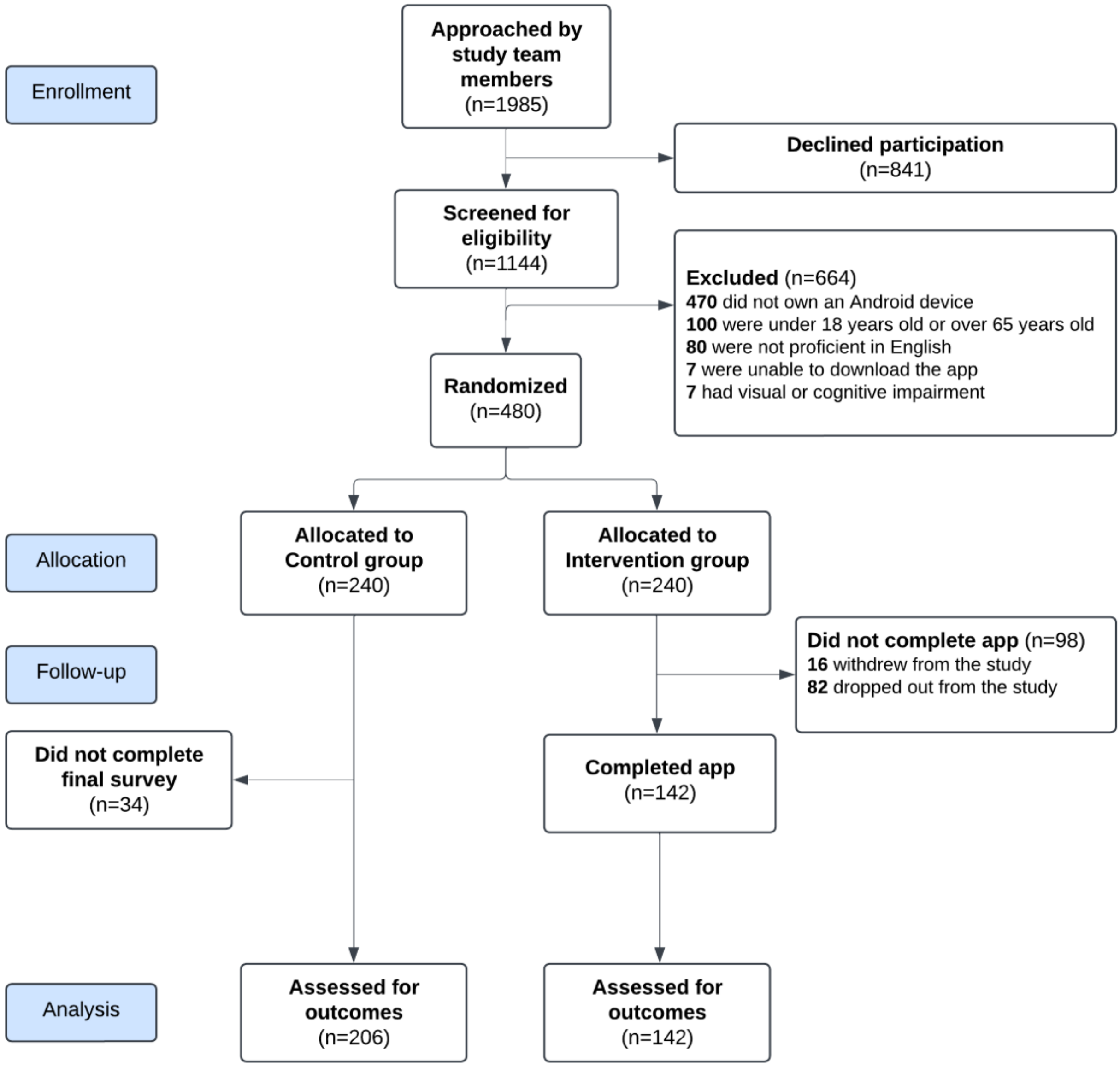
This study was approved by the National Healthcare Group Domain Specific Review Board (2022/00479) in Singapore. Informed consent was obtained from participants after the data collector had fully explained the study to them. Each participant was given a unique research ID linked to their personal details (eg, name and contact details). Only National Healthcare Group Domain Specific Review Board–approved study team members have access to participant data. All identifiable information is stored in password-protected files in the institutions' server. Data collected from the surveys and app contained only the participants' research IDs. Participants were reimbursed SGD 5 (US $3.82) for completing the baseline and demographic survey, SGD 10 (US $7.64) for completing the postintervention survey, and SGD 25 (US $19.10) for completing the app intervention and satisfaction survey (intervention group participants only).

Results

Participant Flow

We approached 1985 patients and visitors of polyclinics for 3 months, and 57.6% (1144/1985) of patients consented to take part in our study. Of these consenting participants, 41.1% (470/1144) of them did not own an Android phone, 8.7% (100/1144) did not meet our age requirements, 7.0% (80/1144) of them were not proficient in the English language, 0.6% (7/1144) of them had difficulty downloading the app, and 0.6% (7/1144) of them had visual or cognitive impairment. We block-randomized 42.0% (480/1144) of participants who were screened for eligibility into 2 study arms. Of these randomly assigned participants, 85.8% (206/240) of them in the control group and 59.1% (142/240) of them in the intervention group completed the study and had their outcomes assessed. About 40.8% (98/240) of intervention group participants did not complete the game quest (16 withdrew; 82 dropped out) in the app (CONSORT [Consolidated Standards of Reporting Trials] flowchart shown in Figure 6).

Figure 6. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of a 2-arm randomized controlled trial conducted on participants recruited from government-funded primary care clinics between January 2023 and March 2023 in Singapore. A total of 1985 people were approached, 1144 were screened for eligibility, and 480 were recruited into the study and randomized equally into the control or intervention group. The outcomes of 142 intervention participants and 206 controls were eventually assessed.



Recruitment

Recruitment occurred between January 2023 and March 2023, and follow-up was completed by early June 2023. We ceased recruitment when we achieved 240 participants per arm.

Baseline Data

Table 1 shows participants' baseline characteristics. The randomization produced comparable groups as there were no significant between-group differences in participants' sociodemographic characteristics at baseline. Participants had a mean age of 36.9 (SD 12.2) years. Our participant profile fits Singapore's population profile in terms of ethnicity (252/348, 72.4% Chinese; 61/348, 17.5% Malay; and 24/348, 6.9% Indian), education level (128/348, 36.8% had tertiary education), and housing type (275/348, 79.0% residing in conventional public housing flats). There were also no significant between-group differences in the mean baseline knowledge score of 8.4/11 (SD 1.6), but the attitude-perception score was significantly higher in controls (mean 9.6, SD 2.6) than in intervention participants (mean 8.8, SD 3.0). In addition, 60.1% (209/348) of participants had no exposure to advertisements on antibiotic use at baseline.

The randomization produced comparable groups as there were no significant between-group differences in participants' sociodemographic characteristics at baseline. Our participant profile fits Singapore's population profile in terms of ethnicity and housing type.

Table 1. Baseline characteristics of the randomized controlled trial participants recruited from government-funded primary care clinics between January 2023 and March 2023 (N=348).

Baseline characteristics of respondents	All participants	Intervention group (n=142)	Control group (n=206)	*P* value
Sociodemographic factors				
Age (years), mean (SD)	36.9 (12.2)	35.9 (12.0)	37.6 (12.3)	.21
Sex, n (%)				.64
Male	181 (52.0)	76 (53.5)	105 (51.0)	
Female	167 (48.0)	66 (46.5)	101 (49.0)	
Race, n (%)				.79
Chinese	252 (72.4)	106 (74.7)	146 (70.9)	
Malay	61 (17.5)	23 (16.2)	38 (18.5)	
Indian	24 (6.9)	8 (5.6)	16 (7.8)	
Other races	11 (3.2)	5 (3.5)	6 (2.9)	
Educational level, n (%)				.24
Tertiary education[a]	128 (36.8)	47 (33.1)	81 (39.3)	
Nontertiary education	220 (63.2)	95 (66.9)	125 (60.7)	
Marital status, n (%)				.36
Married	169 (48.7)	65 (45.2)	101 (49.3)	
Unmarried	178 (51.3)	77 (54.2)	104 (50.7)	
Housing, n (%)				.98
HDB[b] 3-room flat and below	81 (23.3)	33 (23.2)	48 (23.3)	
HDB 4 or 5-room flat	194 (55.8)	80 (56.3)	114 (55.3)	
Above HDB 5-room/private property	73 (21.0)	29 (20.4)	44 (21.4)	
Knowledge, attitude, and perceptions on antibiotics and AMR[c]				
Advertisement on antibiotic use, n (%)				.34
Had seen advertisements	139 (39.9)	61 (43.0)	78 (37.9)	
Had NOT seen advertisements	209 (60.1)	81 (57.0)	128 (62.1)	
Knowledge score, mean (SD) (maximum score: 11)	8.4 (1.6)	8.5 (1.6)	8.3 (1.7)	.16
Attitude-perception score, mean (SD) (maximum score: 14)	9.3 (2.8)	8.8 (3.0)	9.6 (2.6)	.004

[a]Tertiary education: university and above.

[b]HDB flat: flats developed by Singapore's public housing authority.

[c]AMR: antimicrobial resistance.

Change in Knowledge Score

The mean standardized knowledge score improved from 77.6% at baseline to 82.4% postgame and 83.5% 6-10 weeks post intervention among intervention group participants, while controls had a modest improvement of 75.4% from baseline to 78.1% at 6-10 weeks post baseline (Figure S1 in Multimedia Appendix 1).

The multivariable regression analyses (Table 2) showed that intervention group participants had greater improvement in knowledge than controls (between-group differences: 0.58, 95% CI 0.28-0.87) after accounting for age, educational level, and exposure to advertisements on antibiotics and AMR.

Table 2. Multivariable weighted least squares regression analysis of the association between the intervention and knowledge score at 6-10 weeks postintervention.

Model variables	Outcome: 6-10 weeks of postintervention knowledge score		
	Effect size (95% CI)	*P* value	VIF[a]
Randomization group[b] (reference: control)			
Intervention	0.58 (0.28 to 0.87)[c]	<.001	1.04
Knowledge			
Baseline knowledge score	0.24 (0.17 to 0.32)[c]	<.001	1.08
Age group (years; reference: 18-25 years)			
26-35	–0.75 (–1.19 to –0.31)[c]	.001	1.91
36-50	–0.43 (–0.90 to 0.03)	.07	2.49
51-65	–0.45 (–0.99 to 0.08)	.09	2.05
Marital status (reference: not married)			
Married	–0.32 (–0.67 to 0.03)	.07	1.51
Education status (reference: nontertiary education)			
Tertiary education	0.77 (0.45 to 1.09)[c]	<.001	1.13
Advertisement on antibiotics and AMR[d] (reference: have not seen advertisements)			
Have seen advertisements	0.39 (0.09 to 0.69)[c]	.01	1.05

[a]VIF: variance inflation factor.

[b]Intervention group participants had greater improvement in knowledge than controls after accounting for age, educational level, and exposure to advertisements on antibiotics and AMR.

[c]*P*<.05.

[d]AMR: antimicrobial resistance.

Change in Attitude-Perception Score

The mean standardized attitudes-perception score improved from 62.6% at baseline to 77.7% post intervention among intervention group participants, while controls had a modest improvement of 68.9% from baseline to 71.9% post intervention (Figure S1 in Multimedia Appendix 1). The proportion of intervention group participants with the "correct" attitudes and perceptions increased for all 14 statements post intervention (Figure S2 in Multimedia Appendix 1). However, intervention group participants continue to have misconceptions about antibiotic use after completing the game quest. The proportion of participants with correct responses for the statement "I need antibiotics to recover from serious symptoms from the common cold and flu" increased from 29.6% (42/142) pre- to 46.5% (66/142) postintervention, and for "I need antibiotics if I continue to have flu symptoms after two weeks" increased from 22.5% (32/142) pre- to 38.7% (55/142) postintervention.

The multivariable regression analyses (Table 3) showed that intervention group participants had greater attitudes-perception improvement than controls (0.98 [0.44-1.52]) after accounting for marital status and ethnicity.

Table 3. Multivariable weighted least squares regression analysis of the association between the intervention and attitude-perception scores at 6-10 weeks postintervention.

Model variables	Outcome=6-10 weeks of postintervention attitude-perception score		
	Effect size (95% CI)	*P* value	VIF[a]
Randomization group[b] (reference: control)			
Intervention	0.98 (0.44 to 1.52)[c]	<.001	1.30
Attitude-perception			
Baseline attitude-perception score	0.46 (0.39 to 0.53)[c]	<.001	1.26
Education status (reference: nontertiary education)			
Tertiary education	0.56 (–0.13 to 1.26)	.11	1.67
Ethnicity (reference: Chinese)			
Malay	–2.86 (–3.56 to –2.15)[c]	<.001	1.84
Indian	–1.28 (–2.19 to –0.36)[c]	.006	1.71
Others	–0.19 (–2.10 to 1.72)	.85	1.03
Marital status (reference: not married)			
Married	–1.39 (–1.99 to –0.78)[c]	<.001	1.67

[a]VIF: variance inflation factor.

[b]Intervention group participants had better attitudes-perception improvement than controls after accounting for marital status and ethnicity.

[c]*P*<.05.

App Satisfaction

The majority of participants agreed that the "SteWARdS Antibiotic Defence" app improved their awareness on antibiotic use (135/142, 95.1%) and AMR (136/142, 95.8%). About 73.9% (105/142) of the participants agreed that the app is easy to use, 70.4% (100/142) agreed that the app was enjoyable, and 85.2% (121/142) would recommend the app to others (Table 4).

Table 4. Satisfaction in using the "SteWARdS Antibiotic Defence" app (N=142).

App satisfaction	Responses, n (%)
User experience[a]: How far would you agree that the app:	
Improved awareness on antibiotic use	135 (95.1)
Improved awareness on antimicrobial resistance	136 (95.8)
Is easy to use	105 (73.9)
Is enjoyable to use	100 (70.4)
User satisfaction	
Would recommend the app to others	121 (85.2)

[a]User experience was measured on a 5-point Likert scale. We report the proportion of respondents who agreed with the statements on user experience. The majority of participants agreed that the "SteWARdS Antibiotic Defence" app improved their awareness on antibiotic use and AMR. More than 70% agreed that the app is easy to use, enjoyable, and would recommend the app to others.

Participants who would recommend "SteWARdS Antibiotic Defence" found the app user-friendly, fun, enjoyable, and interesting. Some participants also lauded the app's unique learning modality and felt that the intervention addressed an important public health topic, raised their awareness of AMR, and improved their knowledge on appropriate antibiotic use (Table S3 in Multimedia Appendix 1).

> *It provides information about antibiotic resistance in an easy to understand, engaging way.*

> *It is much more engaging compared to reading or [looking] through long videos.*

Participants who would not recommend the app mainly had a poor user experience and were frustrated with technical difficulties in navigating the app. Some technologically astute players found the game boring, repetitive, and unfulfilling.

> *There were technical issues that made completing the game annoying.*

One participant found no value in recommending the app to friends or family member working in health care. Those who found the games difficult to play (mostly older adults) or experienced usability issues (eg, small fonts) also expressed dissatisfaction with the intervention.

Seniors may not be able to understand how to play the games.

Suggestions for App Improvement

Suggestions to improve the quality of “SteWARdS Antibiotic Defence” include the following (Table S4 in Multimedia Appendix 1):

Refining Game Mechanics

Participants pointed out that the minigames should focus on reinforcing concepts the study team intended to teach. For example, 1 minigame leans toward self-care instead of teaching about antibiotics use and AMR. Some participants also suggested improving the game’s instructional clarity, calibrating the in-game economy to optimize the game’s difficulty level, and increasing the number of interactive elements.

The upgrades in the Tower Defence Game is overpowered and makes it too easy.

Let users know how much the coins earned from “endless runner” are worth. This ties into the other game “tower defence” because you need coins to build the towers.

Enhancing the User Interface

Participants suggested having a consistent art style, larger word fonts to cater to older adults, and graphics with a closer resemblance to bacteria and viruses.

Maybe the educational message displayed before each game can be read out.

The 3D games can have better art style to match with the other 2D games and environment.

Having More App Features

Suggested app features include adding music and sound effects within the minigames, allowing players to create personal profiles, and in-app game demonstrations.

More sound effects would be nice.

Bug Fixes

Participants suggested to improve the app’s responsiveness, sensitivity, speed, and fixing the glitches.

Developer to look into the app responsiveness and the graphics.

Discussion

Main Findings

We evaluated the effectiveness of an evidence-based serious game mobile app in improving the KAP of appropriate antibiotic use and AMR among Singapore residents. To our knowledge, this is the first RCT assessing the effectiveness of a serious game mobile app intervention for public education on appropriate antibiotic use and AMR. We also assessed the effectiveness of the app in short-term knowledge retention at 6-10 weeks post intervention, which has often been overlooked. Furthermore, we assessed the app in a real-world context in the community and not in the controlled environment of an experimental setting.

We observed a statistically significant effect size of 0.58 (95% CI 0.28-0.87) in knowledge improvement, which concurs with other serious game studies on chronic disease management among young people [26] and healthy lifestyle promotion [27]. However, a similar Singapore study on dengue prevention observed significant improvements of participants’ knowledge, attitudes, and practices from baseline but no between-group mean differences [28]. The difference in observations is likely attributed to the exposure of relevant information sources among controls in the dengue study, which attenuated the effects of the app intervention. Tertiary-educated adults were positively associated with higher postintervention scores. They may fare better at learning the more difficult AMR concepts as studies have shown that AMR concepts are hard to teach. Participants with no or limited understanding of the etiology of antibiotic resistance would face difficulty connecting AMR with antibiotic misuse [29]. Furthermore, misconceptions about antibiotic use are generally hard to rectify.

We also observed a statistically significant effect size of 0.98 (95% CI 0.44-1.52) in attitude-perception improvement with the “SteWARdS Antibiotic Defence” app. Other studies have found mixed results in the effectiveness of serious games in improving the adoption of healthy behaviors, such as fruit and vegetables intake among adolescents [15]. The effectiveness of serious games in changing attitudes and perceptions are hard to measure as studies are often multifaceted. Therefore, it is essential to have well-designed interventions targeted at the correct population, well-accepted by participants, and with attainable goals for interventions to be successful. While we recruited participants between 18 and 65 years of age, the app was originally developed for young adults as this group of people was found to have the lowest knowledge in antibiotic use and AMR [7].

Our “SteWARdS Antibiotic Defence” app received an overall positive rating, but the ratings should be interpreted with caution as the satisfaction survey was only completed by participants who finished the game quest in the app. It is unclear whether participants dropped out of the intervention due to dissatisfaction with the app or a lack of time to complete the quest. Other studies assessing the effectiveness of serious games in health education have observed high satisfaction among participants [30,31]. However, there is limited evidence on the satisfaction with health education delivered via serious games in the general population and even fewer studies report participant satisfaction. Hence, the feedback and suggestions we gathered from participants are valuable in identifying population-specific shortcomings of the intervention and informing app enhancements that could increase the success of intervention scale-ups.

Despite a modest improvement in KAP scores in the intervention group, improved health literacy may not translate to desired behavioral change [13,32]. Health literacy interventions have led to improved health outcomes in some studies, but the links between health literacy improvement and behavioral change have not been established with higher-quality studies [33].

Effective interventions are also integral to attaining successful behavioral change. One way to improve the learning outcomes, suggested by our participants, is to weave the intended learning concepts into the game mechanics rather than displaying informational text boxes for players to read. Such enhancements would possibly improve user engagement and lead to better learning outcomes [34].

A few limitations exist in our study. First, measurement errors in KAP scores might occur if intervention group participants had requested their friends or family members to help with study completion. Nonetheless, controls could also have had help in completing the postintervention survey. Therefore, any bias which attenuates the intervention's true effect is likely toward the null. Second, there is a ceiling effect for participants with near-perfect baseline scores, which could attenuate the positive effects of the intervention. Third, we randomized participants prior to baseline data collection because the knowledge questions were embedded in the app. Fourth, there may be a learning effect in the intervention group as these participants completed the knowledge questionnaire 3 times. Hence, the positive effects observed in the intervention group could be a combination of the serious game and learning effect from the immediate postgame survey. Including a study group with a traditional form of learning (eg, brochures) may enable us to assess the effect of the serious game app more accurately. Finally, we included only attendees at 2 government-funded primary care clinics (polyclinics). Despite the limitations in recruitment, the sociodemographic distribution of our participants is representative of the Singaporean population. Hence, our findings are generalizable to the rest of the country.

In addition, we noticed several older adults rejecting study participation due to a lack of interest in completing the mobile app quest. Their lack of interest implies an additional effort to change their mindsets, and serious game interventions may not be the best modality for educating older adults on public health issues. Reminders were also integral in driving the app completion rate. Hence, the actual uptake rate may be lower in the population. Future research should assess the effectiveness of interventions that facilitate long-term knowledge retention and long-lasting behavioral change in antibiotic use and AMR [35].

Conclusions

Our educational serious game app improves participants' KAP on appropriate antibiotic use and AMR. Public education apps should be engaging, educational, easy to use, and have an attractive user interface. Interventions should also aim to achieve long-term knowledge retention and long-lasting behavioral change.

Acknowledgments

ZH and OJT had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The study team would like to acknowledge Tan Soon Peng, Khoo Jin Rong, Siti Shazwani Binte Sadimin, Nur Atiqah Binte Surya Akmaja, and Karthiga Natarajan for their assistance in data collection and study administration. We would also like to thank the staff and student interns from Temasek Polytechnic's School of Engineering and School of Informatics & IT for their contribution in maintaining the "SteWARdS Antibiotic Defence" app and database during the study. This work is supported by the National Centre for Infectious Diseases (National Centre for Infectious Diseases Catalyst grant FY2022HZ).

Data Availability

Deidentified data that support the findings of this study are available on request from AC or ZH.

Disclaimer

The research presented in this paper is solely the responsibility of the authors and does not reflect the views of the funder.

Authors' Contributions

ZH codeveloped the app, conceived the manuscript, obtained the catalyst grant, analyzed the data, and drafted the manuscript. OJT administered the study, analyzed the data, and drafted the manuscript. WET provided inputs to the study design and support for running the trial. AC codeveloped the app and provided inputs to the study design of the trial. All authors critically reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional materials to support the findings of the study.
[DOCX File , 1338 KB - games_v12i1e59848_app1.docx]

Multimedia Appendix 2
CONSORT eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 8471 KB - games_v12i1e59848_app2.pdf]

References

1. Thangaraju P, Venkatesan S. WHO Ten threats to global health in 2019: antimicrobial resistance. Cukurova Med J 2019;44(3):1150-1151. [doi: 10.17826/cumj.514157]
2. Dadgostar P. Antimicrobial Resistance: Implications and Costs. Infect Drug Resist 2019;12:3903-3910 [FREE Full text] [doi: 10.2147/IDR.S234610] [Medline: 31908502]
3. Poudel AN, Zhu S, Cooper N, Little P, Tarrant C, Hickman M, et al. The economic burden of antibiotic resistance: A systematic review and meta-analysis. PLoS One 2023;18(5):e0285170 [FREE Full text] [doi: 10.1371/journal.pone.0285170] [Medline: 37155660]
4. Wozniak TM, Barnsbee L, Lee XJ, Pacella RE. Using the best available data to estimate the cost of antimicrobial resistance: a systematic review. Antimicrob Resist Infect Control 2019;8(1):26 [FREE Full text] [doi: 10.1186/s13756-019-0472-z] [Medline: 30733860]
5. O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. Review on Antimicrobial Resistance. 2014. URL: https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf [accessed 2024-07-26]
6. Bell BG, Schellevis F, Stobberingh E, Goossens H, Pringle M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. BMC Infect Dis 2014 Jan 09;14(1):13 [FREE Full text] [doi: 10.1186/1471-2334-14-13] [Medline: 24405683]
7. Guo H, Hildon ZJ, Lye DCB, Straughan PT, Chow A. The associations between poor antibiotic and antimicrobial resistance knowledge and inappropriate antibiotic use in the general population are modified by age. Antibiotics (Basel) 2021;11(1):47 [FREE Full text] [doi: 10.3390/antibiotics11010047] [Medline: 35052924]
8. Lim JM, Duong MC, Cook AR, Hsu LY, Tam CC. Public knowledge, attitudes and practices related to antibiotic use and resistance in Singapore: a cross-sectional population survey. BMJ Open 2021 Sep 13;11(9):e048157 [FREE Full text] [doi: 10.1136/bmjopen-2020-048157] [Medline: 34518255]
9. Chua AQ, Kwa ALH, Tan TY, Legido-Quigley H, Hsu LY. Ten-year narrative review on antimicrobial resistance in Singapore. Singapore Med J 2019 Aug;60(8):387-396 [FREE Full text] [doi: 10.11622/smedj.2019088] [Medline: 31482178]
10. Singh SR, Chua AQ, Tan ST, Tam CC, Hsu LY, Legido-Quigley H. Combating antimicrobial resistance in Singapore: a qualitative study exploring the policy context, challenges, facilitators, and proposed strategies. Antibiotics (Basel) 2019;8(4):201 [FREE Full text] [doi: 10.3390/antibiotics8040201] [Medline: 31671826]
11. Lim JM, Singh SR, Duong MC, Legido-Quigley H, Hsu L, Tam C. Impact of national interventions to promote responsible antibiotic use: a systematic review. J Antimicrob Chemother 2020 Jan 01;75(1):14-29 [FREE Full text] [doi: 10.1093/jac/dkz348] [Medline: 31834401]
12. Hawkins O, Scott AM, Montgomery A, Nicholas B, Mullan J, van Oijen A, et al. Comparing public attitudes, knowledge, beliefs and behaviours towards antibiotics and antimicrobial resistance in Australia, United Kingdom, and Sweden (2010-2021): A systematic review, meta-analysis, and comparative policy analysis. PLoS One 2022;17(1):e0261917 [FREE Full text] [doi: 10.1371/journal.pone.0261917] [Medline: 35030191]
13. Price L, Gozdzielewska L, Young M, Smith F, MacDonald J, McParland J, et al. Effectiveness of interventions to improve the public's antimicrobial resistance awareness and behaviours associated with prudent use of antimicrobials: a systematic review. J Antimicrob Chemother 2018 Jun 01;73(6):1464-1478 [FREE Full text] [doi: 10.1093/jac/dky076] [Medline: 29554263]
14. Sharifzadeh N, Kharrazi H, Nazari E, Tabesh H, Edalati Khodabandeh M, Heidari S, et al. Health Education Serious Games Targeting Health Care Providers, Patients, and Public Health Users: Scoping Review. JMIR Serious Games 2020 Mar 05;8(1):e13459 [FREE Full text] [doi: 10.2196/13459] [Medline: 32134391]
15. Andrew L, Barwood D, Boston J, Masek M, Bloomfield L, Devine A. Serious games for health promotion in adolescents - a systematic scoping review. Educ Inf Technol (Dordr) 2023;28(5):5519-5550 [FREE Full text] [doi: 10.1007/s10639-022-11414-9] [Medline: 36373044]
16. Molnar A, Kostkova P. Learning about hygiene and antibiotic resistance through mobile games: evaluation of learning effectiveness. 2018 Presented at: Proceedings of the International Conference on Digital Health; 2018 April 23-26; Lyon, France p. 95-99. [doi: 10.1145/3194658.3194682]
17. Anderson T, Prue G, McDowell G, Stark P, Brown Wilson C, Graham Wisener L, et al. Co-design and evaluation of a digital serious game to promote public awareness about pancreatic cancer. BMC Public Health 2024 Feb 22;24(1):570 [FREE Full text] [doi: 10.1186/s12889-024-18050-7] [Medline: 38388909]
18. Eley CV, Young VL, Hayes CV, Verlander NQ, McNulty CAM. Young People's Knowledge of Antibiotics and Vaccinations and Increasing This Knowledge Through Gaming: Mixed-Methods Study Using e-Bug. JMIR Serious Games 2019 Feb 01;7(1):e10915 [FREE Full text] [doi: 10.2196/10915] [Medline: 30707096]

19. Farrell D, Kostkova P, Weinberg J, Lazareck L, Weerasinghe D, Lecky DM, et al. Computer games to teach hygiene: an evaluation of the e-Bug junior game. J Antimicrob Chemother 2011 Jun;66(Suppl 5):v39-v44. [doi: 10.1093/jac/dkr122] [Medline: 21680586]
20. Smartphone penetration rate as share of the population in Singapore from 2010 to 2021 and a forecast up to 2028. Statista. 2023 Sep 12. URL: https://www.statista.com/statistics/625441/smartphone-user-penetration-in-singapore/#:~:text=The%20smartphone%20penetration%20share%20reached,use%20and%20engagement%20of%20smartphones [accessed 2024-07-26]
21. Milestones of Singapore’s Smart Nation story. Smart Nation Digital Government Group. 2022. URL: https://www.smartnation.gov.sg/about-smart-nation/our-journey/milestones [accessed 2024-07-26]
22. Lim J. Myth Or Magic: The Singapore Healthcare System. Madison, WI: Select Publishing; 2013.
23. World Health Organization. Antibiotic Resistance: Multi-Country Public Awareness Survey. Geneva, Switzerland: World Health Organization; 2015.
24. Snow V, Mottur-Pilson C, Gonzales R, American College of Physicians-American Society of Internal Medicine. Principles of appropriate antibiotic use for treatment of nonspecific upper respiratory tract infections in adults. Ann Intern Med 2001 Mar 20;134(6):487-489. [doi: 10.7326/0003-4819-134-6-200103200-00014] [Medline: 11255525]
25. Huang Z, Tang WE, Guo H, Natarajan K, Lee TH, Yeo TW, et al. An Evidence-Based Serious Game App for Public Education on Antibiotic Use and Antimicrobial Resistance: Protocol of a Randomized Controlled Trial. JMIR Res Protoc 2023 Mar 28;12:e45833 [FREE Full text] [doi: 10.2196/45833] [Medline: 36976619]
26. Charlier N, Zupancic N, Fieuws S, Denhaerynck K, Zaman B, Moons P. Serious games for improving knowledge and self-management in young people with chronic conditions: a systematic review and meta-analysis. J Am Med Inform Assoc 2016 Jan;23(1):230-239 [FREE Full text] [doi: 10.1093/jamia/ocv100] [Medline: 26186934]
27. DeSmet A, Van Ryckeghem D, Compernolle S, Baranowski T, Thompson D, Crombez G, et al. A meta-analysis of serious digital games for healthy lifestyle promotion. Prev Med 2014 Dec;69:95-107 [FREE Full text] [doi: 10.1016/j.ypmed.2014.08.026] [Medline: 25172024]
28. Tan A, Koh E, Sankari U, Tang J, Goh CK, Tan NC. Effects of a serious game on knowledge, attitude and practice in vector control and dengue prevention among adults in primary care: A randomised controlled trial. Digit Health 2022;8:20552076221129099 [FREE Full text] [doi: 10.1177/20552076221129099] [Medline: 36185390]
29. Beslon G, Batut B, Parsons D, Schneider D, Knibbe C. An alife game to teach evolution of antibiotic resistance. 2013 Presented at: ECAL : The Twelfth European Conference on Artificial Life; 2014 May 28; France p. 43-50. [doi: 10.1162/978-0-262-31709-2-ch007]
30. Zairi I, Ben Dhiab M, Mzoughi K, Ben Mrad I, Kraiem S. Assessing medical student satisfaction and interest with serious game. Tunis Med 2021;99(11):1030-1035 [FREE Full text] [Medline: 35288906]
31. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious Gaming and Gamification Education in Health Professions: Systematic Review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
32. Asharani PV, Lau JH, Roystonn K, Devi F, Peizhi W, Shafie S, et al. Health Literacy and Diabetes Knowledge: A Nationwide Survey in a Multi-Ethnic Population. Int J Environ Res Public Health 2021 Sep 03;18(17):9316 [FREE Full text] [doi: 10.3390/ijerph18179316] [Medline: 34501905]
33. Walters R, Leslie SJ, Polson R, Cusack T, Gorely T. Establishing the efficacy of interventions to improve health literacy and health behaviours: a systematic review. BMC Public Health 2020 Jun 30;20(1):1040 [FREE Full text] [doi: 10.1186/s12889-020-08991-0] [Medline: 32605608]
34. Drummond D, Hadchouel A, Tesnière A. Serious games for health: three steps forwards. Adv Simul (Lond) 2017;2(1):3 [FREE Full text] [doi: 10.1186/s41077-017-0036-3] [Medline: 29450004]
35. Huang F, Blaschke S, Lucas H. Beyond pilotitis: taking digital health interventions to the national level in China and Uganda. Global Health 2017 Jul 31;13(1):49 [FREE Full text] [doi: 10.1186/s12992-017-0275-z] [Medline: 28756767]

Abbreviations

AMR: antimicrobial resistance
CONSORT: Consolidated Standards of Reporting Trials
KAP: knowledge of, attitude toward, and perceptions
NPC: nonplayer character
RCT: randomized controlled trial

Edited by T Leung; submitted 23.04.24; peer-reviewed by P Sterkenburg, G Mboowa; comments to author 26.05.24; revised version received 29.05.24; accepted 12.07.24; published 05.09.24.

Please cite as:
Huang Z, Ow JT, Tang WE, Chow A
An Evidence-Based Serious Game App for Public Education on Antibiotic Use and Resistance: Randomized Controlled Trial
JMIR Serious Games 2024;12:e59848
URL: https://games.jmir.org/2024/1/e59848
doi:10.2196/59848
PMID:39235853

Original Paper

The Role of Relevance in Shaping Perceptions of Sleep Hygiene Games Among University Students: Mixed Methods Study

Zilu Liang[1], PhD; Edward Melcer[2], PhD; Kingkarn Khotchasing[1], BEng; Samantha Chen[3], BSc; Daeun Hwang[3], BSc; Nhung Huyen Hoang[1], MEng

[1]Ubiquitous and Personal Computing Lab, Kyoto University of Advanced Science, Kyoto, Japan

[2]School of Computer Science, Carleton University, Ottawa, ON, Canada

[3]University of California, Santa Cruz, Santa Cruz, CA, United States

Corresponding Author:
Zilu Liang, PhD
Ubiquitous and Personal Computing Lab
Kyoto University of Advanced Science
18 Yamanouchi Gotanda-cho, Ukyo-ku
Kyoto, 615-8577
Japan
Phone: 81 754966510
Email: liang.zilu@kuas.ac.jp

Abstract

Background: Sleep games are an emerging topic in the realm of serious health game research. However, designing features that are both enjoyable and effective at engaging users, particularly university students, to develop healthy sleep habits remains a challenge.

Objective: This study aims to investigate user preferences for 3 sleep game prototypes, that is, Hero's Sleep Journey, Sleep Tamagotchi, and Sleepland, and to explore their popularity and perceived utility in promoting sleep health.

Methods: A mixed methods approach was used in this study. Quantitative and qualitative data were collected through a co-design workshop involving 47 university students. Participants were presented with storyboard cards of game features and were asked to provide an overall rating on each game, as well as ratings for individual features. They were also encouraged to provide free-form comments on the features and suggest improvements. In addition, participants were asked to express their preferences among the 3 games regarding which game they would most like to play and which one they found most useful for promoting sleep health.

Results: Surprisingly, while Hero's Sleep Journey was the most popular choice among participants, Sleep Tamagotchi was perceived as the most beneficial for improving sleep health. Relevance emerged as an overarching theme in the qualitative data analysis, with 3 interconnected dimensions: psychological relevance to users' personal lives, logical relevance to sleep health, and situational relevance to users' circumstantial context. We discussed how the 3 dimensions of relevance address the autonomy and relatedness constructs outlined in the self-determination theory and proposed 3 design recommendations.

Conclusions: Our serious sleep game prototypes demonstrated the potential to engage university students to develop healthy sleep hygiene. Future sleep game designs should aim to create a sense of relevance to users' personal lives, sleep health goals, and situational contexts. Rather than a one-size-fits-all approach, it is essential to develop a wide range of game genres and features to cater to diverse users. Aligning game features with sleep health goals and educating users on the design rationale through sleep knowledge are also important aspects. Furthermore, allowing users to customize their game experience and manage technology boundaries is necessary to nurture a sense of control and autonomy in the process of forming good sleep hygiene.

(JMIR Serious Games 2024;12:e64063) doi:10.2196/64063

KEYWORDS

serious games; sleep hygiene; sleep technologies; co-design; relevance; self-determination theory; digital health; persuasive technology; behavior change

Introduction

Background

Sleep plays an important role in sustaining physical and mental health [1]. Not having enough good sleep is associated with an increased risk of cardiovascular diseases, metabolic disorders, temporary and permanent cognitive impairment, and mental health issues [2,3]. However, attaining adequate sleep is often challenging for many university students, and poor sleep quality is frequently associated with poor sleep hygiene in this population [4-7]. Sleep hygiene encompasses 4 interconnected dimensions: establishing a regular sleep-wake cycle, creating a nighttime routine, optimizing the sleep environment, and cultivating healthy daily habits [8]. Sleep research has demonstrated that behavior changes targeting sleep hygiene positively impact sleep health among clinical populations [8,9], indicating the potential application of behavior change technology for promoting healthy sleep behavior. Despite this potential, the development of such technologies in the domain of sleep health has generally been limited in focus and scope [10].

There have been previous attempts to incorporate playful elements into improving sleep hygiene to make the process more engaging and effective [11]. These efforts can be broadly categorized into tangible approaches and mobile apps. Tangible approaches use physical tools or materials, such as worksheets [12], tactile game boards [13], and aroma tokens [14]. In addition, various alarm clocks have been introduced with designs to awaken users in effective and creative ways. Examples include Ruggie (Windustries Limited), which requires users to step on it; Gun O'Clock (Bandai Namco), which necessitates shooting a target; and Clocky (CLOCKY, LLC) and the Helicopter Alarm Clock (Toyo Trading), which must be chased to deactivate. These approaches challenge users by incorporating active tasks to turn off the alarm.

With respect to mobile apps, many attempts have been made to develop various sleep games. Some focus on relaxation by integrating peaceful and calm elements within the game design (eg, Harmony [Logitech Europe SA], SpinTree 3D [Tabasco Games], and Sheep Sleep [Superpea Ltd]), while others guide users through relaxation techniques such as breathing exercises (eg, Loona [KEYi Technology Co, Ltd] and Mindllama Breathe Sleep better [Llama Luna Apps]). Certain games promote mindfulness regarding bedtime routines [15], screentime before bed [16], and daily habits related to sleep [14] or incentivize the maintenance of healthy bedtime habits (eg, Sleep Town [Seekrtech] and Pokémon Sleep [The Pokémon Company]). Despite this variety, most of these apps function as alarm clocks, offering playful methods to wake users up and regulate their sleep patterns. They typically present specific missions (eg, Early Bird Alarm Clock [ONEYEAR], Sleep as Android: Smart alarm [Urbandroid Team], AlarmClock Xtreme [Agilesoft Resource], ChallengeAlarm [Nicole Ocean], Alarmy [Delight Room Co, Ltd], and TurboAlarm [Francisco Javier Castaño Gómez]) or incorporate minigames, such as simple shooting games or puzzles, requiring players to achieve certain goals to deactivate the alarm (eg, AlarmMon [Malang Studio Co Ltd], Vmons [Vmons App], Wakey [Kanetik], AlarmBuddy [Alarm Buddy], and Unicorn Alarm [NETIGEN Apps]).

While there are numerous games and gamified apps to enhance sleep hygiene, there remains significant room for improvement. First, as noted earlier, existing solutions are primarily designed for mobile apps or custom tangible artifacts. At this point, smartwatches offer unique advantages over other types of technology, such as reduced blue light exposure, timely interventions, and the capacity to track different health indicators [17,18] and sleep metrics [19], with reasonable accuracy [20]. Nevertheless, surprisingly little work has explored the potential of smartwatch games for promoting sleep hygiene.

Second, current sleep games and gamification systems primarily target the nighttime routine and, to a lesser degree, the regularity of sleep-wake cycles and sleep environment optimization. Only 1 study has explored the design space surrounding daily activities, particularly caffeine intake, for sleep hygiene improvement [14]. Targeting daytime activities that promote better sleep represents a vast design space, including strategies, such as receiving early morning sunlight exposure, having meals early, and avoiding late naps or caffeine consumption close to bedtime [21-25]. Furthermore, it is still unclear how incorporating daily activities to improve sleep health will be received by users of sleep hygiene technologies.

Finally, most current services aim to improve sleep quality through gamified apps rather than actual games that provide a fully playful and immersive experience. The potential of games for behavior changes and persuasive technologies is well documented [26] but surprisingly underexplored in the domain of sleep health.

This Study

This work seeks to explore how smartwatch-based sleep health games are perceived by potential users and could in turn be used as effective tools for improving sleep hygiene. We achieve this by tapping into the unexplored dimensions of sleep hygiene games using a participatory design approach, aiming to provide insights and design implications that contribute to the future development of serious sleep hygiene games.

Methods

Conceptualization and Design of Smartwatch Sleep Games

Overview

In prior work, we adopted a co-design approach that gathered a cohort of university students to participate in the brainstorming and design of sleep health technologies targeted at addressing the sleep health needs of this population more accurately [4]. A total of 51 university students participated in 3 co-design workshops. Thematic analysis of the qualitative data generated 9 themes, including health tracking, sleep environment optimization, sleep literacy, community, recommendation, gamification, generative artificial intelligence, materializing sleep with learning, and personalization. For more details on the workshop and resulting initial design insights, refer to the study by Liang et al [4]. Building on these design insights and evidence-based sleep health knowledge, we developed 3 low-fidelity prototypes of smartwatch sleep games that addressed various themes identified as important by the participants of the initial workshop (Figure 1). The game features for each prototype and their designs are described in more detail in the subsequent sections.

Figure 1. Prototype mockups of the 3 smartwatch sleep games.

Hero's Sleep Journey: Power Up Hero With Your Sleep to Fight Monsters

Hero's Sleep Journey (HSJ) focuses on leveraging the powerful motivational effects of role playing and avatars [27-29]. It features a hero avatar as the main visual theme and is centered on adventure and mission completion. Users play as a hero who grows stronger through good sleep hygiene and fights monsters to protect their village. Some of the core game features are summarized in Textbox 1.

Textbox 1. Summary of core game features of Hero's Sleep Journey.

Core game features of Hero's Sleep Journey

- Regain health points through sleep: players restore their hero's vitality by ensuring adequate sleep, with an emphasis on sleep duration and quality.
- Fight monsters through healthy daily habits: players fight monsters that attack the village by performing various healthy daily habits, such as exercising, walking, stretching, and so forth. Successfully defeating monsters rewards players with in-game currency, which can be used to hire additional heroes or change avatars, as well as buy accessories and weapons that can further improve their hero's abilities.
- Complete daily sleep hygiene challenges: players can complete a small number of special challenges each day to gain unique rewards for their hero, such as extra currency, exclusive weapons or accessories not available in the shop, and stat increases that can further improve their hero's abilities. Examples of special challenges include but are not limited to walking a set amount of distance, taking a certain number of steps in an hour, or reaching a specified heart rate.

Sleep Tamagotchi: Raise Virtual Pets With Your Sleep

Sleep Tamagotchi (ST) leverages the charm of a virtual pet to engage players in improving their sleep hygiene [30,31]. The central theme involves a cute digital pet that guides users in developing and maintaining healthy sleeping habits. Some of the core game features are summarized in Textbox 2.

Textbox 2. Summary of core game features of Sleep Tamagotchi.

Core game features of Sleep Tamagotchi

- Complete daily routines with pet: players can complete various daily activities with their pet, such as stretching, walking, and exercising at appropriate times of the day. These activities are proven to benefit night sleep according to sleep science studies [6,7].
- Prompting healthy sleep practices: players will be prompted with notifications and in-game animations of their pet performing desired activities at appropriate times of day or night (eg, walking, stretching, taking a bath, getting ready for bed, and sleeping).
- Customization rewards for healthy sleep practices: players earn rewards and gifts for their pet through sleeping well, maintaining a consistent sleep schedule, and forming good sleeping habits.

Sleepland: Building Village With Your Sleep

Sleepland focuses on using the motivational power of construction games, that is, games where players combine or configure objects to organize and build things [32]. Specifically, it features building construction as its main visual theme, allowing players to create and develop their own land or town. Users are tasked with the challenge of restoring a Japanese village using their *sleep power*, which is accumulated through healthy sleep hygiene activities. Some of the core game features are summarized in Textbox 3.

Textbox 3. Summary of core game features of Sleepland.

Core game features of Sleep Tamagotchi

- Engaging game narrative: A guardian spirit visits players and asks for the player's help in fixing an old countryside village so that they can have a beautiful place to watch over and protect again. The player has daily interactions with this guardian spirit and slowly learns about the village, its people, and its history over time as they progress through the game.
- Healthy sleep hygiene rewards: When players sleep, they can meet goals and reinforce healthy behaviors to gain points. Meeting goals for the first time, such as sleeping earlier or reducing screen time, will allow them to repair the village and expand its boundaries. In addition, maintaining existing healthy sleep behaviors will reward the player with in-game currency, which can be used to build within the village and customize its look.
- Repair the village: The repair stage involves fixing habits that users want to change, such as waking up earlier, sleeping earlier, reducing screen time, or drinking more water for the first time. Meeting these goals allows users to repair places and items in the village.
- Build within the village: Players can build new habits by repeating healthy behavior over time. As players build up these new habits, they receive in-game currency that allows them to build new things within the village to customize its look, layout, and feel.

Co-Design Workshop

Participants

We distributed flyers and posters around the campus of the Kyoto University of Advanced Science (KUAS). A total of 45 participants (self-identified as female participants: n=7, 16%; and self-identified as male participants: n=38, 84%) attended the workshop. The age of the participants ranged between 18 and 35 years. In total, 9 (20%) participants had prior experience using a smartwatch or wristband for self-tracking. All participants were undergraduate or graduate students enrolled in the Faculty of Engineering at KUAS during the time of the workshop.

Workshop Flow

Before beginning the main activities of the workshop, we provided an introduction to the workshop goal as well as background information on serious games, wearable sleep-tracking technologies, and sleep hygiene. Participants were given an icebreaker activity to encourage collaboration and discussion among one another. Following the icebreaker, participants were briefed on the main findings from the initial workshop and given background information about the sleep game. They were then presented with storyboard cards depicting features from all 3 prototypes. In total, 2 examples of these storyboard cards are shown in Figures 2 and 3. Each card included 4 questions on the backside: "How useful is this feature for improving sleep?"; "Would you use this feature?"; "Could anything be improved?"; and "Free comments." The first 2 questions required participants to rate the feature on a Likert scale ranging from 1 to 4, with higher ratings indicating a more positive appraisal of the feature. The last 2 questions are open-ended. Participants formed groups of 4 to 6 people to evaluate each feature of the 3 prototypes and were encouraged to openly share their feedback within their groups. Once all features were evaluated, the cards were collected.

Figure 2. Storyboard of the friendly nudge feature in Sleep Tamagotchi.

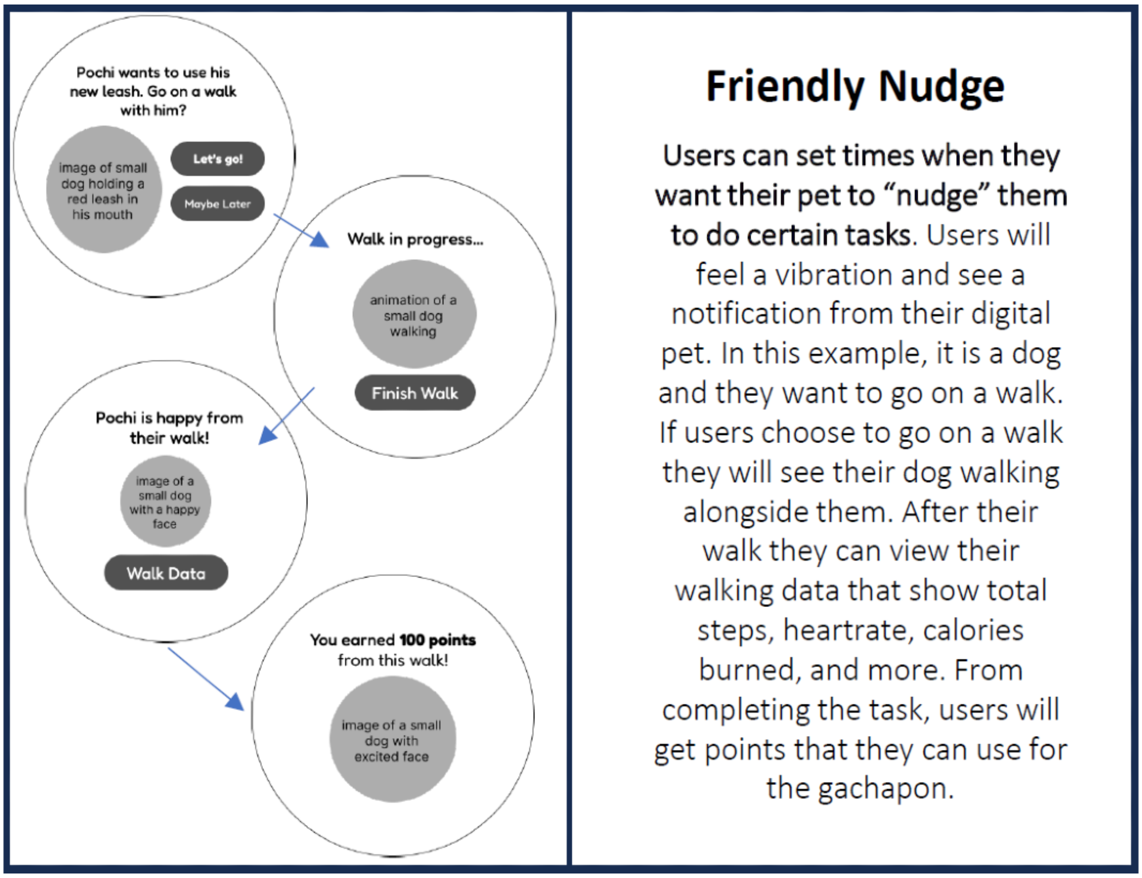
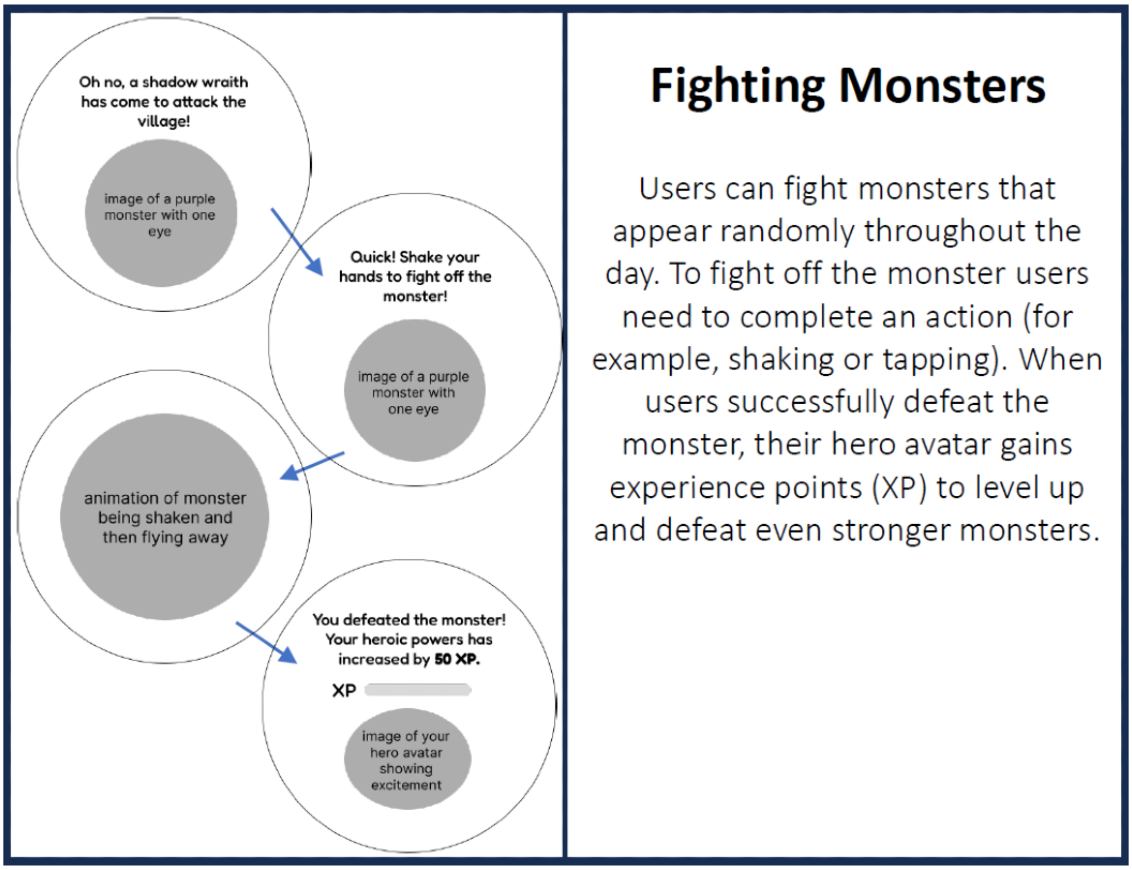


Figure 3. Storyboard of the fighting monster feature in Hero's Sleep Journey.



At the conclusion of the workshop, participants completed the 5-factor player traits questionnaire [33], a validated instrument with 25 items, administered in paper format. This questionnaire assessed participants' preferences for different game elements and playing styles. Using this tool provided a deeper insight into the participants' inclinations toward each sleep game. Alongside the questionnaire, participants were also surveyed about the amount of time they spent playing games per week as well as their preferences among the 3 games regarding which one they found most useful for promoting sleep health and which game they would most like to play.

Data Analysis

This study used a mixed methods approach, resulting in a comprehensive collection of both quantitative and qualitative data from the workshop. Initially, all data were collected on paper and subsequently digitalized into spreadsheets for analysis. For the quantitative data analysis, histograms were created to illustrate participants' ratings of each game's overall appraisal and individual features. The qualitative data, consisting of participants' free-form comments, were analyzed using thematic analysis. In total, 3 authors independently conducted the coding, and discrepancies were resolved through discussion. Instead of using a predefined coding schema, codes were inductive and identified iteratively by repeatedly reviewing the comments. These codes were then grouped into 8 themes, which will be detailed in the *Qualitative Findings* section.

Ethical Considerations

The study was approved by the ethics review board of Kyoto University of Advanced Science (23E04). Participants signed informed consent forms before the start of the workshop and received an Amazon gift card (approximately US $20) upon the completion of the workshop. Data were anonymized to protect the privacy of the participants.

Results

Participants

Understanding player traits is essential for designers to create games that resonate with their target audience. We received 96% (43/45) of the responses to the 5-factor player traits questionnaire. Nearly half (21/43, 49%) of the respondents primarily demonstrate the esthetics trait, highlighting their appreciation for a game's visual and auditory elements, such as graphics, sound, and art style. The second most common trait is the narrative trait (11/43, 26%), revealing their preference for complex storylines within games. In addition, 21% (9/43) demonstrated dominant social traits, reflecting their preference for social interaction during gameplay and their desire to feel connected to other players. As for game-playing hours, most respondents (31/43, 72%) reported spending a modest amount of time playing games, typically between 1 and 10 hours per week, which aligns with the global average of around 8.5 hours per week [34]. In addition, 19% (8/43) spent 11-20 hours per week, while only 7% (3/43) spent more than 21 hours per week on gaming.

Quantitative Findings

We received responses from 96% (43/45) of participants regarding their preferences for the 3 sleep game prototypes. Our results revealed a notable discrepancy between the popularity and perceived usefulness of sleep games. When asked which game prototype they would most like to use, 56% (24/43) of the respondents selected HSJ, 30% (13/43) chose ST, and only 14% (6/43) opted for Sleepland. By contrast, when asked which game prototype they found most helpful for improving sleep, 67% (29/43) identified ST, 23% (10/43) picked HSJ, and only 9% (4/43) selected Sleepland.

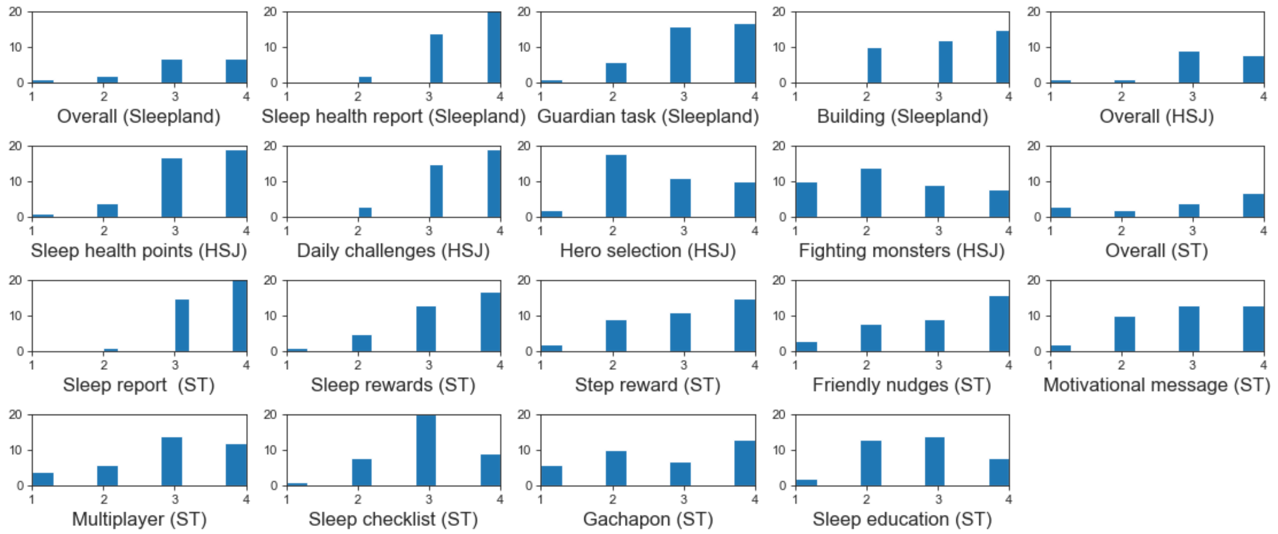
For the perceived utility of individual features, we received ratings and feedback from 91% (41/45) of participants for both Sleepland and HSJ and 84% (38/45) of participants for ST. Table 1 shows the portion of respondents who answered "yes" to the question "Would you use this feature?" The most popular features across 3 sleep games are sleep reports, sleep rewards, and daily challenges, which have a clear and direct relation to changing sleep hygiene behaviors and providing concise feedback on that progress. In contrast, the least popular features are fighting monsters and hero selection in HSJ and sleep education in ST. These features either obfuscated the focus on sleep hygiene or required extraneous amounts of effort to process.

A similar tendency was observed in the histograms of respondents' ratings for each game prototype and individual feature, as illustrated in Figure 4. Consistent with participants' willingness to use the features, the hero selection and fighting monster features of HSJ as well as the sleep checklist and sleep education features of ST received lower ratings from participants. In contrast, the sleep health report and Guardian's task features of Sleepland, the sleep health points and daily challenges features of HSJ, and the sleep report and sleep rewards features of ST received higher ratings, indicating greater appreciation for these features. Meanwhile, features such as Gachapon and motivational nudges of ST exhibited more variance in their scores.

Table 1. The proportion of respondents who answered "yes" to the question "Would you use this feature?"

Game and individual feature	Respondents, n (%)
Sleep Tamagotchi (n=38)	
Sleep report	33 (87)
Sleep rewards	32 (84)
Motivational message	27 (71)
Sleep checklist	27 (71)
Step reward	26 (68)
Friendly nudges	26 (68)
Multiplayer	25 (66)
Gachapon	24 (63)
Sleep education	19 (50)
Hero's Sleep Journey (n=41)	
Sleep health points	37 (90)
Daily challenges	34 (83)
Hero selection	26 (63)
Fighting monsters	18 (44)
Sleepland (n=41)	
Sleep health report	36 (88)
Guardian task	33 (80)
Building	31 (76)

Figure 4. Histogram of participants' ratings to the question "How useful is this feature for improving sleep health?" (1=not useful and 4=very useful). HSJ: Hero's Sleep Journey; ST: Sleep Tamagotchi.



Qualitative Findings

Thematic analysis identified 8 themes, with 2 that arose from participants' suggested improvements for each game or feature.

Game Genre Impacts the Popularity of Sleep Games

Our observations suggest that the genre of the game influences participants' perceptions of its features. Many participants who indicated their preference for HSJ were originally fans of role-playing games (RPGs), drawn to the fantasy environment embodied by the theme and storyline of the HSJ game. For instance, the participants said the following:

> *(The HSJ game) has a clear energy and having to sleep to fight is pretty cool.* [P24]

> *It's a good idea to have a role-playing game. Having a main hero character has always been the 'apple of the eye' for people in movies and games. So having a hero against a villain is a good idea for games.* [P35]

Similarly, participants who favored Sleepland were originally fond of building games, while those who chose ST felt connected to animals and pets. They said the following:

> *I love world-building games where you can create your own little corner of peace!* [P4]

> *I like gathering resources and using them to build something up, and this would be a fun way to get points and work towards a satisfying village.* [P10]

> *I really enjoy games related to pets that we have to take care of or fight for.* [P23]

Different Preferences Between Competitive and Noncompetitive Players

Another primary factor influencing participants' preferences for different sleep game prototypes was their orientation toward competitive or noncompetitive gameplay. The HSJ game distinguished itself as the most interactive game among the 3, featuring level-up mechanics and customization that appealed to competitive players. P34 noted the following:

> *The hero character is very interesting with the upgrade potential. I could make my own customizable character different from my friends.* [P34]

In addition, a few participants found HSJ useful for improving sleep due to its motivating and engaging nature. In total, 4 participants who voted for HSJ believed that the daily challenges would motivate them to "be more mindful of sleep health" and they would "try to sleep well so that the hero can win." P13 mentioned that "progress can be made much faster (in HSJ) compared to the other two (games) due to the element of competition."

Conversely, not all participants identified as competitive players. Some preferred Sleepland and ST precisely for their passive nature. Compared with HSJ, Sleepland was considered less complicated because, as P3 said, "the tasks can be automated, and it takes minimum input from the users." The pet theme of ST was also perceived as more lighthearted compared with the other 2 games, as P5 noted, "I don't think I would be a competitive sleeper, which is the requirement for Hero's Sleep Journey." Unlike HSJ, which excites players, ST causes less stress as, according to P5, "it seems like a fun app in an easygoing manner. Not too competitive. And feels like a very gently encouraging app."

Emotional Impact of Sleep Games

Participant feedback highlights the emotional impact of sleep games. Interacting with cute pets in ST seems to create an affection effect that reinforces positive behavior change by fostering a sense of companionship and mirroring behavior. Participants reported feeling connected to virtual pets, with some expressing willingness to "do anything for a cute virtual pet, even sleeping early" or "try to complete any tasks that a virtual pet gives me." Furthermore, the friendly nudge feature in ST was praised for its companionship effect, potentially enhancing users' motivation. P5 commented as follows:

> *It might be a bit boring to do walks or exercises on your own, so it would be nice to have some encouragement from a virtual pet.* [P5]

Virtual pets may also have a mirroring effect, as stated by P40:

> *When you are taking care of the animals routine, you can have a perceptive of your own routine.* [P40]

A similar effect was also noted in the HSJ game. One participant associated the healing during sleep with the healing process of the hero character, thus considering an RPG the most suited for sleep health intervention.

Nevertheless, we also observed potential negative emotions that sleep games may elicit, particularly in relation to reward features. While rewards can serve as external motivation for some participants, others expressed concerns that earning health points or completing tasks in HSJ may induce unnecessary stress. In addition, rewards based on sleep quality may inadvertently trigger rumination, since factors like the amount of deep sleep are often beyond an individual's control.

Alignment With Sleep Health Enhances Perceived Utility of Game Features

Surprisingly, the most popular game is not necessarily perceived as the most useful. While HSJ seemed to be the most popular, ST was perceived as the most useful. The primary reason for ST's perceived utility lies in its features, which are more focused on sleep health and less distractive from the sleep goal.

More than one-third (13/38, 34%) of the participants considered the ST features highly pertinent to sleep health, stating that ST "has the most sleep and helpful driven features out of the three" (P12). They acknowledged that "there are lots of features that could help me improve my sleep health even if the features or tasks might not be interesting" (P22). Sleep report and sleep education were 2 features that participants highly appraised for their potential positive impact on users' awareness of sleep health. They considered the sleep education contents useful and informative, and the sleep reports would allow users to "look back in the long run" to "analyze sleep patterns" so that "users can be in touch with this for a longer period without getting bored" (P43). These features help users stay focused on improving sleep health without overengaging with features that are not critical for sleep health. As P30 stated, "the other games will make you get distracted and prioritize your game more than your sleep."

Similarly, due to the perceived direct link to sleep, the sleep reward features in all 3 prototypes, which manifested differently as a building feature in Sleepland, a sleep health points feature in HSJ, and a sleep reward feature in ST, were highly praised by participants. In contrast, the step reward feature in ST, while also a form of reward, garnered less interest. Given the unclear relationship between step count and sleep quality improvement, participants were less inclined to engage with this feature.

Game Features Distracting From Sleep Health

We observed that some interactive features, despite being fun and engaging, were considered distracting from the goal of improving sleep health. Participants noted that the diversity of tasks in HSJ is likely to keep users engaged. P1 stated the following:

> *The range of possible tasks is the most diverse in Hero's sleep journey, so it is most likely to keep me engaged the longest.* [P1]

However, participants also acknowledged that these interactive features are not necessarily tied to improvement in sleep quality, just as one participant mentioned:

> *It looks the most fun to me. It also feels like the most interactive one. Although I don't feel like it would help improve my sleep like the other two.* [P12]

Similar complaints were noted about Sleepland by P12, "it (Sleepland) feels more like Clash of Clans than a sleep improving app."

Features Influencing Perceived Autonomy

Participants' preferences for game features were influenced by their perception of autonomy while engaging with those features. They highly valued the flexibility of tasks in Sleepland and ST, emphasizing the importance of being able to "choose their tasks" and "do them at any time" to avoid interference with their work or class schedule. In contrast, the hero selection feature in HSJ received criticism for requiring frequent user choices. Excessive customization imposed a sense of forced action that compromises autonomy. Similarly, the sleep checklist feature imposes a sense of forced action. Participants stated that they would be "too tired at the end of the day to check the list" (P12) and "people have different strategies in winding down for sleep" (P22).

The daily challenge features, including the friendly nudge feature in ST, a daily challenge feature in HSJ, and the Guardian's task feature in Sleepland, may also compromise autonomy. For example, participants found that the 10,000-step goal was too laborious and challenging to achieve, preferring apps that do not dictate their actions. P45 preferred to do things by herself, and she did not "need the app to tell me what to do." In addition, concerns were raised regarding the potential frequent notifications resulting from the daily challenge features. Participants expressed worries that these notifications could be disruptive and cause annoyance rather than motivation.

Suggested Improvements

Overview

Participants offered valuable feedback for improving the sleep games, as summarized in Textbox 4. The suggestions largely fall into 2 categories: enhancing the game experience and aligning the game with sleep health goals and outcomes. Suggestions for improvement included customization options, diverse reward structures, and integration of community features. In addition, participants emphasized the importance of aligning game features with sleep health goals, such as discouraging oversleeping and providing personalized sleep recommendations.

Textbox 4. Summary of participants' suggestions for improvement.

Suggestions for enhancing the game experience

- Avoid excessive notifications
- customization (game characters, daily challenge list, and goals)
- personalized reward (bonus rewards for streaks and higher rewards for people with a worse sleep baseline)
- community and social features
- support multisensory interaction (voice input and vibration)

Suggestions for better alignment with sleep health

- Provide personalized sleep recommendations
- align daily challenges and rewards with the day-night cycle
- discourage oversleeping

Suggestions for Enhancing Game Experience

Participants highlighted the need to maintain personal space and avoid excessive notifications, which can become intrusive and bothersome. Allowing users to customize their notification preferences, including whether and when to receive notifications, was considered essential. Customization in other areas of the game was also valued. Along the same line, daily challenges should be tailored to match a user's capability and daily routine. As P1 stated, "a nice middle ground must be found regarding the difficulty of the daily challenges so as to not demotivate the player." Diversifying game resources, such as introducing various hero and pet types and various forms of in-game currency such as gold and mana, was suggested to enhance the gaming experience. Participants also proposed the idea of assigning different heroes with unique skills that tackle a set of "daily challenges tailored to specific aspects of sleep hygiene" (P27). For example, 1 hero could "focus on exercise for better sleep," while another could "focus on sleep preparation" (P27).

Participants highly valued the sleep reward features in all 3 games but suggested that rewards should not follow a linear structure. They proposed awarding bonus points for maintaining streaks and offering higher rewards to those with a low sleep baseline as strategies to encourage consistent, healthy habits and recognize users' progress. In addition, some participants expressed interest in incorporating penalties for not adhering to good sleep schedules or neglecting daily challenges. They also appreciated the concept of progressive difficulty in the game, with the challenge levels increasing as users advanced. This approach may help maintain engagement and interest while fostering a sense of accomplishment over time.

They also advocated for more community and social features in the games, such as fostering collaboration, competition, team play, and enabling users to share or trade items with other players. For instance, 1 participant suggested allowing users to "visit and interact with other users' villages" in the Sleepland game.

Furthermore, participants valued the use of multisensory interaction. They proposed the integration of pet sounds, ringtones, and vibration patterns as effective ways to boost user engagement and immersion in the ST game.

Suggestions for Better Alignment With Sleep Health

Participants also offered a variety of suggestions to align game features with sleep health, that is, using more endogenous designs [35,36]. These proposals aim to integrate gaming experiences with healthy sleep habits and encourage users to develop better sleep practices grounded in scientific evidence in sleep research.

While recognizing the value of checking sleep statistics, participants emphasized the importance of providing personalized recommendations based on users' current sleep patterns and predicted future sleep quality. Several participants emphasized the importance of aligning daily challenges and rewards with the day-night cycle. For the Sleepland game, P35 suggested implementing a timeline for tasks, such as "sleeping only at night to gain points and exercise in the morning to earn points." P3 recommended that "sleeping during specific hours of the day should allow the hero to heal faster." To minimize distractions from the app during sleep hours, P10 proposed that "if the village is viewed during designated sleep time" in the Sleepland game, "some repairs should be undone."

Oversleeping appears to be a common issue among university students. Participants underscored the importance of discouraging oversleeping, even suggesting penalties for doing so. As P35 noted, "sleeping too much (more than 7-8 hours) should result in point deduction."

By contrast, a few participants expressed concerns about rewarding players based on sleep quality, arguing that rewards should instead focus on sleep habits and discipline, such as maintaining a consistent sleep schedule. They noted that sleep quality can be beyond a person's control, whereas habits and discipline are more manageable.

Discussion

Principal Findings

All 3 sleep games incorporate various persuasive strategies, including self-monitoring, reminders, and rewards, addressing the themes that emerged in our initial co-design workshop [4]. However, not all games or their individual features were perceived as equally useful or engaging. Our findings underscore the pivotal role of the relevance of game features in participants' perceptions of the utility of sleep games. Relevance can be broadly defined as something "important to the matter at hand." The search for relevance is a fundamental aspect of human cognition, and things are relevant to an individual when they have a positive cognitive effect [37]. The relevance of game experience has been recently explored in [38]. In the field of serious games, the importance of relevance has been studied in learning games [39,40]. Previous research has found that relevance influences motivation, adherence, and initial technology adoption. While the importance of relevance has been somewhat recognized in health games [41], no study has thoroughly explored this topic.

Through our analysis, 3 interconnected aspects of relevance emerged: relevance to users' personal experience, relevance to sleep health, and relevance to users' situational context. We mapped these 3 dimensions of relevance in sleep games to different types of relevance identified in the literature: psychological relevance, logical relevance, and situational relevance [42]. Through the lens of self-determination theory [43,44], we discuss how these dimensions of sleep game relevance can fulfill users' psychological needs, particularly autonomy and relatedness, thereby promoting healthy sleep behavior change.

Relevance of Sleep Game to Players' Personal Experience and Preferences

Overview

Participants' inclination to engage with sleep games is heavily influenced by their prior game experience and personal preferences, highlighting the importance of creating psychological relevance in the design of serious sleep games. Our findings indicate that participants preferred sleep games that aligned with their favorite game genre. For example, those fond of RPGs favored HSJ, while others attracted to pet-related themes were more inclined toward ST. Participants' preferences for sleep games were also shaped by their orientation toward competitive or noncompetitive gameplay: competitive players preferred HSJ, and noncompetitive players leaned toward ST and Sleepland.

Several participants emphasized their sense of connection to the game characters as a main reason for their willingness to engage with the game, further highlighting the importance of sleep hygiene games to create psychological relevance. In HSJ, for example, the hero healing feature serves as a metaphor for the restorative effects of sleep, framing the player's actions as instrumental in restoring the hero's vitality. Similarly, users care for virtual pets in the ST game, prompting them to reflect on their own schedules as they maintain a regular schedule for the virtual pet. This creates a digital twin effect [4], aligning with the strategy of similarity used in other health-interventional games [26]. The sense of psychological relevance reduces psychological barriers and enhances the sense of autonomy for behavior change. In contrast, the guardian's task feature in Sleepland, which involves repairing damaged villages, did not seem to foster the same association with the healing benefits of sleep for humans, possibly due to a lack of similarity and relevance between repairing a lifeless object and human recovery. The importance of establishing an emotional connection in serious games has been previously highlighted in [45-47]. Our findings contribute further evidence within the context of sleep health. Game design elements that foster psychological relevance to players' prior game experiences and personal lives, particularly those that elicit feelings of autonomy and relatedness, may be crucial to the initial adoption of sleep game technologies.

Design Recommendation 1: Diversify Game Genres and Create Relatable Characters

Sleep games need not adhere to a one-size-fits-all design. Instead, offering a variety of game genres allows users to choose those that resonate most with their preferences and personal experiences, thereby transforming the external demand for healthy sleep behavior change into a more self-determined, personally endorsed, and volitionally initiated process. Furthermore, creating game characters that users feel connected to will enhance perceptions of self-relevance and promote a sense of relatedness [48-50]. This, in turn, boosts intrinsic motivation to engage with the game and healthy behavior change.

Relevance of Sleep Game Features to Sleep Health

Overview

The second dimension of relevance for sleep games revolves around whether the game features logically contribute to improving sleep. As a form of serious game, sleep games use game design elements to influence players' behavior toward better sleep health. Aligning game design with the serious context poses a well-acknowledged challenge in serious game research [51,52], and sleep games are no exception.

Our findings indicate that perceived relevance to sleep health significantly influences the utility value of game features, which in turn leads players to feel more competent in the act of improving sleep health. Features that directly focus on sleep, such as sleep education and sleep rewards in ST, were highly regarded for their relevance and potential to improve sleep. This relevance was considered a major reason why ST could be more effective than the other 2 games. Conversely, features, such as fighting monsters in HSJ, while enjoyable, were not directly linked to sleep health and were consequently rated as the least popular among all HSJ features. Furthermore, several participants expressed concern that features irrelevant to sleep health might distract users from their sleep goals, potentially harming rather than improving sleep. This concern reflects the tension between the health care context and the entertainment nature of games highlighted in previous studies [53]. Addressing this tension requires careful alignment between the serious context and the game features following established design frameworks [41,53]. Indeed, participants made pertinent suggestions on how to better align the game design with sleep health goals, including aligning daily challenges and rewards with the day-night cycle and discouraging oversleeping.

Interestingly, we observed that the perceived utility of sleep game features does not always match their actual utility. Some features, despite having a direct positive impact on nighttime sleep supported by scientific evidence, were not perceived as relevant by many participants. Features such as the daily challenges in HSJ and friendly nudges in ST encourage users to engage in daytime activities conducive to better sleep. These activities represent meaningful behaviors for improving sleep and empower users to achieve their sleep goals. However, not all participants immediately recognized the relevance of these activities and were hesitant to engage with them. For instance, while some participants understood the connection between daytime physical activity and nighttime sleep, others did not. While goal mechanisms have been proven effective in serious games, particularly in educational settings [54], our study shows that it is equally, if not more important, to communicate the rationale behind these goals to users, particularly regarding their logical relevance to improving sleep.

We also found that perceived relevance alone is not sufficient for users to engage with a feature. Despite its perceived relevance, the sleep education feature in ST was the least popular. Instead of presenting educational content in long text, the sleep education feature could be gamified and presented in smaller chunks through gameplay, as exemplified in [55].

Design Recommendation 2: Align Game Features With Sleep Health and Help Users Understand the Rationale Behind Design

Educational content of sleep knowledge can be embedded within the game's narrative and tasks; for example, introducing a character that provides sleep tips as part of the storyline or using interactive elements, such as quizzes, minigames, and challenges that reward users with in-game benefits for learning about sleep health [55].

Relevance of Sleep Game Features to Players' Situational Context

Overview

The third dimension of relevance for sleep games involves the situational or circumstantial context of users, akin to the notion of situational relevance discussed in prior research [56]. This notion of "life context" has also become an important issue identified in recent behavior change literature [57]. A key finding of this study is that frequent game interactions during the daytime are not always welcomed, as participants' willingness to engage with those features varied across the day. While sleep games do not necessarily need to be played at night [14], it is important to maximize the situational relevance of daytime interactions to avoid compromising the sense of autonomy.

All 3 games implemented features that address daytime activities. There is a large body of evidence showing that numerous daytime activities can influence nighttime sleep [4,5], making daytime interactions a natural design target for sleep games. However, many participants expressed concerns about being disrupted by notifications irrelevant to the activities at hand, especially during classes or other busy times. Consequently, features likely to generate excessive notifications, such as friendly nudges and sleep education in ST, were rated as the least popular features among the workshop participants. Several participants mentioned that they do not need the sleep game app to dictate their actions, indicating a compromised sense of autonomy due to potentially excessive notifications.

One way to address the potential interruptions caused by excessive notifications is through boundary management [58,59], which ensures that notifications are pushed at appropriate timing and are relevant to the users' situational context. This can be achieved through direct approaches, such as user customization, or indirect approaches, such as leveraging contextual information and users' behavior patterns to determine optimal notification timing [60]. For example, many participants were reluctant to engage in sleep games immediately upon waking, as they preferred to prioritize their daily routines over gaming. This observation aligns with previous research indicating that people's acceptance of notifications varies throughout the day [61] and that they are less likely to attend to notifications in the morning compared with other times of the day [62].

Ensuring the relevance of sleep game features to users' situational context addresses their psychological needs for autonomy. Besides customizing notifications, participants expressed a desire for autonomy in completing daily challenges to gain a sense of competence. Users prefer daily challenges to be achievable rather than feeling frustrated about not being able to accomplish them. For example, instead of setting a default goal of 10,000 steps, step goals should be customizable because users may have busy schedules, prefer biking over walking, or prefer indoor exercise over outdoor walking on some days. On the basis of these findings, we propose the third design recommendation mentioned subsequently.

Design Recommendation 3: Support Boundary Management Through Manual and Automatic Customization

Allow users to choose when and how often they receive notifications. Make it possible for users to opt to receive notifications only in the morning, evening, or at specific times that suit their daily routines. Contextual information from other apps, such as calendar, GPS, and sleep profile, can be used to customize notification settings automatically, avoiding sending notifications during busy periods, meetings, or events. Providing a variety of choices and allowing users to configure a daily challenge list according to their daily schedule is another way to enhance the situational relevance of sleep game features.

Limitations

This study has several limitations. First, our participants predominantly consist of male international students enrolled in a Japanese university. As a result, the findings may not be readily applicable to other demographic groups. Second, all participants had prior experience playing video games, which could potentially limit the generalizability of our results to individuals who do not engage in gaming. Third, due to the small sample size, we were not able to quantify the associations between player traits and their preferences for sleep game features. We will expand the participant pool to include a more diverse demographic and investigate the preferences of individuals with limited or no prior gaming experience.

Conclusions

This study examined university students' preferences for 3 sleep game prototypes and their perceived utility in promoting sleep health. Results showed a notable divide between the popularity and perceived utility of the games. HSJ was favored for enjoyment, while ST was seen as more effective for improving sleep health. This highlights the potential of serious games to enhance sleep hygiene, especially when game features are perceived as relevant. Key factors influencing perceived utility include prior gameplay preferences, perceived relevance to sleep health, and avoiding excessive interactions. Our findings suggest focusing on psychological, logical, and situational relevance in game design. We recommend diversifying game genres, aligning features with sleep health, and managing technology boundaries effectively.

Acknowledgments

This study was supported by JSPS Grants-in-Aid for Early-Career Scientists (21K17670) and Kyoto University of Advanced Science (KUAS) Overseas Collaboration Grant. The authors would like to extend their gratitude to the graduate and undergraduate students at the University of California, Santa Cruz (UCSC) who assisted in prototyping the 3 sleep games, especially Bianca Villamor, Jilly Li, Mandy Cai, Bridget Ho, and Hebah Haque. The authors also thank KUAS graduate students Cleber Franca Carvalho, Kevin Kristofer Kosasih, and Prasan Priya Yapa for their help with the logistics of the co-design workshop. Finally, the author would also like to thank the workshop participants.

Conflicts of Interest

None declared.

References

1. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 01;37(1):9-17 [FREE Full text] [doi: 10.5665/sleep.3298] [Medline: 24470692]
2. Russell K, Allan S, Beattie L, Bohan J, MacMahon K, Rasmussen S. Sleep problem, suicide and self-harm in university students: a systematic review. Sleep Med Rev 2019 Apr;44:58-69 [FREE Full text] [doi: 10.1016/j.smrv.2018.12.008] [Medline: 30721844]

3. Luyster FS, Strollo PJJ, Zee PC, Walsh JK, Boards of Directors of the American Academy of Sleep Medicine and the Sleep Research Society. Sleep: a health imperative. Sleep 2012 Jun 01;35(6):727-734 [FREE Full text] [doi: 10.5665/sleep.1846] [Medline: 22654183]
4. Liang Z, Melcer E, Khotchasing K, Hoang NH. Co-design personal sleep health technology for and with university students. Front Digit Health 2024 Apr 9;6:1371808 [FREE Full text] [doi: 10.3389/fdgth.2024.1371808] [Medline: 38655450]
5. Minghelli B. Sleep disorders in higher education students: modifiable and non-modifiable risk factors. North Clin Istanb 2022;9(3):215-222 [FREE Full text] [doi: 10.14744/nci.2021.44520] [Medline: 36199865]
6. Brown FC, Buboltz WCJ, Soper B. Relationship of sleep hygiene awareness, sleep hygiene practices, and sleep quality in university students. Behav Med 2002 Jan;28(1):33-38. [doi: 10.1080/08964280209596396] [Medline: 12244643]
7. Brick CA, Seely DL, Palermo TM. Association between sleep hygiene and sleep quality in medical students. Behav Sleep Med 2010 Mar 26;8(2):113-121 [FREE Full text] [doi: 10.1080/15402001003622925] [Medline: 20352547]
8. Irish LA, Kline CE, Gunn HE, Buysse DJ, Hall MH. The role of sleep hygiene in promoting public health: a review of empirical evidence. Sleep Med Rev 2015 Aug;22:23-36 [FREE Full text] [doi: 10.1016/j.smrv.2014.10.001] [Medline: 25454674]
9. Maniraj P, Jegadeesan P, Ananthakrishnan S. Sleep hygiene a neglected paediatric identity - a cross-sectional study. J Family Med Prim Care 2023 Dec;12(12):3075-3078 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_2160_22] [Medline: 38361902]
10. Orji R, Moffatt K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. Health Informatics J 2018 Mar 31;24(1):66-91 [FREE Full text] [doi: 10.1177/1460458216650979] [Medline: 27245673]
11. Liang Z, Melcer E, Hoang NH, Khotchasing K. Sleep hygiene games and gamification: where are we and where are we heading? Research Gate Preprint posted online on June 10, 2024 [FREE Full text] [doi: 10.13140/RG.2.2.32935.56481]
12. de Almondes KM, Leonardo ME. Study protocol of sleep education tool for children: serious game "perfect bedroom: learn to sleep well". Front Psychol 2018 Jun 26;9:1016 [FREE Full text] [doi: 10.3389/fpsyg.2018.01016] [Medline: 29997540]
13. Pusateri J, Leng J, Timczyk J, Chen X, Wang Q, Shah K, et al. Toward a design theory of sleepy games. In: Proceedings of the Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2019 Presented at: CHI PLAY '19; October 22-25, 2019; Barcelona, Spain. [doi: 10.1145/3341215.3356268]
14. Pusateri J, Leng J, Wang Q, Chen X, Hammer J. Designing games for healthy sleep. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3313831.3376413]
15. Shin JY, Li T, Peng W, Lee HR. Bedtime pals: a deployment study of sleep management technology for families with young children. In: Proceedings of the 2023 ACM Designing Interactive Systems Conference. 2023 Presented at: DIS '23; July 10-14, 2023; Pittsburgh, PA. [doi: 10.1145/3563657.3596068]
16. Cai J, Chen B, Wang C, Jia J. Wander: a breath-control audio game to support sound sleep. In: Proceedings of the Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play. 2021 Presented at: CHI PLAY '21; October 18-21, 2021; Virtual Event, Austria. [doi: 10.1145/3450337.3483461]
17. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Hosseinzadeh S, Sajadi Tabar SS. Smartwatches in healthcare medicine: assistance and monitoring; a scoping review. BMC Med Inform Decis Mak 2023 Nov 03;23(1):248 [FREE Full text] [doi: 10.1186/s12911-023-02350-w] [Medline: 37924029]
18. Chang L, Lu J, Wang J, Chen X, Fang D, Tang Z, et al. SleepGuard: capturing rich sleep information using smartwatch sensing data. Proc ACM Interact Mob Wear Ubiquitous Technol 2018 Sep 18;2(3):1-34. [doi: 10.1145/3264908]
19. Liang Z, Ploderer B, Liu W, Nagata Y, Bailey J, Kulik L, et al. SleepExplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors. Pers Ubiquit Comput 2016 Sep 16;20(6):985-1000. [doi: 10.1007/S00779-016-0960-6]
20. Liang Z, Chapa-Martell MA. Accuracy of Fitbit wristbands in measuring sleep stage transitions and the effect of user-specific factors. JMIR Mhealth Uhealth 2019 Jun 06;7(6):e13384 [FREE Full text] [doi: 10.2196/13384] [Medline: 31172956]
21. Brown TM, Brainard GC, Cajochen C, Czeisler CA, Hanifin JP, Lockley SW, et al. Recommendations for daytime, evening, and nighttime indoor light exposure to best support physiology, sleep, and wakefulness in healthy adults. PLoS Biol 2022 Mar 17;20(3):e3001571 [FREE Full text] [doi: 10.1371/journal.pbio.3001571] [Medline: 35298459]
22. BaHammam AS, Pirzada A. Timing matters: the interplay between early mealtime, circadian rhythms, gene expression, circadian hormones, and metabolism-a narrative review. Clocks Sleep 2023 Sep 06;5(3):507-535 [FREE Full text] [doi: 10.3390/clockssleep5030034] [Medline: 37754352]
23. Ye L, Hutton Johnson S, Keane K, Manasia M, Gregas M. Napping in college students and its relationship with nighttime sleep. J Am Coll Health 2015 Jan 14;63(2):88-97. [doi: 10.1080/07448481.2014.983926] [Medline: 25397662]
24. Mograss M, Abi-Jaoude J, Frimpong E, Chalati D, Moretto U, Tarelli L, et al. The effects of napping on night-time sleep in healthy young adults. J Sleep Res 2022 Oct 06;31(5):e13578. [doi: 10.1111/jsr.13578] [Medline: 35253300]
25. Gardiner C, Weakley J, Burke LM, Roach GD, Sargent C, Maniar N, et al. The effect of caffeine on subsequent sleep: a systematic review and meta-analysis. Sleep Med Rev 2023 Jun;69:101764 [FREE Full text] [doi: 10.1016/j.smrv.2023.101764] [Medline: 36870101]

26. Ndulue C, Orji R. Games for change—a comparative systematic review of persuasive strategies in games for behavior change. IEEE Trans Games 2023 Jun;15(2):121-133. [doi: 10.1109/tg.2022.3159090]
27. M Tuah N, Wanick V, Ranchhod A, B Wills G. Exploring avatar roles for motivational effects in gameful environments. EAI Endorsed Trans Creative Technol 2017 Sep;4(10):153055 [FREE Full text] [doi: 10.4108/eai.4-9-2017.153055]
28. Grasse KM, Melcer EF, Kreminski M, Junius N, Ryan J, Wardrip-Fruin N. Academical: a choice-based interactive storytelling game for enhancing moral reasoning, knowledge, and attitudes in responsible conduct of research. In: Bostan B, editor. Games and Narrative: Theory and Practice. Cham, Switzerland: Springer; 2022.
29. Dickey MD. Game design and learning: a conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. Educ Technol Res Dev 2006 Sep 15;55(3):253-273. [doi: 10.1007/S11423-006-9004-7]
30. Fujinami K. Facilitating unmotivated tasks based on affection for virtual pet. In: Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). 2019 Presented at: PERCOMW 2019; March 11-15, 2019; Kyoto, Japan. [doi: 10.1109/percomw.2019.8730777]
31. Lin C, Faas T, Dombrowski L, Brady E. Beyond cute: exploring user types and design opportunities of virtual reality pet games. In: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. 2017 Presented at: VRST '17; November 8-10, 2017; Gothenburg, Sweden. [doi: 10.1145/3139131.3139132]
32. Bostan B. Player motivations: a psychological perspective. Comput Entertain 2009 Jun 24;7(2):1-26. [doi: 10.1145/1541895.1541902]
33. Tondello GF, Arrambide K, Ribeiro G, Cen AJ, Nacke LE. “I don’t fit into a single type”: a trait model and scale of game playing preferences. In: Proceedings of the IFIP Conference on Human-Computer Interaction. 2019 Presented at: INTERACT 2019; September 2-6, 2019; Paphos, Cyprus. [doi: 10.1007/978-3-030-29384-0_23]
34. Combs V. 8 hours and 27 minutes. That’s how long the average gamer plays each week. TechRepublic. 2021 Mar 10. URL: https://www.techrepublic.com/article/8-hours-and-27-minutes-thats-how-long-the-average-gamer-plays-each-week/ [accessed 2024-05-15]
35. Mestadi W, Nafil K, Touahni R, Messoussi R. An assessment of serious games technology: toward an architecture for serious games design. Int J Comput Games Technol 2018 Aug 01;2018:1-16. [doi: 10.1155/2018/9834565]
36. Athavale S, Dalvi G. Endogen: framework for designing endogenous educational games. In: Proceedings of DiGRA 2020 Conference: Play Everywhere. 2020 Presented at: DiGRA 2020; June 2-6, 2020; Tampere, Finland.
37. Wilson D, Sperber D. Relevance theory. In: Horn LR, Ward G, editors. The Handbook of Pragmatics. Hoboken, NJ: Blackwell Publishing Ltd; 2006.
38. Sharma R, Melcer E, Kao D. Exploring relevance, meaningfulness, and perceived learning in entertainment games. In: Proceedings of the 2022 DiGRA International Conference: Bringing Worlds Together. 2022 Presented at: DiGRA ’22; July 7-11, 2022; Kraków, Poland.
39. Foster AN. Games and motivation to learn science: personal identity, applicability, relevance and meaningfulness. J Interact Learn Res 2008;19(4):597-614 [FREE Full text]
40. Gee JP. What video games have to teach us about learning and literacy. Comput Entertain 2003 Oct;1(1):20. [doi: 10.1145/950566.950595]
41. Verschueren S, Buffel C, Vander Stichele G. Developing theory-driven, evidence-based serious games for health: framework based on research community insights. JMIR Serious Games 2019 May 02;7(2):e11565 [FREE Full text] [doi: 10.2196/11565] [Medline: 31045496]
42. Saracevic T. RELEVANCE: a review of and a framework for the thinking on the notion in information science. J Am Soc Inf Sci 2007 Mar 22;26(6):321-343. [doi: 10.1002/asi.4630260604]
43. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/S11031-006-9051-8]
44. Grasse KM, Kreminski M, Wardrip-Fruin N, Mateas M, Melcer EF. Using self-determination theory to explore enjoyment of educational interactive narrative games: a case study of academical. Front Virtual Real 2022 Jun 16;3 [FREE Full text] [doi: 10.3389/frvir.2022.847120]
45. Isbister K, Flanagan M, Hash C. Designing games for learning: insights from conversations with designers. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2010 Presented at: CHI '10; April 10-15, 2010; Atlanta, GA. [doi: 10.1145/1753326.1753637]
46. Isbister K. How Games Move Us: Emotion by Design. Cambridge, MA: MIT Press; 2016.
47. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
48. Kao D, Ratan R, Mousas C, Magana AJ. The effects of a self-similar avatar voice in educational games. Proc ACM Hum Comput Interact 2021 Oct 06;5(CHI PLAY):1-28. [doi: 10.1145/3474665]
49. Kartsanis N, Murzyn E. Me, my game-self, and others: a qualitative exploration of the game-self. In: Proceedings of the International Conference on Interactive Technologies and Games (ITAG). 2016 Presented at: iTAG 2016; October 26-27, 2016; Nottingham, UK. [doi: 10.1109/itag.2016.12]

50. Birk MV, Atkins C, Bowey JT, Mandryk RL. Fostering intrinsic motivation through avatar identification in digital games. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI '16; May 7-12, 2016; San Jose, CA. [doi: 10.1145/2858036.2858062]
51. Dimitriadou A, Djafarova N, Turetken O, Verkuyl M, Ferworn A. Challenges in serious game design and development: educators' experiences. Simul Gaming 2020 Aug 06;52(2):132-152. [doi: 10.1177/1046878120944197]
52. Thomas TH, Sivakumar V, Babichenko D, Grieve VL, Klem ML. Mapping behavioral health serious game interventions for adults with chronic illness: scoping review. JMIR Serious Games 2020 Jul 30;8(3):e18687 [FREE Full text] [doi: 10.2196/18687] [Medline: 32729836]
53. Lukka L, Palva JM. The development of game-based digital mental health interventions: bridging the paradigms of health care and entertainment. JMIR Serious Games 2023 Sep 04;11:e42173 [FREE Full text] [doi: 10.2196/42173] [Medline: 37665624]
54. Laine TH, Lindberg RS. Designing engaging games for education: a systematic literature review on game motivators and design principles. IEEE Trans Learn Technol 2020 Oct 1;13(4):804-821. [doi: 10.1109/tlt.2020.3018503]
55. Seaver C, Bowers C, Beidel D, Holt L, Ramakrishnan S. A game-based learning approach to sleep hygiene education: a pilot investigation. Front Digit Health 2024 Apr 12;6:1334840 [FREE Full text] [doi: 10.3389/fdgth.2024.1334840] [Medline: 38680214]
56. Wilson P. Situational relevance. Inf Storage Retr 1973 Aug;9(8):457-471. [doi: 10.1016/0020-0271(73)90096-x]
57. Rapp A, Boldi A. Exploring the lived experience of behavior change technologies: towards an existential model of behavior change for HCI. ACM Trans Comput Hum Interact 2023 Sep 25;30(6):1-50. [doi: 10.1145/3603497]
58. Gross T, Mueller AL. NotificationManager: personal boundary management on mobile devices. In: Proceedings of the IFIP Conference on Human-Computer Interaction. 2021 Presented at: INTERACT 2021; August 30-September 3, 2021; Bari, Italy. [doi: 10.1007/978-3-030-85610-6_15]
59. Ashforth BE, Kreiner GE, Fugate M. All in a day's work: boundaries and micro role transitions. Acad Manag Rev 2000 Jul;25(3):472-491. [doi: 10.2307/259305]
60. Kamal R, McCullagh P, Cleland I, Nugent C. A hybrid model based on behavioural and situational context to detect best time to deliver notifications on mobile devices. In: Proceedings of the International Conference on Intelligent Emerging Methods of Artificial Intelligence & Cloud Computing. 2021 Presented at: IEMAICLOUD 2021; April 26-29, 2021; Online. [doi: 10.1007/978-3-030-92905-3_2]
61. Auda J, Weber D, Voit A, Schneegass S. Understanding user preferences towards rule-based notification deferral. In: Proceedings of the Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI EA '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3170427.3188688]
62. López G, Guerrero LA. A conceptual framework for smart device-based notifications. J Ambient Intell Human Comput 2020 Feb 26;13(11):5433-5445. [doi: 10.1007/s12652-020-01801-w]

Abbreviations

HSJ: Hero's Sleep Journey
KUAS: Kyoto University of Advanced Science
RPG: role-playing game
ST: Sleep Tamagotchi

Edited by T Leung, N Ahmadpour; submitted 11.07.24; peer-reviewed by M Aresta, L He; comments to author 26.08.24; revised version received 28.08.24; accepted 22.09.24; published 08.10.24.

Please cite as:
Liang Z, Melcer E, Khotchasing K, Chen S, Hwang D, Hoang NH
The Role of Relevance in Shaping Perceptions of Sleep Hygiene Games Among University Students: Mixed Methods Study
JMIR Serious Games 2024;12:e64063
URL: https://games.jmir.org/2024/1/e64063
doi:10.2196/64063
PMID:

XSL•FO
RenderX

Original Paper

Social Gaming to Decrease Loneliness in Older Adults: Recruitment Challenges and Attrition Analysis in a Digital Mixed Methods Feasibility Study

Bas D L Châtel[1,2,3], MSc; Jeroen H M Janssen[1,3], MSc; Geeske M E E Peeters[1,3], PhD; Rense Corten[4], PhD; Rob Tieben[5], PhD; Menno Deen[6], PhD; Elmy J M Hendriks[1], MD; Marcel G M Olde Rikkert[1,3,7], MD, PhD

[1]Department of Geriatric Medicine, Radboud University Medical Center, Nijmegen, Netherlands
[2]Computational Science Lab, University of Amsterdam, Amsterdam, Netherlands
[3]Radboudumc Alzheimer Center, Radboud University Medical Center, Nijmegen, Netherlands
[4]Department of Sociology/Interuniversity Center for Social Science Theory and Methodology, Utrecht University, Utrecht, Netherlands
[5]Games for Health, Eindhoven, Netherlands
[6]Super Menno Monster, Utrecht, Netherlands
[7]Department of Geriatric Medicine, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, Netherlands

Corresponding Author:
Bas D L Châtel, MSc
Department of Geriatric Medicine
Radboud University Medical Center
Geert Grooteplein Zuid 10
Nijmegen, 6500 HB
Netherlands
Phone: 31 0243616772
Email: bastiaan.chatel@gmail.com

Abstract

Background: Digital mental health interventions could sustainably and scalably prevent and reduce loneliness in older adults. We designed an app containing 29 text-based games and a questionnaire-administering chatbot to stimulate intergenerational contact.

Objective: This study aims to evaluate the feasibility of a social gaming app in reducing loneliness among older adults by evaluating recruitment strategies, data collection procedures, and gameplay activity.

Methods: This mixed methods study recruited participants via newsletters, articles, and a social media campaign. We used semistructured interviews and descriptive analysis of questionnaire answers and game data to assess feasibility. Key measures included recruitment reach and efficiency, participant demographics, in-app activity, and app usability and engagement feedback.

Results: The social media campaign reached 192,641 potential participants, resulting in 1363 game downloads. A total of 155 participants (aged 65 years and older: n=34, 21.9% and aged less than 65 years: n=121, 78.1%) provided informed consent, yielding a conversion rate of 0.08%. The recruitment campaign focusing on distanced playful interaction had a significantly (*P*<.001) higher click-through rate (1.98%) than a campaign focusing on research participation (click-through rate=0.51%; *P*<.001). The overall conversion rate from advertisement exposure to research participation was 0.08%. Participants had a mean age of 48 (SD 16) years. The 65 years and older group averaged 70 (SD 5) years, while the less 65 years group averaged 42 (SD 12) years. Additionally, 45.2% (57/126) reported at least moderate levels of loneliness at baseline. Of the initial 554 players, 91 (16.4%) remained active after the first week, and 32 (5.8%) remained active for more than 90 days. Active participants tended to interact with those within their own age group, as indicated by a Pearson correlation of r=0.31 between the ages of the message sender and receiver. Interviews with 12 (48%) participants (aged 65-79 years; female: n=12, 83%) revealed barriers such as excessive chatbot questions and a mismatch between the target group and app design focus. Questionnaire completion rates dropped from 46% at baseline to 10% at follow-up.

Conclusions: These findings underscore the challenges of recruitment and retention for older adults in a fully digital social gaming intervention. Effective recruitment strategies and targeted app design are crucial for engagement. Based on these insights, future interventions should focus on simplified interfaces, clear guidance for gameplay, and addressing the specific needs and preferences of older adults, thereby enhancing the effectiveness of digital mental health interventions.

(JMIR Serious Games 2024;12:e52640) doi:10.2196/52640

KEYWORDS

loneliness; digital health; serious gaming; older adults; recruitment; feasibility study

Introduction

In the Netherlands, people younger than 55 years (20%) and 95 years (62%) experience moderate to severe loneliness [1] according to the De Jong Gierveld Loneliness Scale (DJG) [2,3]. Loneliness increases with age due to social network changes, availability of family members and friends, opportunities for transportation, and health problems [4-6]. Loneliness, in turn, exacerbates these health factors [7], thus creating a negative feedback loop.

There are 2 reasons for limited robust evidence for loneliness interventions: a lack of randomized, blinded trials [8-14] and the difficulty with recruitment. Lonely participants are hard to include in studies, as people typically do not self-identify as lonely or fear the stigma of loneliness [15].

Local interventions are often not scalable without demanding vast resources and are often not transferable to other contexts [16-20]. Digital mental health interventions (DMHIs) might increase efficiency and reach, standardize data gathering, and reduce the costs of care and workforce needed [21]. Such DMHIs have been successfully used in several psychotherapeutic, cognitive remediation, and telepsychiatric interventions and applied to healthy aging and disorders such as attention-deficit/hyperactivity disorder and dementia [22]. The use of DMHIs to reduce loneliness in older adults is scarce. During the COVID-19 pandemic, DMHIs were successfully used to reduce loneliness. However, these DMHIs were not easily scalable and did not focus on strengthening the social networks of an older population [16,23,24].

Applied gaming is a type of DMHI using the idea of “serious gaming” to achieve goals other than entertainment [22]. It adds playful design and game dynamics elements to an intervention to, for example, promote engagement, enhance intervention efficiency and efficacy, and possibly induce behavioral change [25]. Social gaming, a specific form of serious gaming, stimulates the creation of friendships and social interaction [26-29] and can mitigate feelings of loneliness [29]. However, game mechanics and personal factors such as stress levels may affect individual outcomes [30,31]. For example, 1 study found that individuals playing to increase social interaction may experience less loneliness, but people playing for diversion from daily life may feel lonelier [31]. As such, social games need to be designed and evaluated purposefully, as the effect of gaming on loneliness may vary across individuals.

This study evaluated a loneliness-targeting gaming intervention co-designed with game developers, design experts, and representatives from aging communities. Our research aim was to describe the recruitment and inclusion process and search for options for improved efficiency in games for health studies.

Methods

Ethical Considerations

The local ethics committee of the Radboud University Medical Center (2020-6645) reviewed the study and judged it not to fall within the remit of the Medical Research Involving Human Subjects Act. Therefore, the ethics committee approved the study based on the Dutch Code of Conduct for Health Research and Responsible Use, the Dutch Personal Data Protection Act, and the Medical Treatment Agreement Act. Participants provided informed consent in-app before data collection. No compensation was provided to participants, and data was securely stored in a database compliant with European laws (SO 9001, 27001 and NEN 7510), ensuring privacy and confidentiality according to the Dutch Personal Data Protection Act.

Participants

The research aimed to streamline the recruitment and inclusion process and identify ways to enhance efficiency in health-related gaming studies. This study did not involve specific sample size calculations for the recruitment aim, although our original study focused on measuring social interaction through gaming. For details on the sample size calculation of the original study, see Multimedia Appendix 1 [32-34].

We recruited participants via newsletters, studies, workshops, project websites, and all project partners’ networks and social media channels (mainly through Facebook campaigns). Recruitment through the social media campaign lasted from August 2020 to February 2021 (ie, 6 months). The app remained available after the campaign ended. Recruitment efforts mainly targeted potential players aged 65 and older. However, younger players were also targeted to allow younger players to invite, assist, and play with older players. For example, we organized a gaming cocreation workshop with children and included one of the games they created in the app.

No exclusion criteria were applied to prevent in-game social network formation restrictions, allowing everyone to download the app. Participants were divided into 2 groups for analysis: older adults are those aged 65 years and older (“65+”) and younger adults are those younger than 65 years (“65–”) in accordance with the United Nations’ definition [35]. We chose 65+ as the cutoff age because this period is characterized by retirement (the average age at retirement is 65 years and 11 months [36]) and good health, often referred to as “the third age” [37]. This group is also considered well-suited for preventative measures against loneliness later in life [38].

After downloading the app, agreement to the terms of conditions was mandatory to use the app, which included sharing backend data (eg, message type and timestamp and game session information). Participants were then provided information, asked for informed consent, and administered questionnaires by a chatbot. Finally, Dutch-speaking participants 65 years or older could enroll in poststudy interviews assessing user experience regarding game and study design. For more information on participant demographics, see Multimedia Appendix 2.

Playing Together App

Development Process

The PlayingTogether ("SamenSpelen"; Games for Health) app development used an iterative design process between the research team, the game developers, and the target group (Multimedia Appendix 3).

Game Description

PlayingTogether is a photography and text-based social gaming platform on iOS and Android consisting of 29 mini-games designed to elicit personal social interaction. These games are chat groups allowing synchronous and asynchronous interaction, themed around a specific set of rules. Examples include playing Hangman ("Galgje") and exchanging photographs of objects starting with the last letter of the thing depicted in the previous photo ("Fotoslang").

User Interface (UI)

There were 2 major UI versions during this study. The first version was available at launch containing a black-and-white appearance with line art icons as game avatars. The black and white color scheme allowed less distraction regarding design and game mechanics during iterative feedback rounds with target users, as gaining feedback on these aspects was the main goal in this initial phase.

The design closely resembled commonly used messenger apps to facilitate user-friendliness, as over 70% of Dutch people aged 65-75 years used messaging apps in 2019 [39]. However, we focused game design on compatibility with all age groups to make the app more appealing to younger family members and promote inter-generational gameplay. Figures S4A-S4C in Multimedia Appendix 4 show the first version's home page, game repository, and conversational chatbot, respectively. Subsequently, according to a user voting campaign via Facebook, we introduced a new UI, as seen in Figures S4D-S4F in Multimedia Appendix 4. The updated design was developed to aim for a fun and playful feeling for all ages.

Data Collection

We collected data through semistructured interviews and the app. The app collected data through a questionnaire administering chatbot, helping to boost completion rates as participants can stay in the app to answer questionnaires in a conversational format. In addition, it provides game data by storing backend information regarding gameplay.

Recruitment Data

The numbers of impressions and clicks were stored to monitor the reach versus the effectiveness of the recruitment campaign. Impressions refer to the number of people seeing the ad, while clicks refer to the number of people clicking on the link provided within the ad.

Questionnaires

The chatbot was preprogrammed to administer 5 questionnaires at baseline. Participants started by answering a Personal Information Questionnaire (PI) containing questions regarding, for example, marital status, work status, living situation, and education. Next, a mobility score was calculated using the Life Space Assessment [40], which has been translated into several languages with acceptable validity and reliability [41,42], and more recently translated, though not yet validated, for Dutch older adults [43]. This score ranges from 0=totally bed-bound to 120=independently traveled out of town daily [44]. Next, loneliness scores were measured at baseline and repeated after 1, 2, and 3 months as the primary outcome measure by DJG, which was developed and validated in the Netherlands [2,45] and translated to several different languages later on [46]. It consists of a 6-item scale for emotional loneliness and a 5-item scale for social loneliness, with the total loneliness score being categorized into 4 levels: 0-2=not lonely, 3-8=moderately lonely, 9-10=severely lonely, and 11=very severely lonely. The Older Persons and Informal Caregivers Survey, developed and validated in the Netherlands [47,48], provided information on demographics, morbidity, quality of life, functional limitations, emotional well-being, social functioning, and health services use [47]. Finally, the Network Domain Identification and Significance (NDIS) questionnaire, developed and validated in the Netherlands [32], provided information on participants' social networks at baseline and end of participation. Participants aged 65+ had to answer all questionnaires, while younger participants only answered the PI and DJG questionnaires.

Game Data

Game data entails a session creation timestamp, a list of player IDs, session IDs, and game names. We stored message length, timestamp, sender ID, and type (text or photo) for every message sent within a session. Social network data is stored implicitly through message receivers and senders. Player IDs were linked with profile creation time at first login and, if given, their age.

Interviews

Participants who agreed to be contacted for a qualitative evaluation were asked to participate in a semistructured telephone interview in March 2021. All interviews were voice recorded and conducted by a female project member (EH) according to an interview guide (Multimedia Appendix 5) with primarily open-ended questions structured around the following main topics: barriers and facilitators regarding the app, games, and chatbot. Participants provided informed consent before their interviews, field notes were taken during the interviews, and the audio recordings were transcribed verbatim and anonymized afterward. Data saturation was reached after 12 interviews.

Analysis

Assessment of App Download Funnel

We applied A/B testing to compare social media recruitment strategies where *A* focused on participation in research and *B* focused on distanced playful interaction. Each focus was divided into 3 different emphases. For focus *A*, the emphasis was on A_1: reducing loneliness in older adults, A_2: studying the effect of gaming on valuable interpersonal contact, and A_3: the impact of gaming on well-being. Focus *B* involved the emphasis on B_1: the need to play together due to the COVID-19 pandemic, B_2: the app facilitating more family time, and B_3: encouraging play on special occasions often used for social interaction (eg, Christmas).

We used Facebook's A/B testing functionality to analyze the effectiveness of these recruitment strategies. This method randomly advertised the 2 focuses and their underlying emphases while recording their total number of impressions and the number of resulting clicks. Then, the click-through rate (CTR), that is, clicks divided by impressions, and total conversion rate (CR), that is, the total number of app downloads divided by impressions, were calculated from these numbers to measure recruitment strategy effectiveness. Finally, a *z* test for independent proportions was used to calculate whether CTRs significantly differed between advertisement focus *A* and *B*.

To assess recruitment effectiveness outside social media channels, the chatbot also asked participants how they learned about the app or study: advertisement or paper, newsletter, acquaintance, project website, social media, or other.

Assessment of In-App Participant Funnel

We assessed the total number of active profiles at critical points in the participant journey to uncover bottlenecks in app usage. This entailed evaluating the number of people who gave consent to the terms and conditions, agreed to read and provide informed consent to the questionnaire chatbot, answered the questionnaires at different time points, and sent messages.

Analysis of In-Game Collected Data

Python (Python Software Foundation) was used for a descriptive analysis of sample characteristics, usage dynamics, and social network information. The degree distributions are described using the mean, minimum, maximum, and SD. Furthermore, the median is reported with the first and third quartiles. A preliminary descriptive analysis of app effectiveness was performed by assessing DJG scores over time. Furthermore, the Pearson correlation *r* is used on message data to relate the age of the message's sender and receiver, providing an outcome between r=–1 and r=1. A value of 1 means that everyone only communicates with their own age group, –1 means that everyone avoids communicating with their own age group, and 0 implies that communication is random regarding age. When analyzing network structure, we refer to components, that is, islands of connected players in which no connections exist between players on other islands. Using these components could increase an individual's social network naturally as it provides information to introduce the individual to previously unknown friends of friends. Therefore, nonfully connected components (ie, not everyone in the component knows one another) could be used for targeted introductions and network growth.

Analysis of Interviews

The interview transcripts were coded using ATLAS.ti (ATLAS.ti Scientific Software Development GmbH) and analyzed thematically to identify relevant and overarching themes [49]. Two researchers (JJ and EH) independently coded the interviews, discussing codes and differences afterward.

Results

Overview

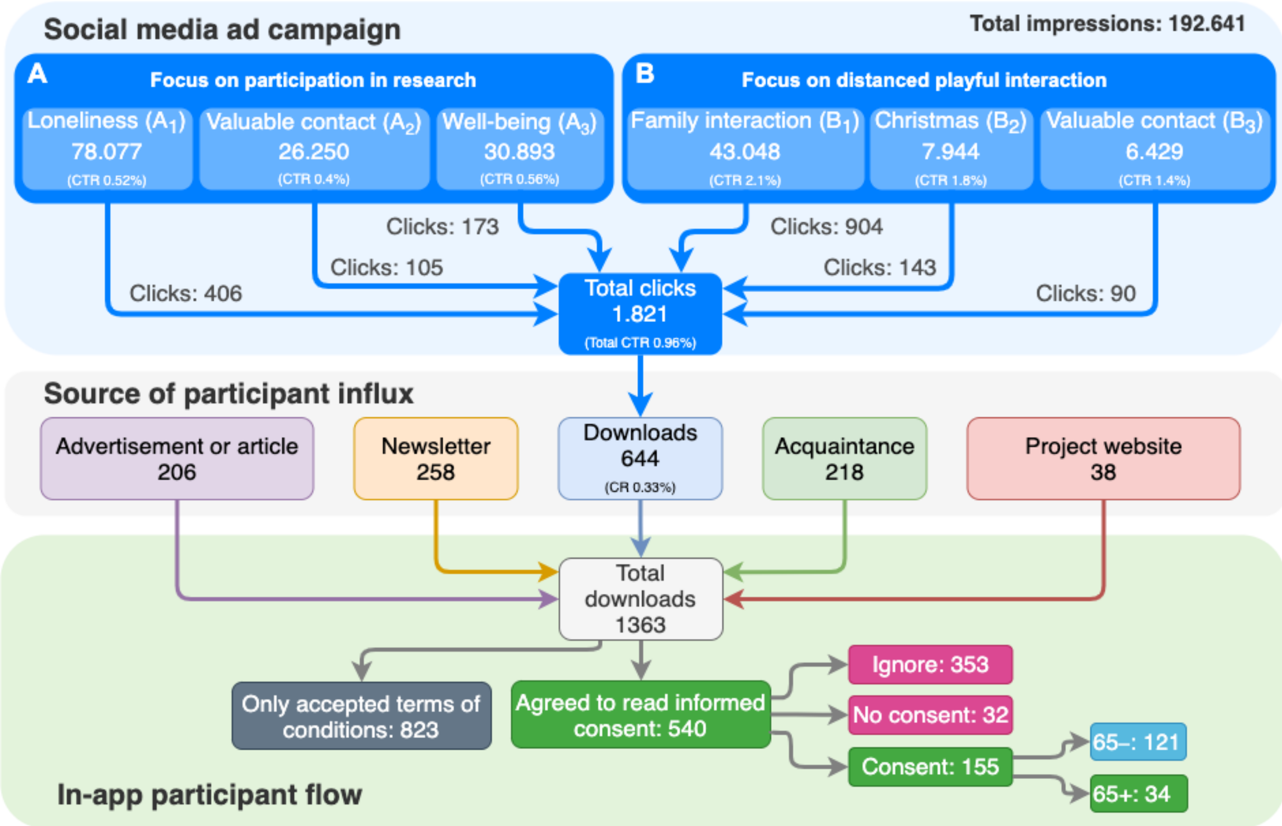
Among the 155 participants who provided informed consent, the average age for all participants was 48 (SD 16) years. The "65+" group had an average age of 70 (SD 5) years, while the "65–" group had an average age of 42 (SD 12) years. Both groups exhibited characteristics of being predominantly healthy, highly educated, married, female, living independently with others, and having children (Multimedia Appendix 2). In total, 57 out of 126 (45.2%) participants who completed the DJG questionnaire at baseline reported at least moderate levels of loneliness. The subgroup of 13 out of 27 (48.1%) participants aged 65 and older, experienced moderate loneliness, while 8 out of those 27 (30%) reported severe loneliness. Twenty-five participants gave consent to be contacted by the researchers, of whom 12 (48%; female: n=12, 83%) consented to be interviewed.

Participant Recruitment

Overview

Figure 1 depicts a higher CTR for focus (P<.001) on distanced playful interaction (CTR=1.98%) than focus *A* on research participation (CTR=0.51%). However, the reach of focus *A* was greater (70.2% of total impressions). Focus *B* generated more clicks, with 49.64% of the total number of clicks from the campaign emphasizing family interaction (B_1: CTR=2.10%, number of impressions=22.35%). The total CR of the social media campaign was 0.33%.

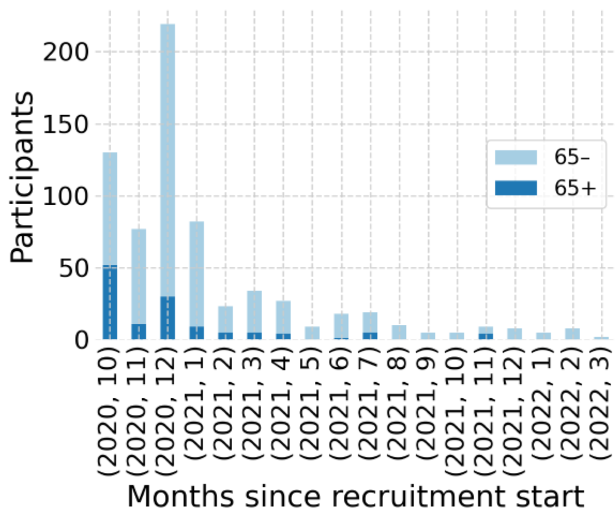
Figure 1. The flow of potential participants from ad exposure to informed consent. On the top level, the number of impressions is shown with the number of clicks and CTR per ad focus and emphasis. The middle level depicts the origin of how participants found the app. The bottom level illustrates in-app participant flow through different consent stages, ending with the number of participants 65 years or older (65+) and participants younger than 65 (65–). CTR: click-through rate.



Of the 1363 downloads, 823 participants only agreed to the mandatory terms and conditions, and 540 participants also agreed to receive research participation information, where the chatbot asked for informed consent. Of these 540 participants, 353 (65.3%) ignored the chatbot, 32 (5.9%) refused, and 155 (28.7%) gave informed consent. Among the 155 participants, 121 (78.1) were 65–, while 34 (21.9) were 65+. The total CR from advertisement exposure to research participation was 0.080% (ie, number of consent/impressions=155/192.641). As channels other than social media were also used, this total CR is an upper bound.

Regarding timing, Figure 2 shows an increase in participants at the app's launch due to initial recruitment efforts. A participant inflow peak follows due to the recruitment campaign between December 18, 2020, and January 4, 2021, followed by a decreased growth, stagnating at around 10 new participants per month. This stagnation indicates the importance of active and continuous recruitment efforts.

Figure 2. Participant influx over time per age group since the start of recruitment.



Feedback Through Qualitative Interviews

Some interviewees indicated that they did not invite someone else because the invitation process was unclear. Furthermore, some indicated that the app was not enjoyable enough to invite someone. One of the participants said:

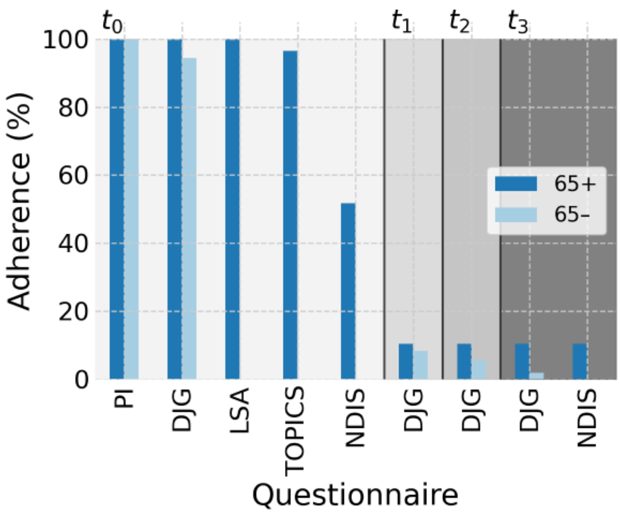
> *It was a barrier for me that I was going to involve others in something that annoyed myself.* [Participant, female, 78]

Data Collection Procedures and Intervention Acceptability

Overview

Questionnaire adherence was relatively steady for PI (N=29, N=109), DJG, and the Life Space Assessment questionnaires at baseline. However, there was a drop from 28 to 15 participants (a 46% drop) who completed the NDIS questionnaire after completion of the Older Persons and Informal Caregivers Survey questionnaire (Figure 3). During the following months, questionnaire completion rates were further reduced to 3 participants (10% relative to baseline), after which the 65+ group's completion rate remained stable while the 65– group continued to decline.

Figure 3. Questionnaire adherence compared to baseline (t0), with follow-up questionnaires each consecutive month. PI: Personal Information Questionnaire; DJG: De Jong-Gierveld Loneliness Scale; LSA: Life Space Assessment; TOPICS: The Older Persons and Informal Caregivers Survey; NDIS: Network Domain Identification and Significance.



Feedback Through Qualitative Interviews

Most interviewees expressed that the chatbot presented information pleasantly and explained the study goal well. However, to some, the chatbot felt slightly artificial, childish, or communicated too quickly. In addition, for some questions, the goal was judged unclear, so interviewees did not feel motivated to answer these questions, as 1 participant illustrated:

> *I did not exactly understand what it was about. So was that questionnaire, with all those contacts, to see if I have contacts? Or to see if I have many contacts? Or was it about finding people with whom I wanted to play a game? That was not clear.* [Participant, male, 77]

Interviewees also indicated that the questionnaires contained too many questions, making the average questionnaire duration unpleasant. One of the participants indicated:

> *I was filling out questionnaires all the time, and I didn't like that.* [Participant, male, 77]

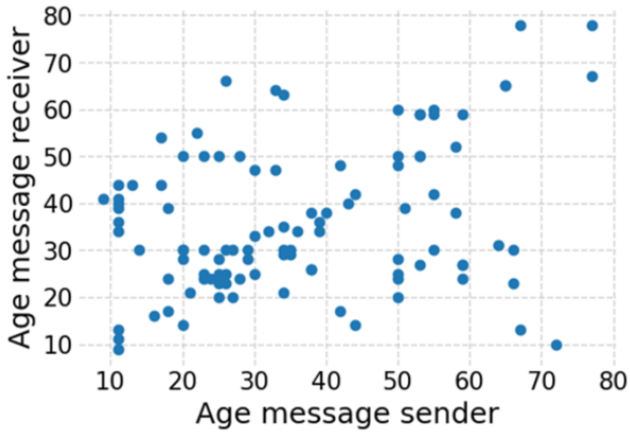
When asked how often interviewees were willing to answer questions, they indicated that monthly would be fine, provided it was not too long. Additionally, the NDIS required participants to disclose information about others, which many interviewees did not appreciate.

Participant Gameplay

Overview

Out of the initial 554 participants, 91 (16.4%) remained active after the first week, and 32 out of the 554 (5.8%) participants had more than 90 days of activity. Furthermore, assessing design choices toward intergenerational play, we found participants tended to play with players in their age stratum (*r*=0.31; Figure 4 and Multimedia Appendix 6).

Figure 4. Age of message sender and receiver. Each dot represents a message from the sender (x-axis) to the receiver (y-axis), based on their respective age.



Feedback Through Qualitative Interviews

Interviewees mentioned that the games were amusing, intuitive, and clearly explained. The games were often considered somewhat slow, old-fashioned, and childish in terms of content and appearance. Some indicated that the games should be more challenging and usually stopped playing after a few times. One of the participants mentioned:

> *Yes, I find it a bit, simple. They are party games, and there is little challenge in it.* [Participant, female, 66]

The number of games was considered enough by some and too low by others. Furthermore, interviewees mentioned not seeing the added benefit of the games in this app over established digital games. Interviewees mentioned that the games are most suitable for playing with children rather than with other adults. However, the interviewees who invited someone else generally invited a good friend or partner of the same generation.

Discussion

Principal Findings

This study investigated the recruitment and inclusion process of a serious game app to reduce loneliness in the older population. Due to the high dropout during the recruitment and onboarding process, substantial overshooting of the intended sample size was required. Of the 193,361 potential participants reached through social media, plus an unknown number through other channels, only 0.08% of participants provided informed consent, indicating significant challenges in recruitment. The A/B testing revealed that the campaign focusing on distanced playful interaction had a significantly higher CTR (1.98%) than the campaign focusing on research participation (CTR=0.51%). In-app activity analysis indicated that 16.4% of participants (n=91) remained active after the first week, but only 5.8% (n=32) stayed active for over 90 days. Interviews highlighted barriers such as excessive chatbot questions and a mismatch between the app's design and the target group's preferences. These findings underscore the difficulties in recruitment and retention.

Even though these low participation rates are considered typical in digital epidemiological research [50,51], especially in studies where enrollment is completely web-based [52], evidence for higher participation rates among older adults (though inconclusive) increased expectations [51]. Due to the nature and goal of our app, it is also possible that the stigma on loneliness [15] played a role in the low recruitment rate. Furthermore, the study's timing during the COVID-19 quarantining period made in-person recruitment strategies (eg, workshops and talks) more challenging, if not impossible.

Most growth came from social media, recruiting more than 200 people in 1 month through Facebook, suggesting participants can be reached. However, few participants completed the entire study. Additionally, when aiming to reach new participants through existing ones, the game did not result in players inviting others once they logged into the app. To that end, improvements to increase game satisfaction must be implemented. This mouth-to-mouth recruitment is considered powerful and should be prioritized [53]. One of the possibilities could be introducing a feature to recommend the game to a Facebook friend and receive in-game rewards. Another way to improve recruitment is to optimize organic user acquisition by, for example, increasing visibility at the top of distributor channel lists like the Apple App Store and Google Play Store, updating metadata often and using metadata keywords based on seasonal contexts [53-55].

After the first week post download, 16.4% of participants remained active; 5.8% remained active after the first 90 days. The questionnaire retention rate remained around 10% compared to the baseline. These retention rates comply with gaming industry standards where a 7-day retention rate of 8.6% and 30-day retention rate of 3.5% in European countries are regular among highly rated gaming apps [56]. Internet-based interventions also experience low retention rates compared to traditional interventions, especially in older populations [57-59]. However, intervention research attrition usually needs to be much lower than that of the gaming industry; attrition of 45% is already described as high [58-61]. Furthermore, player types and participation reasons play a role in retention [62], as differences could exist between players joining for gameplay versus those joining to participate in research. This mismatch between gaming and research standards might imply that initial sample sizes must be much higher than anticipated.

Data collection procedures (ie, backend data collection and chatbot) were considered pleasurable and straightforward. However, language use needs to be adapted per age category as older participants found the chatbot speaking too childlike. The chatbot's tone is essential as negative emotions regarding the chatbot might affect interaction willingness [63], possibly resulting in communication breakdown and increasing the number of missing values [64]. Furthermore, the questionnaires were considered too long and complicated. However, as the attrition rates after 1 month of play were similar between older adults and the younger group, it seems that questionnaire duration is not the only limiting factor, as the younger group did not need to answer all questionnaires. Additionally, even though interviewees preferred answering questionnaires in the app or on the computer, research shows that most older adults prefer answering questionnaires on paper or, if digitally, on the computer [65]. Therefore, our method of administering questionnaires could have been a barrier.

The games were designed to facilitate inter-generational contact, and participants confirmed they were most suited for children and older adults playing together. However, we found that most people played within their age stratum (ie, good friends and partners), corresponding with the literature suggesting that people of different ages do not uniformly intermingle [66]. This age segregation is not only a matter of preferences [66-69], but also a matter of institutional, spatial, and cultural segregation based on age [67,70-72]. Indicating that the design choices toward intergenerational gameplay did not have their intended effect and possibly even had an adverse impact as older users felt the app was too childish. However, the app is possibly only perceived as "childish" as participants did not interact with it within the context of its intergenerational intention.

Strengths and Limitations

The first strength of this study is the use of A/B testing as it helped understand recruitment through social media by providing insights into the discrepancy between the extensive reach needed and app downloads realized. Furthermore, it allowed for adjusting recruitment strategies and optimizing recruitment. A second strength is that we provide a detailed qualitative and quantitative evaluation of an innovation built with stakeholders (ie, older persons and game innovation companies), allowing for iterative analysis and improvement in a controlled and verifiable manner. A third strength regards the data collection during the COVID-19 pandemic. Loneliness has increased during COVID-19 [73,74] and many people, especially older adults, had less in-person social contact [75]. Therefore, the need for digital inclusion of older adults increased [76,77]. Furthermore, research shows that older adults use digital technology to improve social connectedness [78,79] and change how they use and adopt new technology during the COVID-19 pandemic [79]. The PlayingTogether app has the potential to fill the need for digital inclusion as it aims to make digital communication easy and accessible, possibly positively influencing some players to engage more with the game.

A first limitation is the timing of data collection during the COVID-19 pandemic as it made in-person recruitment strategies more difficult and therefore reduced inclusion numbers and lessons learned about the effectiveness of in-person recruitment strategies (eg, workshops and presentations). Consequently, recruitment shifted toward digital strategies. Though this made calculating the advertisement reach possible, it possibly caused the substantial filtering of 99.02% which is still a lower estimate as reach increases when adding unknown contributions through different channels. A second limitation is the representativeness of the target group in the in-game sample. Participants were few (ie, N=155), relatively young, highly educated, and not lonely, so the participant pool did not reflect the target group well. By campaigning toward a broad audience, we believed younger players would invite older potential participants; as lonely, more senior adults are hard to reach [80]. Though we successfully recruited a younger target group, this indirect recruitment did not occur, highlighting the difficulty of reaching a diverse group which hampers the game design process and assessment as the target group is underrepresented. A third limitation is that the app was in iterations of active development during the study, including an update on UI. This is also the case in most other eHealth innovations, as the iterations of innovation cannot wait until the end of a (slow) trial evaluation with long-term follow-up. However, making direct comparisons between players became more challenging with several app versions. Nevertheless, the interviewees' experiences did not differ between versions, indicating that the underlying problems were more profound than those covered by the updates. Finally, our intervention was designed too broad, with the intention to attract both older adults and their grandchildren. We tested game mechanics with older adults and grandchildren, but the intersection of games and questionnaires was not tested. Furthermore, the new UI, although it received the most votes in a digital campaign, was not rigorously tested by older adults and the younger generations. Therefore, the app might have felt "childish" for older adults, and too simple for younger generations, which made the app unsuccessful in connecting the different generations. This might have been overcome by more rigorous pilots and repeated tests within this intended context.

Future Directions

Based on the data and insights gained in this study, we deem the following matters important to address before further (large-scale) evaluation or implementation of social games in alleviating loneliness in older adults. First, game experiences must provide clear guidance to start and continue gameplay. One option could be to introduce volunteers scouted to keep communication energetic. They could catalyze communication and "break the ice" while monitoring participant activity so necessary changes can promptly be implemented. Second, the intervention needs a sharply defined target group that thoroughly tests the app and the intervention procedures. Our approach of using a mixed target group, comprising both older and younger generations, posed challenges in effectively engaging either demographic. Third, substantial initial recruitment investments and continuous advertisement strategies are essential for long-lasting digital effectiveness trials. Current efforts are insufficient for scaling up, as shown by a 99.92% drop from exposure to download and a 90% dropout rate after 1 month.

In-person recruitment could enhance engagement and reduce dropout rates.

There is increasing literature on digital gaming interventions and loneliness [81] and web-based interventions containing a gaming feature and loneliness [82]. However, we believe that further large-scale evaluation only becomes viable once these concerns are addressed. This study adds insights into the gap between the reach of potential participants and the actual recruitment outcome. Furthermore, we have shown how to design, develop, recruit for, and improve games for older adults and where these aspects might need to be improved for future endeavors.

Conclusions

We describe the development process and participant funnel from advertisement to app usage to provide context and guidance for future work. These stages often remain unreported even when crucial for acceptance and, thus, effectiveness. As such, this study underlines the difficulty of recruiting and retaining older adults in a social gaming intervention focused on loneliness. We showed that combining the target group with a hybrid solution to attract different generations leads to difficulty in providing a user experience where all users understand the app's intent and feel appropriately addressed. Furthermore, we have shown the discrepancy between advertisement reach and the number of included participants. This filtering process shows that casting a wide net does not necessarily result in high participation rates, indicating that expectations and recruitment strategies should be realistically assessed when initiating such interventions. More focused small-scale, in-person recruitment strategies might therefore be more effective in increasing participation and decreasing attrition.

Acknowledgments

The authors would like to thank the participants for generously sharing their experiences with them. The research is part of the Social Health Games project, in collaboration with Games for Health and Cooperation Dela.

This work was supported by the Netherlands Organization for Scientific Research (grant 645.003.002).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy restrictions in accordance with the study sponsor but are available from the corresponding author upon reasonable request.

Authors' Contributions

BDLC, JHMJ, GMEEP, RC, RT, MD, and MGMOR conceptualized the trial. JHMJ, BDLC, and EJMH collected and analyzed the data. BC wrote the manuscript, to which all authors gave critical feedback. All authors approved the final draft of the manuscript.

Conflicts of Interest

RT and MD work with Games for Health, which was both a research partner and a commercial partner. Research-wise, they performed design research to explore and validate how to design for social interaction and quality contact to reduce loneliness. Commercial-wise, they investigated if and how a feasible business model could be created for this scope. This dual role was integral to the study's development and implementation, and it is disclosed to maintain transparency and integrity in reporting the research findings.

Multimedia Appendix 1
Sample size calculation.
[DOCX File , 28 KB - games_v12i1e52640_app1.docx]

Multimedia Appendix 2
Participant demographics.
[DOCX File , 299 KB - games_v12i1e52640_app2.docx]

Multimedia Appendix 3
Stages in app design process. An overview of the design process's different stages (project specification, development, and testing). We first specified a problem statement, followed by identifying the target group. Then a minimal viable product prototype was made and released, and a series of tests and feedback rounds with participants followed.
[PNG File , 454 KB - games_v12i1e52640_app3.png]

Multimedia Appendix 4
User interface of the app. Panels A, B, and C show the first user interface in gray tones. Panels D, E, and F show the user interface based on what users found most appealing through a feedback campaign. Panels A and D show the home screen with the

participants' chatgroups. Panels B and E depict the game repository screen with a short explanation of each game and a carousel through which the user can swipe for more games. Panels C and F depict the appearance of a chatgroup, in this case, that of the chatbot administering questionnaires.
[PNG File , 853 KB - games_v12i1e52640_app4.png]

Multimedia Appendix 5
Interview guide.
[DOCX File , 16 KB - games_v12i1e52640_app5.docx]

Multimedia Appendix 6
In-depth analysis of participant gameplay.
[DOCX File , 260 KB - games_v12i1e52640_app6.docx]

References

1. van Campen C, Vonk F, van Tilburg TG. Kwetsbaar en eenzaam? Risico's en bescherming in de ouder wordende bevolking. Den Haag: Sociaal en Cultureel Planbureau; 2018.
2. de Jong-Gierveld J, Kamphuls F. The development of a Rasch-type loneliness scale. Appl Psychol Meas 2016;9(3):289-299. [doi: 10.1177/014662168500900307]
3. de Jong Gierveld J, van Tilburg T. Living arrangements of older adults in the Netherlands and Italy: coresidence values and behaviour and their consequences for loneliness. J Cross Cult Gerontol 1999;14(1):1-24. [doi: 10.1023/a:1006600825693] [Medline: 14617893]
4. Shiovitz-Ezra S. Confidant networks and loneliness. In: Börsch-Supan A, Brandt M, Litwin H, Weber G, editors. Active Ageing and Solidarity Between Generations in Europe : First Results From Share After the Economic Crisis. Berlin: De Gruyter; 2013.
5. van Tilburg TG. The size of the supportive network in association with the degree of loneliness. In: Knipscheer CPM, Antonucci TC, editors. Social Network Research: Substantive Issues and Methodological Questions. Amsterdam: Swets & Zeitlinger; 1990:137-150.
6. Litwin H, Stoeckel KJ. Confidant network types and well-being among older Europeans. Gerontologist 2014;54(5):762-772 [FREE Full text] [doi: 10.1093/geront/gnt056] [Medline: 23749390]
7. Crewdson J. The effect of loneliness in the elderly population: a review. Healthy Aging Clin Care Elderly 2016;8(1):1-8. [doi: 10.4137/HACCE.S35890]
8. Mann F, Bone J, Lloyd-Evans B, Frerichs J, Pinfold V, Ma R, et al. A life less lonely: the state of the art in interventions to reduce loneliness in people with mental health problems. Soc Psychiatry Psychiatr Epidemiol 2017;52(6):627-638 [FREE Full text] [doi: 10.1007/s00127-017-1392-y] [Medline: 28528389]
9. Quan NG, Lohman MC, Resciniti NV, Friedman DB. A systematic review of interventions for loneliness among older adults living in long-term care facilities. Aging Ment Health 2020;24(12):1945-1955. [doi: 10.1080/13607863.2019.1673311] [Medline: 31602993]
10. Coll-Planas L, Nyqvist F, Puig T, Urrútia G, Solà I, Monteserín R. Social capital interventions targeting older people and their impact on health: a systematic review. J Epidemiol Community Health 2017;71(7):663-672. [doi: 10.1136/jech-2016-208131] [Medline: 27834223]
11. Cohen-Mansfield J, Perach R. Interventions for alleviating loneliness among older persons: a critical review. Am J Health Promot 2015;29(3):e109-e125. [doi: 10.4278/ajhp.130418-LIT-182] [Medline: 24575725]
12. O'Rourke HM, Collins L, Sidani S. Interventions to address social connectedness and loneliness for older adults: a scoping review. BMC Geriatr 2018;18(1):214 [FREE Full text] [doi: 10.1186/s12877-018-0897-x] [Medline: 30219034]
13. Seepersad SS. Helping the "poor get richer"—successful internet loneliness intervention programs. In: Sha'ked A, Rokach A, editors. Addressing Loneliness: Coping, Prevention and Clinical Interventions. London, England: Routledge Taylor & Francis Group; 2015:231-240.
14. Gardiner C, Geldenhuys G, Gott M. Interventions to reduce social isolation and loneliness among older people: an integrative review. Health Soc Care Community 2018;26(2):147-157 [FREE Full text] [doi: 10.1111/hsc.12367] [Medline: 27413007]
15. Rokach A. Loneliness updated: an introduction. J Psychol 2012;146:1-6. [doi: 10.1080/00223980.2012.629501] [Medline: 22303608]
16. Boucher E, McNaughton E, Harake N, Stafford J, Parks A. The impact of a digital intervention (Happify) on loneliness during COVID-19: qualitative focus group. JMIR Ment Health 2021;8(2):e26617 [FREE Full text] [doi: 10.2196/26617] [Medline: 33498011]
17. Fairburn CG, Patel V. The impact of digital technology on psychological treatments and their dissemination. Focus (Am Psychiatr Publ) 2018;16(4):449-455 [FREE Full text] [doi: 10.1176/appi.focus.16405] [Medline: 32021582]

18. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO world mental health surveys. Psychol Med 2014;44(6):1303-1317 [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
19. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. Mhealth 2018;4:6 [FREE Full text] [doi: 10.21037/mhealth.2018.03.02] [Medline: 29682510]
20. Ebert DD, Cuijpers P, Muñoz RF, Baumeister H. Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research. Front Psychiatry 2017;8:116 [FREE Full text] [doi: 10.3389/fpsyt.2017.00116] [Medline: 28848454]
21. Khanna MS, Carper MM. Digital mental health interventions for child and adolescent anxiety. Cogn Behav Pract 2022;29(10):60-68. [doi: 10.1016/j.cbpra.2021.05.003]
22. Vajawat B, Varshney P, Banerjee D. Digital gaming interventions in psychiatry: evidence, applications and challenges. Psychiatry Res 2021;295:113585. [doi: 10.1016/j.psychres.2020.113585] [Medline: 33303223]
23. Gilbody S, Littlewood E, McMillan D, Chew-Graham CA, Bailey D, Gascoyne S, et al. Behavioural activation to prevent depression and loneliness among socially isolated older people with long-term conditions: the BASIL COVID-19 pilot randomised controlled trial. PLoS Med 2021;18(10):e1003779 [FREE Full text] [doi: 10.1371/journal.pmed.1003779] [Medline: 34637450]
24. Loveys K, Sagar M, Pickering I, Broadbent E. A digital human for delivering a remote loneliness and stress intervention to at-risk younger and older adults during the COVID-19 pandemic: randomized pilot trial. JMIR Ment Health 2021;8(11):e31586 [FREE Full text] [doi: 10.2196/31586] [Medline: 34596572]
25. Cheng VWS, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
26. Hussain Z, Griffiths M. A qualitative analysis of online gamingocial interaction, community, and game design. Int J Cyber Behav Psychol Learn 2014;4(2):41-57. [doi: 10.4018/ijcbpl.2014040104]
27. Lenhart A, Kahne J, Middaugh E, Macgill A, Evans C, Vitak J. Teens, video games, and civics: teens' gaming experiences are diverse and include significant social interaction and civic engagement. PEW Internet & American Life Project 2008:1-76.
28. Evans J, Evans S, Shank D, Fallon Q. Motivations for social interaction: the case of pokémon go after the fad ended. Soc Sci Q 2020;102(1):547-551. [doi: 10.1111/ssqu.12880]
29. Pallavicini F, Pepe A, Mantovani F. The effects of playing video games on stress, anxiety, depression, loneliness, and gaming disorder during the early stages of the COVID-19 pandemic: PRISMA systematic review. Cyberpsychol Behav Soc Netw 2022;25(6):334-354 [FREE Full text] [doi: 10.1089/cyber.2021.0252] [Medline: 35639118]
30. Kaye LK. Understanding the “social” nature of digital games. Entertain Comput 2021;38(4):100420. [doi: 10.1016/j.entcom.2021.100420]
31. Nebel S, Ninaus M. Does playing apart really bring us together? Investigating the link between perceived loneliness and the use of video games during a period of social distancing. Front Psychol 2022;13:683842 [FREE Full text] [doi: 10.3389/fpsyg.2022.683842] [Medline: 35222171]
32. van Tilburg TG. Delineation of the social network and differences in network size. In: Knipscheer CPM, de Jong Cierveld J, van Tilburg TG, Dykstra PA, editors. Living Arrangements and Social Networks of Older Adults. Amsterdam: VU University Press; 1995.
33. Hulley S, Cummings S, Browner W, Grady D, Newman T. Designing Clinical Research. 4th Edition. Philadelphia, PA: Lippincott Williams & Wilkins; 2013.
34. Konstantinidis EI, Billis AS, Mouzakidis CA, Zilidou VI, Antoniou PE, Bamidis PD. Design, implementation, and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens. IEEE J Biomed Health Inform 2016;20(1):189-200. [doi: 10.1109/JBHI.2014.2378814] [Medline: 26731797]
35. United Nations Department of Economic and Social Affairs Population Division. World Population Ageing. New York, NY: United Nations; 2021.
36. Pensioenleeftijd werknemers nadert 66 jaar. CBS. URL: https://www.cbs.nl/nl-nl/nieuws/2024/18/pensioenleeftijd-werknemers-nadert-66-jaar [accessed 2024-07-15]
37. Carr DC, Komp K. Gerontology in the Era of the Third Age: Implications and Next Steps. New York, NY: Springer; 2011.
38. Schoenmakers EC, Van Tilburg TG, Fokkema T. Awareness of risk factors for loneliness among third agers. Ageing Soc 2013;34(6):1035-1051 [FREE Full text] [doi: 10.1017/s0144686x12001419]
39. CBS. Steeds meer ouderen maken gebruik van sociale media Internet. 2020. URL: https://www.cbs.nl/nl-nl/nieuws/2020/04/steeds-meer-ouderen-maken-gebruik-van-sociale-media [accessed 2022-09-22]
40. Stalvey B, Owsley C, Sloane M, Ball K. The life space questionnaire: a measure of the extent of mobility of older adults. J Appl Gerontol 2016;18(4):460-478. [doi: 10.1177/073346489901800404]
41. Mümken SA, Gellert P, Stollwerck M, O'Sullivan JL, Kiselev J. Validation of the German Life-Space Assessment (LSA-D): cross-sectional validation study in urban and rural community-dwelling older adults. BMJ Open 2021;11(7):e049926 [FREE Full text] [doi: 10.1136/bmjopen-2021-049926] [Medline: 34230022]

42. Curcio CL, Alvarado BE, Gomez F, Guerra R, Guralnik J, Zunzunegui MV. Life-Space Assessment scale to assess mobility: validation in Latin American older women and men. Aging Clin Exp Res 2013;25(5):553-560. [doi: 10.1007/s40520-013-0121-y] [Medline: 23949973]
43. Life Space Assessment (LSA-NL). UMCG. 2024. URL: https://www.umcg.nl/-/uno-umcg/life-space-assessment-lsa-nl [accessed 2024-07-11]
44. Hashidate H, Shimada H, Shiomi T, Shibata M, Sawada K, Sasamoto N. Measuring indoor life-space mobility at home in older adults with difficulty to perform outdoor activities. J Geriatr Phys Ther 2013;36(3):109-114. [doi: 10.1519/JPT.0b013e31826e7d33] [Medline: 22976813]
45. de Jong-Gierveld J, van Tilburg TG. Manual of the Loneliness Scale. In: Methoden en technieken. Amsterdam, Netherlands: Vrije Universiteit; 1990.
46. Uysal-Bozkir Ö, Fokkema T, MacNeil-Vroomen JL, van Tilburg TG, de Rooij SE. Translation and validation of the De Jong Gierveld Loneliness Scale among older migrants living in the Netherlands. J Gerontol B Psychol Sci Soc Sci 2017;72(1):109-119. [doi: 10.1093/geronb/gbv044] [Medline: 26248622]
47. Lutomski JE, Baars MAE, Schalk BWM, Boter H, Buurman BM, den Elzen WPJ, et al. The development of The Older Persons and Informal Caregivers Survey Minimum DataSet (TOPICS-MDS): a large-scale data sharing initiative. PLoS One 2013;8(12):e81673 [FREE Full text] [doi: 10.1371/journal.pone.0081673] [Medline: 24324716]
48. Lutomski JE, van Exel NJA, Kempen GIJM, van Charante EPM, den Elzen WPJ, Jansen APD, et al. Validation of the care-related quality of life instrument in different study settings: findings from The Older Persons and Informal Caregivers Survey Minimum Dataset (TOPICS-MDS). Qual Life Res 2015;24(5):1281-1293. [doi: 10.1007/s11136-014-0841-2] [Medline: 25381121]
49. Glaser BG, Strauss A. Discovery of Grounded Theory: Strategies for Qualitative Research. England: Routledge; 1968:364.
50. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
51. Fiordelli M, Fadda M, Amati R, Albanese E. Older adults' motivations to participate or not in epidemiological research. qualitative inquiry on a study into dementia in Switzerland. PLoS One 2021;16(2):e0247141 [FREE Full text] [doi: 10.1371/journal.pone.0247141] [Medline: 33577607]
52. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
53. Lassila E. “Free”-to-play game: governing the everyday life of digital popular culture. Crit Perspect Account 2022;87:102434. [doi: 10.1016/j.cpa.2022.102434]
54. Zykov SV. IT Crisisology Casebook: Smart Digitalization for Sustainable Development. Singapore: Springer; 2022:43-63.
55. Panosian H. Marketing your app. In: Learn iOS Application Distribution: Successfully Distribute Apps. Berkeley, CA: Springer; 2017:167.
56. Mobile app trends report: a growing mobile economy. Liftoff. 2019. URL: https://liftoff.io/wp-content/uploads/2019/10/2019-Mobile-App-Trends-Report.pdf [accessed 2022-11-13]
57. Melville KM, Casey LM, Kavanagh DJ. Dropout from internet-based treatment for psychological disorders. Br J Clin Psychol 2010;49:455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
58. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
59. Bouwman T, van Tilburg T, Aartsen M. Attrition in an online loneliness intervention for adults aged 50 years and older: survival analysis. JMIR Aging 2019;2(2):e13638 [FREE Full text] [doi: 10.2196/13638] [Medline: 31518268]
60. Cattan M, White M, Bond J, Learmouth A. Preventing social isolation and loneliness among older people: a systematic review of health promotion interventions. Ageing Soc 2005;25(1):41-67. [doi: 10.1017/S0144686X04002594]
61. Shah SGS, Nogueras D, van Woerden HC, Kiparoglou V. Evaluation of the effectiveness of digital technology interventions to reduce loneliness in older adults: systematic review and meta-analysis. J Med Internet Res 2021;23(6):e24712 [FREE Full text] [doi: 10.2196/24712] [Medline: 34085942]
62. Fleming T, Merry S, Stasiak K, Hopkins S, Patolo T, Ruru S, et al. The importance of user segmentation for designing digital therapy for adolescent mental health: findings from scoping processes. JMIR Ment Health 2019;6(5):e12656 [FREE Full text] [doi: 10.2196/12656] [Medline: 31066705]
63. Perez-Marin D, Pascual-Nieto I, editors. Conversational Agents and Natural Language Interaction: Techniques and Effective Practices. Hershey, PA: IGI Global; 2011.
64. Rapp A, Curti L, Boldi A. The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots. Int J Hum Comput 2021;151(3):102630. [doi: 10.1016/j.ijhcs.2021.102630]
65. Fitzgerald D, Hockey R, Jones M, Mishra G, Waller M, Dobson A. Use of online or paper surveys by Australian women: longitudinal study of users, devices, and cohort retention. J Med Internet Res 2019;21(3):e10672 [FREE Full text] [doi: 10.2196/10672] [Medline: 30869647]
66. Uhlenberg P, De Jong Gierveld J. Age-segregation in later life: an examination of personal networks. Ageing Soc 2004;24(1):5-28. [doi: 10.1017/s0144686x0300151x]

67. Hagestad GO, Uhlenberg P. Should we be concerned about age segregation? Some theoretical and empirical explorations. Res Aging 2016;28(6):638-653. [doi: 10.1177/0164027506291872]
68. Ling R. Exclusion or self-isolation? Texting and the elderly users. Inf Soc 2008;24(5):334-341.
69. Ling R. Should we be concerned that the elderly don't text? Inf Soc 2008;24(5):334-341. [doi: 10.1080/01972240802356125]
70. Portacolone E, Halpern J. "Move or Suffer": is age-segregation the new norm for older Americans living alone? J Appl Gerontol 2016;35(8):836-856. [doi: 10.1177/0733464814538118] [Medline: 25186312]
71. Zhu Y, Zhang Y, Wang Y, Trivic Z. The square dance in China: how sensory design can foster inter-generational interaction and improve older adults' wellbeing. J Urban Des Ment Health 2021;7:15 [FREE Full text]
72. Noon RB, Ayalon L. Older adults in public open spaces: age and gender segregation. Gerontologist 2018;58(1):149-158. [doi: 10.1093/geront/gnx047] [Medline: 28535273]
73. Seifert A, Hassler B. Impact of the COVID-19 pandemic on loneliness among older adults. Front Sociol 2020;5:590935 [FREE Full text] [doi: 10.3389/fsoc.2020.590935] [Medline: 33869519]
74. van Tilburg TG. Emotional, social, and existential loneliness before and during the COVID-19 pandemic: prevalence and risk factors among dutch older adults. J Gerontol B Psychol Sci Soc Sci 2022;77(7):e179-e184 [FREE Full text] [doi: 10.1093/geronb/gbab101] [Medline: 34097025]
75. Freedman VA, Hu M, Kasper J. Changes in older adults' social contact during the COVID-19 pandemic. J Gerontol B Psychol Sci Soc Sci 2022;77(7):e160-e166 [FREE Full text] [doi: 10.1093/geronb/gbab166] [Medline: 34529083]
76. Weil J, Kamber T, Glazebrook A, Giorgi M, Ziegler K. Digital inclusion of older adults during COVID-19: lessons from a case study of Older Adults Technology Services (OATS). J Gerontol Soc Work 2021;64(6):643-655. [doi: 10.1080/01634372.2021.1919274] [Medline: 33882782]
77. Xie B, Charness N, Fingerman K, Kaye J, Kim MT, Khurshid A. When going digital becomes a necessity: ensuring older adults' needs for information, services, and social inclusion during COVID-19. J Aging Soc Policy 2020;32:460-470 [FREE Full text] [doi: 10.1080/08959420.2020.1771237] [Medline: 32507061]
78. Hill R, Betts LR, Gardner SE. Older adults' experiences and perceptions of digital technology: (dis)empowerment, wellbeing, and inclusion. Comput Human Behav 2015;48:415-423. [doi: 10.1016/j.chb.2015.01.062]
79. Haase KR, Cosco T, Kervin L, Riadi I, O'Connell ME. Older adults' experiences with using technology for socialization during the COVID-19 pandemic: cross-sectional survey study. JMIR Aging 2021;4(2):e28010 [FREE Full text] [doi: 10.2196/28010] [Medline: 33739929]
80. Ige J, Gibbons L, Bray I, Gray S. Methods of identifying and recruiting older people at risk of social isolation and loneliness: a mixed methods review. BMC Med Res Methodol 2019;19(1):181 [FREE Full text] [doi: 10.1186/s12874-019-0825-6] [Medline: 31464586]
81. Hausknecht S, Schell R, Zhang F, Kaufman D. Building seniors' social connections and reducing loneliness through a digital game. 2015 Presented at: Proceedings of the 1st International Conference on Information and Communication Technologies for Ageing Well and e-Health; January 01, 2015; Lisbon, Portugal p. 276-284. [doi: 10.5220/0005526802760284]
82. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018;58(3):467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]

Abbreviations

CR: conversion rate
CTR: click-through rate
DJG: De Jong Gierveld Loneliness Scale
DMHI: digital mental health intervention
NDIS: Network Domain Identification and Significance scale
PI: Personal Information Questionnaire
UI: user interface

Edited by T Leung; submitted 11.09.23; peer-reviewed by V Rashedi, M Pulier; comments to author 30.04.24; revised version received 07.06.24; accepted 17.07.24; published 16.10.24.

Please cite as:
Châtel BDL, Janssen JHM, Peeters GMEE, Corten R, Tieben R, Deen M, Hendriks EJM, Olde Rikkert MGM
Social Gaming to Decrease Loneliness in Older Adults: Recruitment Challenges and Attrition Analysis in a Digital Mixed Methods Feasibility Study
JMIR Serious Games 2024;12:e52640
URL: https://games.jmir.org/2024/1/e52640
doi:10.2196/52640
PMID:39412863

Original Paper

Gamification in the Design of Virtual Patients for Swedish Military Medics to Support Trauma Training: Interaction Analysis and Semistructured Interview Study

Natalia Stathakarou[1,2], MSci; Andrzej A Kononowicz[3], PhD; Erik Mattsson[4], MD; Klas Karlgren[1,5,6], PhD

[1]Department of Learning, Informatics, Management and Ethics (LIME), Karolinska Institutet, Stockholm, Sweden
[2]Faculty of Health Sciences, Tokyo Metropolitan University, Tokyo, Japan
[3]Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College, Kraków, Poland
[4]Division for Logistics, Department for Support Services, National Home Guard Staff, Swedish Armed Forces Headquarter, Stockholm, Sweden
[5]Department of Research, Education and Development and Innovation, Södersjukhuset, Stockholm, Sweden
[6]Department of Health and Functioning, Faculty of Health and Social Sciences, Western Norway University of Applied Sciences, Bergen, Norway

Corresponding Author:
Natalia Stathakarou, MSci
Department of Learning, Informatics, Management and Ethics (LIME)
Karolinska Institutet
Tomtebodavägen 18A
Solna
Stockholm, 171 77
Sweden
Phone: 46 707799671
Email: natalia.stathakarou@ki.se

Abstract

Background: This study explores gamification in the design of virtual patients (VPs) to enhance the training of Swedish military medics in trauma care. The challenges related to prehospital trauma care faced on the battlefield require tailored educational tools that support military medics' education and training.

Objective: The aim of the study is to investigate how to design VPs with game elements for Swedish military medics to support learning in military trauma care. By understanding the reasoning and perceptions of military medics when interacting with VPs, this study aims to provide insights and recommendations for designing VPs with game elements that are specifically tailored to their needs.

Methods: The study involved 14 Swedish military medics of the Home Guard–National Security Forces participating in a tactical combat care course. Participants interacted with 3 different VP cases designed to simulate military trauma scenarios. Data were collected through think-aloud sessions and semistructured interviews. The data were analyzed using interaction analysis, structured by the unawareness, problem identification, explanation, and alternative strategies or solutions (uPEA) framework, and reflexive thematic analysis to explore participants' reasoning processes and perceptions and identify possible game elements to inform the VP design.

Results: Mapping the military medics' reasoning to the uPEA framework revealed that study participants became more creative after making a mistake followed by feedback and after receiving a prompt to make a new decision. The thematic analysis revealed 6 themes: *motivation, "keep on trying"*; *agency in interaction with VPs*; *realistic tactical experience*; *confidence, "I know that the knowledge I have works"*; *social influence on motivation*; and *personalized learning*. Participants suggested that game elements such as scoring; badges; virtual goods; progress bars; performance tables; content unlocking; hints; challenge; control; imposed choice; narrative; avatars; sensation; randomness; difficulty adapting; competition; leaderboards; social pressure; progression; and renovation can promote engagement, motivation, and support confidence in decision-making.

Conclusions: Gamification in the design of VPs represents a promising approach to military medical training, offering a platform for medics to practice medical and tactical decision-making in a risk-free environment. The insights gained by the study may encourage designing VPs with game elements, as well as including possibly wrong decisions, their consequences, and relevant feedback, that may support military medics' reflections and decision-making.

(JMIR Serious Games 2024;12:e63390) doi:10.2196/63390

KEYWORDS

military trauma; gamification; game elements; serious games; virtual patients; trauma; medical training; medical education; medical assessment; emergency care; first aid; basic life support; trauma care; medics; military

Introduction

Background

Virtual patients (VPs) have the potential to support the training of competencies required to operate in military settings [1]. VPs are interactive, computer-based simulations of real-life clinical scenarios designed to support health care and medical training, education, or assessment [2]. Gamification, the use of game elements in nongaming contexts [3], is gaining interest in education. Examples of game elements include scoring, avatars, badges, storytelling, and adaptive difficulty levels [4,5]. The incorporation of these elements in education has shown promise in enhancing engagement, motivation, and improving learning outcomes, particularly when these elements promote learning behaviors and attitudes [6]. Gamification appears to be at least as effective as traditional educational methods and often more effective for improving knowledge, skills, and satisfaction in health professions education [7].

Game elements can be combined with other educational methods or technologies. For instance, when gamification is combined with extended reality technologies, such as virtual reality and augmented reality, it offers a powerful approach to modern education by creating immersive and interactive learning environments that make learning more engaging and motivating. The integration of gamification with extended reality can thus transform traditional pedagogical methods, leading to more meaningful and student-centered learning [8]. Applying game elements in the design of VPs for military trauma training has the potential to provide learners with active learning opportunities to prepare for providing medical services under austere conditions [9].

In Sweden, the Home Guard–National Security Forces [10] is a military reserve force of the Swedish Armed Forces that consists mainly of local rapid response units. In addition to personnel who have completed their national service or basic military training, the Home Guard includes a large proportion of various specialists, including military medics.

The Swedish model of military prehospital emergency care has 3 levels, where the first level expects that a wounded soldier on the battlefield shall receive first aid from another soldier or a military medic as soon as possible [11]. The education and training to become a military medic requires essential skills vital to wartime scenarios, where medics are tasked with providing initial basic life support and assisting military nurses or physicians.

Prehospital trauma care on the battlefield differs significantly from that practiced in the civilian sector, particularly in terms of the types and severity of injuries encountered. Military medical personnel face additional challenges in providing care to their wounded teammates under tactical conditions. Unlike natural disasters or civilian incidents such as traffic accidents, where the immediate danger typically subsides and care can be administered in relative safety; battlefield medics might need to deliver care while under hostile fire; frequently dealing with multiple casualties; limited resources; and harsh environmental conditions such as rain, cold, and darkness.

Hemorrhage is the leading cause of preventable death in both military and civilian trauma, making the outcomes of care provided in tactical settings critically important. As gun violence rises in some European countries, injuries more commonly associated with the battlefield, such as gunshot and blast wounds, are increasingly seen in civilian life, representing a growing public health concern [12]. However, the nature of gunshot wounds caused by military ammunition often differs from those typically encountered in civilian settings, primarily due to the higher velocity and a greater kinetic energy of military rounds, which result in more extensive tissue damage [13].

Treatment guidelines developed for the civilian setting do not necessarily translate well to the battlefield. Sometimes, solutions need to be improvised [14]. Courses such as the Tactical Combat Casualty Care (TCCC) initiative are a set of evidence-based, prehospital trauma care guidelines customized for use on the battlefield. These guidelines may be used by battlefield medics, corpsmen, and pararescuemen to acquire trauma management strategies that integrate medical principles with effective small-unit tactics. The curriculum for this training outlines specific objectives, emphasizing the military medics' need to comprehend the effects of trauma on the human body, administer subsequent treatment, and deliver initial prehospital care. Following the training, military medics are expected to be able to make critical decisions on the battlefield, combining medical care with tactical decision-making [15].

While military medics across different armed forces may have varying duties and organizational structures, their exposure to handling trauma patients is limited. This challenge is further compounded by a lack of research on the learning processes and educational strategies tailored to support the training of military medics [14].

The educational value of VPs can be understood through the experiential learning theory [16], which emphasizes the importance of action and reflection in a safe environment that can tolerate errors [17,18]. VPs may offer a potential solution in preparing military medics by providing a controlled environment where they can practice their skills, make critical decisions, and learn from their mistakes without real-world consequences, ensuring they are better prepared for situations they may face in the field. The risk-free environment offered by VPs has been previously emphasized as their key characteristic in a systematic review of VP descriptions [19]. Through VP simulations, military medics can experience a breadth of scenarios, from the common to the rare and complex, ensuring they are better prepared for situations they may face in the field.

Despite the potential of VP simulations to enhance military medical training, there is a lack of studies focused on incorporating gamification into the design of VPs for military medics. This gap highlights the need to explore how to effectively integrate game elements into VP design to meet the unique demands and learning objectives of military trauma education while increasing engagement.

Aim

The aim of the study is to investigate how to design VPs with game elements for Swedish military medics to support learning in military trauma care. By understanding the reasoning and perceptions of military medics when interacting with VPs, this study aims to provide insights and recommendations for designing VPs with game elements that are specifically tailored to their needs. The study has two research questions:

1. How do Swedish military medics reason when interacting with the VPs?
2. What are the military medics' perceptions of VPs?

Methods

Study Setting and Participants

A course took place from October 21 to 28, 2023, in a military training area located in Väddö, Sweden, as part of the TCCC training for the Swedish Home Guard. The training aimed to teach Home Guard soldiers basic concepts of TCCC (as part of the TCCC Combat Lifesavers course [TCCC-CLS]) based on the published guidelines [20,21]. This was the first time the TCCC-CLS course took place for the Swedish Home Guard, using the US-standardized TCCC training guidelines and material. The course was attended by 30 participants; 14 (47%) participants were recruited to participate in the study with purposeful sampling. Participants with different professional backgrounds were included to achieve variation in the data [22]. In particular, 11 (37%) participants were soldiers of the Home Guard in Sweden with mixed backgrounds and 3 (10%) participants were course instructors.

Data Collection

The data collection took place on October 27, 2023, in Väddö, Sweden. A think-aloud method was used while each participant interacted individually with the 3 VP cases, to answer the first research question. Directly after the think-aloud sessions, each participant discussed the cases in semistructured interviews. Two authors (NS and KK) conducted think-aloud and semistructured interviews with the 14 participants.

Development of the VPs

Three VP cases were developed using the Open Labyrinth open-source system [23]. The content of the VP cases was cocreated with the support of the Home Guard of Sweden in collaboration with the instructors in combat casualty care. The content was designed first in a PowerPoint (Microsoft Corporation) format. To inform the decision points and the content of the VPs, recorded sessions took place with Home Guard members with mixed backgrounds, specifically educators with backgrounds in medicine, nursing, ambulance nursing, combat casualty care, and TCCC-CLS. The 3 cases addressed trauma management caused by gunshot and blast injuries. The cases included key medical and tactical decisions. The medical decisions concerned the use of the <c>ABCDE protocol [21] where the "c" relates to stopping the catastrophic bleeding. Because the medics learn the <c>ABCDE protocol and the military alternative "MARCH" (massive hemorrhage, airway, respirations, circulation, head injury, and hypothermia), the VPs were developed without explicitly referring to the letters of the acronyms, as these can easily be memorized without full understanding of each letter's meaning. For instance, instead of phrasing a decision alternative as "select A," the VPs listed the decision as "select to check the airway."

Design Principles

The VP cases were developed using design principles identified in the literature and the authors' own experience. Photos from exercises and educational materials of the Home Guard in Sweden, along with images generated using OpenAI's DALL-E artificial intelligence system [24], were used. A key design principle was to "emphasize a feeling of presence by keeping the patient visible." To maximize the sense of presence and not let the interaction become an overly theoretical exercise, the trauma patient was kept visible during every decision the learner made as if the patient was in front of the learner.

We aimed to emphasize the participants' control and the feeling of being in charge rather than a spectator. The interaction with the VP focused on the user's next step, enhancing emotional involvement and placing the responsibility for patient care on the learners, requiring them to act quickly and manage the visible trauma. VPs were designed to prompt learners to make decisions as if they were a military medic using first-person language, for example, "I decide to approach the injured soldier." This approach aimed to immerse the learner in the role and responsibility of a military medic. Incorrect medical decisions, such as failing to prioritize stopping catastrophic bleeding, would result in patients dying from hemorrhage. Suboptimal tactical decisions endangered the user's character, potentially resulting in injury or death.

The VPs were designed as linear scenarios [25]; if a wrong decision was made, the learner was asked to try again. By clicking "try again," the learner would be navigated back to the previous key decision node, either medical or tactical, and asked to try again. There was one possible pathway for all VPs, leading to one final case outcome.

A 2-fold scoring system was developed to gamify the VPs, awarding or deducting 10 points for every correct or incorrect medical or tactical decision. Figures S1-S3 in Multimedia Appendix 1 visualize the 3 VPs, with TKD representing tactical key decisions and MKD representing medical key decisions. Table S1 in Multimedia Appendix 1 provides an example of the key decision content of VP1.

The 3 cases were designed with varying levels of difficulty. The first case was the simplest medically, while the third was the most complex. However, the first case involved massive bleeding, making its simplicity not immediately apparent. The scenario provided hints about suspicious objects in the remote area, requiring the user to prioritize tactical safety before assisting the patient. The second case was somewhat more challenging, involving facial bleeding that needed to be controlled in a remote area with an active threat, requiring the medic to approach the patient carefully while taking cover behind stones. In the final case, the patient initially appeared to be in good condition—smiling but slightly confused—yet eventually died due to brain bleeding. In all 3 scenarios, the learner assumed the role of a military medic, tasked with aiding an injured soldier. Each case began with a critical tactical decision, where a wrong choice led to the medic's death, either from stepping on a cluster bomb or coming under enemy fire. Prioritizing safety was essential in all cases, even though it was not always immediately obvious.

Piloting the VP Cases

Two researchers, associated with Karolinska Institutet, both with civilian nursing backgrounds and specific experience with medical education, gamification, and civilian emergency and medical education, respectively, were contacted for piloting the VP cases. The 2 researchers conducted the think-aloud sessions and participated in the semistructured interviews. At the last step, they provided feedback as experts, using experience from gamification, emergency medicine, and medical education. After receiving feedback and improving the VPs, the content was validated with the Home Guard combat casualty care course leader.

Think-Aloud

Think-aloud is a technique where the participant is asked to talk-aloud while solving a problem. Participants are encouraged to verbalize their thoughts as they work through the task. Think-aloud is a method that, in principle, does not lead to much disturbance of the thought process [26]. In our study, participants were prompted to elaborate on their decision-making while playing the cases.

Interaction Analysis

Interaction analysis was used to analyze the video-recorded think-aloud sessions. Interaction analysis brings focus on the important moments in the learning activity, such as the interaction of military medics with the VPs and their reasoning while making tactical and medical key decisions. This method eliminates the researchers' bias by grounding assertions on the video-recorded sessions, making them verifiable [27].

The primary objective was to identify recurring patterns of user actions while interacting individually with the VPs. First, NS and KK reviewed the video-recorded think-aloud sessions. We observed that the participants' comments increased considerably whenever they made a mistake. Therefore, we decided to map the data from the think-aloud sessions to the unawareness, problem identification, explanation and alternative strategies or solutions (uPEA) framework to analyze how they reasoned in relation to becoming aware of problems and dealing with those, particularly after encountering consequences of wrong decisions. Our analysis followed a deductive-inductive approach to qualitative data analysis [28]. The uPEA framework has been previously used to analyze participants' reflections during post-simulation debriefings [29].

Semistructured Interviews

Semistructured interviews were conducted to understand how the military medics perceive the VPs and how game elements can inform the VP design. An interview guide with questions was developed by the first author, reviewed by 2 additional authors (AAK and KK), and the questions can be found in Table S2 in Multimedia Appendix 1.

Thematic Analysis

To analyze the semistructured interviews and identify patterns and themes in the dataset, a reflexive thematic analysis was used following the 6-step approach suggested by Braun and Clarke [30]. The thematic analysis was conducted by the first author (NS) and discussed with AAK and KK in a collaborative and reflexive manner, aiming to achieve richer interpretations of meaning. First, the transcripts were generated using the automatic transcription tool Word (Microsoft Corporation) "transcribe." NS actively revisited all interviews and recordings and went through the automatically generated transcripts to edit them and generate preliminary notes; then codes were created to encapsulate the meaning of the text. In an iterative manner, themes were actively developed and revisited to form the results and emphasize themes relevant to the research questions.

An additional round of coding was dedicated to associating the participants' statements to game elements. Using a framework by Toda et al [4] and Maheu-Cadotte et al [5], which we previously used in our research, we performed this mapping organically following a deductive-inductive approach. We actively connected phrases from participants' discussions to game elements based on our experience in gamification.

Reflexivity

The reflexive thematic analysis as outlined by Braun and Clarke [30] is a widely recognized method for identifying, analyzing, and reporting patterns (themes) within data. Their approach emphasizes the importance of the researcher's reflexivity in the analysis process. NS has a background in health informatics and an interest in educational technologies, particular VPs, and gamification. This background has been informing the way that the interview data were interpreted. KK has a background in interaction design and is a senior researcher within the field of medical education, with a particular interest in technology-enhanced learning. AAK has a background in computer science and is experienced as a researcher in topics around technology-enhanced medical education and clinical reasoning.

Ethical Considerations

Ethics approval was sought from the Swedish Ethical Review Authority (diarienr 2016/1701-31), followed by an amendment in 2020 (diarienr 2020-01660). The Swedish Ethical Review Authority exempted the study from the requirement for ethics approval, as it did not involve handling sensitive personal data, as defined by the Swedish Ethical Review Act.

The video recordings were limited to capturing the laptop screens used, demonstrating participants navigating through the

VPs with synchronized voice audio tracks of the participants. No identifiable data about the observed participants were collected. Recordings began only after each participant received oral and written information about the study. Participation could be interrupted at any time without any consequences. All study participants signed a consent form.

Results

Overview

A total of 14 military medics participated in the think-aloud and interview sessions. All participants completed the 3 VP cases. The VP cases and the nodes visited by the participants are visualized in Figures S4-S6 in Multimedia Appendix 1. Most visited nodes are visualized with dark gray ranging to light gray and white for less-visited nodes and nodes not visited at all. The number within each node shows the number of clicks per node.

Swedish Military Medics' Reasoning When Interacting With the VPs

Overview

The uPEA categories were mapped to the participants' discussions around key decisions as follows: "unawareness" was used to refer to how participants were unaware of a problem that they would subsequently encounter and become aware of. The participants could, for example, comment on the background information, text, and images seemingly unaware of any pending danger or risks just before making a mistake. "Problem" was used to refer to reflections on how the participants realize that they made a suboptimal or incorrect decision. At this point, they discover that the consequences of their decision were not the desired ones. "Explanation" was used to categorize discussions and explanations of why the problem occurred, for instance, to justify a wrong decision made by the participant. The category "alternative strategies or solutions" was used to refer to alternative decisions made after recognizing the problem and considering actions that would avoid similar issues in the future. An overview of different kinds of reflections made by the participants during the different phases is presented in Figure 1.

An example of a participant's (participant 10) reasoning, mapped to the uPEA framework, is presented in Figure 2. The example outlines the participant's reasoning process, starting with his initial unawareness of the risks as he decides to proceed toward the injured patient without hesitation. As he moves forward and steps on a mine, he realizes that he had overlooked the danger, identifying this as the problem (P, which stands for problem identification). He then provides an explanation for his mistake, attributing it to the brevity of his military training (E, which stands for explanation). Finally, he elaborates on how he would act differently in a similar situation in the future (A, which stands for alternative strategies or solutions).

The interaction analysis and the uPEA mapping allowed us to generate the following categories that bring insights about how military medics reason while interacting with the VPs.

Figure 1. An overview of different kinds of reflections observed during the different phases of the unawareness, problem identification, explanation and alternative strategies or solutions framework. VP: virtual patient.

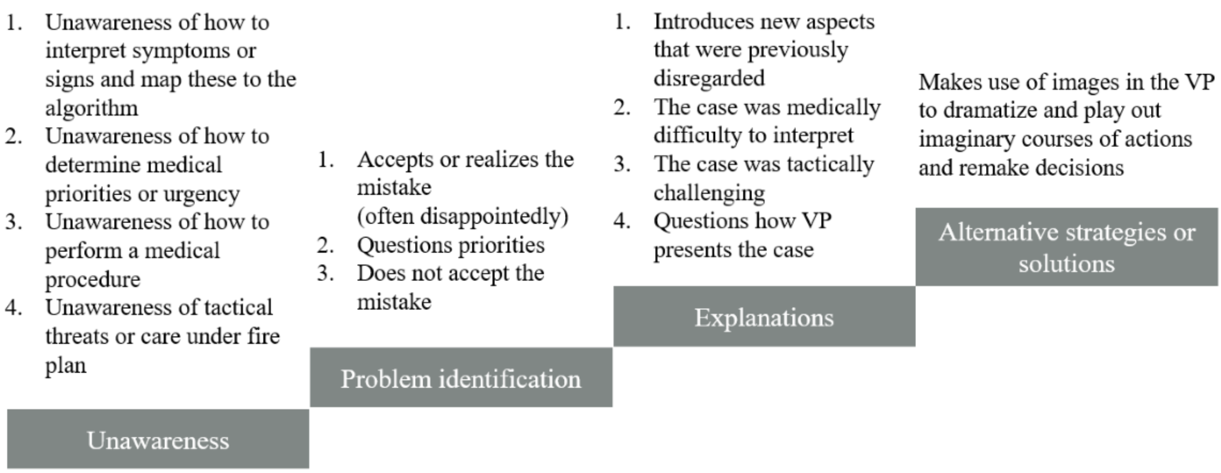
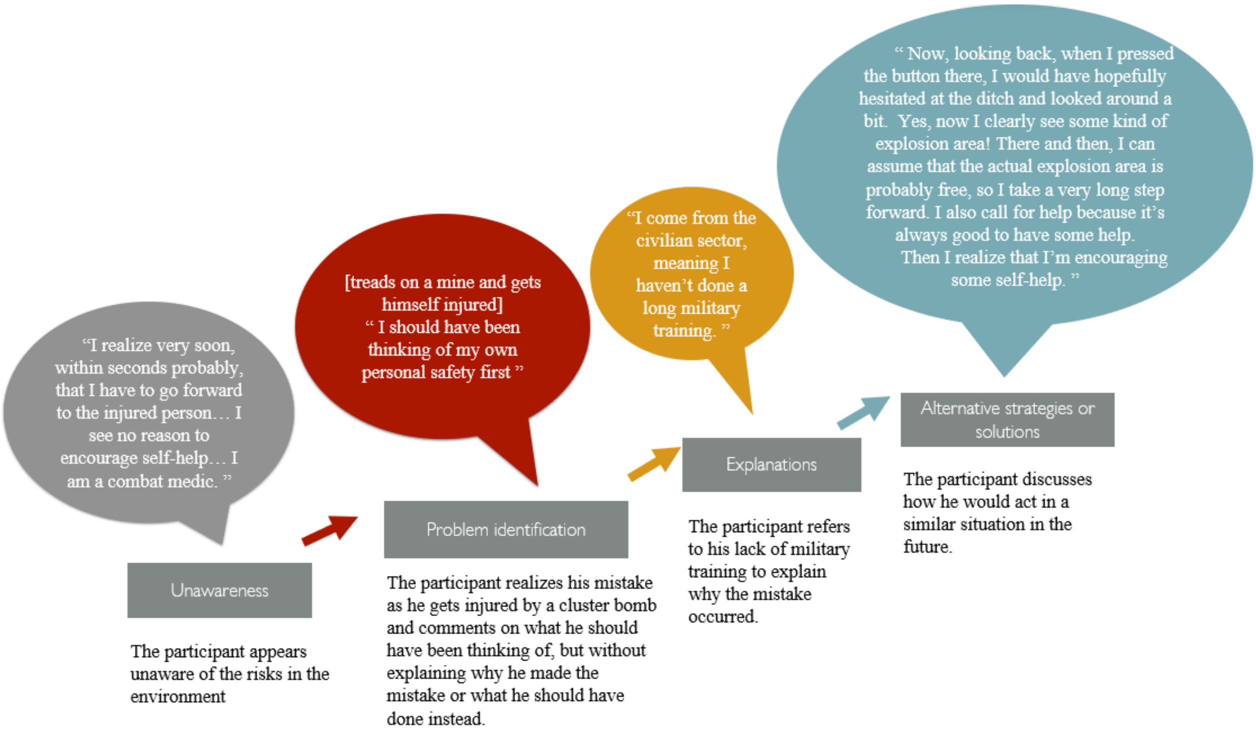


Figure 2. An example of a participant's reasoning mapped to the unawareness, problem identification, explanation and alternative strategies or solutions framework (participant 10).



Unawareness

Overview

Participants were often unaware of factors that contributed to the error that they would later make. They seemed not to reflect on the aspects that they would later bring up as important in their explanations. For instance, participants would exclaim that they should hurry to the wounded patient without encouraging self-help because it was part of their role as a combat medic without mentioning the risks involved. Alternatively, looking at an image of a wounded patient concluded that the patient would probably not die, thus apparently unaware of the risk that the patient had latent fatal injuries. In subsequent sections, we have categorized different themes of unawareness.

Unawareness of How to Interpret Symptoms or Signs and Map These to the Algorithm

In the first VP case, the term "massive bleeding" was used, and all participants recognized the necessity to prioritize it and stop the bleeding. Thus, everyone made the correct medical decision in the first VP. However, in the second case, several military medics were confused when the term "pulsatile bleeding" was introduced instead of the usual term "catastrophic bleeding" and had difficulty mapping the <c>ABCDE algorithm to the case, even if there were hints of the necessity to prioritize to stop the bleeding. This implies that even if the participants have been trained to the protocol with clear examples of what step to prioritize, they had difficulties applying the same concept to a broader case without a clear hint referring to "catastrophic" bleeding:

> *There are many things to do, but I need to prioritize cleaning the mouth so that he can breathe. It would be good to stop the bleeding too...But it [the injury] is on the face and jaw...But there is pulsating bleeding in the injured area...It is not catastrophic bleeding. Then one could try...I am between these two [pointing at two options, to stop the bleeding vs to clean the mouth]. I will select this one thought [to clean the mouth].* [Participant 5]

Unawareness of How to Determine Medical Priorities or Urgency

In the second VP case, the patient with trauma on the face caused confusion because of the anatomical location, as the trauma was around the missing jaw. Many participants thought that they should try to free the airway and prioritize breathing instead of stopping the bleeding:

> *I will try to see whether there are things one can see [in the mouth] so that I can take them out. And then I will try to stop the bleeding, but this is after I check the airways, so this is the order I would follow...* [Participant 4]

In the third VP case, a picture of a smiling patient appears, who has internal brain bleeding. Although the case is considered medically challenging, several medics understood that there might be a head injury, even if the medical urgency was not always clear, which resulted in the patient's death:

> *I think he has a head injury. He is probably not going to die, and he doesn't seem to have anything wrong with his limbs...* [Participant 11]

Unawareness of How to Perform a Medical Procedure

While dealing with the second VP case, some participants thought that it was practically impossible to stop the bleeding around the mouth area. Therefore, they proceeded with other priorities. While the anatomical positioning of the injury might have influenced their decision-making, the practical perception of how to stop the bleeding and its feasibility was unclear to some of the participants:

> *If it is pulsatile bleeding, it is not easy to set something [means like a tourniquet] on the face. So, I will select to free the airways.* [Participant 8]

Unawareness of Tactical Threats or Care Under Fire Plan

Many medics were confused in making a decision between "directing the casualty to move to cover and apply self-aid" and "selecting to run to help the patient." Directing the casualty and encouraging self-aid is one of the priorities of the first phase of the care under fire or threat, according to the TCCC protocol. However, given the design of the VP, which presented a bleeding patient screaming for help, many medics rushed to help the bleeding patient, which resulted in stepping on a cluster bomb and injuring themselves:

> *...I am safely approaching the injured person on my knees. Securing my weapon, I realize very soon, within seconds, probably, that I have to go forward [to the injured person] because I see that he probably bleeds a lot. I see no reason to encourage self-help because he is actually an injured person, and I am a combat medic.* [Participant 10]

Problem Identification

Overview

The VPs provided feedback to the users after each correct or incorrect decision. Sometimes the feedback was just textual, but often the consequences of a decision were quite concrete, so that it was obvious to users that they had made a mistake. If a user was not careful enough and stepped on a mine, such a decision would result in an explosion injuring them, presented in an image. In this section, we describe the reactions and reflections of the participants when it became clear that they had made an incorrect decision. These varied between acceptance, acceptance with a large portion of disappointment, questioning the priorities that should be made, and not accepting the mistake.

Accepts or Realizes the Mistake (Often Disappointedly)

Some participants accepted the mistake after it became obvious to them, through experiencing the consequences or through the VP feedback. Often, realizing and accepting the mistake involved disappointment:

> *...So is it over now? [after stepping on a cluster-bomb, which resulted in an explosion].* [Participant 3]
>
> *…Oh, no! I should have been thinking of my own safety first.* [Participant 5]

Questions Priorities

Reactions to the wrong decision-making also included questioning the priorities presented in the VP. For example, questioning the priorities in the basic management plan for care under fire or threat, where instead of the tactical priority "return fire and take cover," the participants had to select "direct casualty to move to cover and apply self-aid if able or when tactically feasible, move or drag casualty to cover" because, in the case, there were no enemies presented:

> *So, isn't the first priority to kill the enemies?* [Participant 1]

Does Not Accept the Mistake

In some other cases, the participants did not accept the mistake after it became obvious to them:

> *This was wrong, ok. I need to read the text again. Aha, ok, it states that it is pulsatile bleeding. It is just that I read wrongly. Yes, I will stop the bleeding first. It is just that I read wrongly.* [Participant 2]

Explanations

Overview

After having made a mistake, the participants would often stop to reflect on why the mistake took place. Participants offered explanations to why they selected a suboptimal option, such as lack of tactical training, and they sometimes also seemed to rationalize their decisions and attribute mistakes to limitations of the VP scenarios or their interpretations. This process often involved participants justifying their actions by uncovering aspects they had not initially considered.

Introduces New Aspects That Were Previously Disregarded

> *There is no sign of threat in the picture; that's why I thought of running to help [the patient]. There can be more cluster-bombs on the way, and that's why I would follow the same pathway as him. But of course, there are risks on the way...* [Participant 3]

The Case Was Medically Difficult to Interpret

> *It was wrong, yes, but it is a bit difficult to stop the bleeding on the mouth area...That's how I thought [and skipped it]. When the patient looks like that [pointing at the picture], then I should clear the mouth because otherwise he will not have a free airway. But if it is catastrophic bleeding, then of course I will have to press [stop the bleeding] first...But the picture is limiting the understanding of how serious the catastrophic bleeding is.* [Participant 7]

The Case Was Tactically Challenging

> *I come from the civilian sector, meaning I haven't done a long military training.* [Participant 5]

Questions How the VP Presents the Case or Reality

VP limitations that were mentioned to justify wrong decision-making were the (1) picture limitations, (2) a limited number of listed alternatives to choose from, (3) a limited number of options to select at a time, and (4) limited sensory simulation (only sight is involved, and no other senses)

> *In a real situation, one receives more information. One has a smell, a taste; it is like déjà vu. I think I would perform better if I was there, and I had all these things in front of me, so I don't need to get all this information [from the VP text]. It would be a long text to give all these details, i.e. how much he is*

> *bleeding from the face...But one can't quantify that either, right?* [Participant 5]

Alternative Strategies or Solutions: Makes Use of Images in the VP to Dramatize and Play Out Imaginary Courses of Actions and Remake Decisions

At the last stage of the reasoning, some of the military medics shifted from justifying their wrong decision-making to reflecting and remaking correct decisions. Not all participants would reflect on alternative solutions and quickly just selected another VP option. However, others engaged in describing scenarios, dramatizing, and playing out imaginary courses of action that they could have chosen instead of making the error. Two examples of participants immersing themselves in the role of a military medic are presented in blockquotes later and in Figure 2. These 2 participants had previously made the wrong decision to quickly approach and help a wounded soldier, without sufficient caution, thereby becoming injured:

> *I see now that it seems to be in the picture an open territory and that it might be challenging to hide oneself. But there seem to be stones on the ground, and I can imagine that there must be somewhere a bigger stone that I can use to hide myself...If there was a lot of fire going on, I would have approached Sandstedt [the wounded soldier] carefully, ready to exchange fire. If there were enemies around, I would shoot them and prioritize that over Sandstedt, so that we avoid being two dead people.* [Participant 1]

Military Medics' Perceptions of VPs

Overview

The thematic analysis resulted in the following themes: *motivation, "keep on trying"*; *agency in interaction with VPs*; *realistic tactical experience*; *confidence, "I know that the knowledge I have works"*; *social influence on motivation*; and *personalized learning*. Within each theme, corresponding game elements were developed through the second round of coding and linked to the participants' discussions. These game elements were inspired by previous studies [4,5] and were further developed and defined based on the authors' experience. Table 1 presents a summary of the game elements mapped to the themes and defines them within the context of this study.

Table 1. Summary of defined game elements and how they map to the developed themes.

Game elements	Definitions	Themes
Scoring	Assigning points for correct actions or decisions and deducting points for incorrect ones to provide feedback and encourage improvement	1, 4, and 6
Badges	Visual symbols of achievement awarded for completing specific tasks or reaching milestones	1
Virtual goods	Digital items earned through performance that can be collected or used within the VP[a] environment	1
Progress bars	Visual indicators showing the learner's progress through tasks or learning objectives	1
Performance tables	Detailed overviews of the learner's performance across multiple tasks or criteria	1
Content unlocking	Access to new content or pathways based on the learner's performance or decisions	1
Hints	Clues provided to help learners find the correct answers without giving them away directly	1
Challenge	Elements designed to test the learner's abilities and maintain engagement	1
Avatars	Digital representations of the learners or patients within the VP environment	1 and 2
Control	The player's ability to influence the game environment and outcomes through their actions	2
Imposed choice	Providing learners with multiple decision options and requiring them to select a specific one to proceed, preventing further progress in the virtual patient scenario until the correct choice is made	2
Narrative	The structured sequence of events and choices that shape the learner's experience	2
Sensation	The use of sensory stimuli, such as visual and auditory cues, to enhance immersion	3
Randomness	Introducing unpredictable elements into the simulation to create a more realistic and challenging environment	3
Difficulty adaptation	Adjusting the level of difficulty based on the learner's performance	4 and 6
Competition	Encouraging learners to compete with each other, often through leaderboards or similar mechanisms	5
Leaderboards	Rankings that display top scores to motivate learners through competition	5
Social pressure	Influence from peers or the community that can motivate or demotivate learners	5
Renovation	Opportunities to retry decisions and improve outcomes based on feedback	6
Progression	Tracking and visualizing a learner's development and achievements over time	6

[a]VP: virtual patient.

Theme 1: Motivation, "Keep on Trying"

Motivation was a central theme in the discussion with the participants. While referring to the VPs, participants reflected on features and aspects that could enhance motivation during their interactions. Participants discussed scoring and its potential to increase motivation by encouraging continuous efforts to achieve higher scores. Participants also highlighted badges and virtual goods [5] as elements that could boost motivation and reward good performance. For instance, participants mentioned collecting stars or virtual items, such as virtual food, as incentives for performing well.

Visual progress bars [4] and performance tables [5] were suggested to visualize medical and tactical decision-making progress while interacting with the VPs. For instance, the participants suggested that virtual progress bars with color could depict correct, suboptimal, and wrong decision-making. In addition, participants suggested that branched VPs could unlock content [5] and allow them to interact through avatars [5] with different pathways depending on their decisions, enabling them to explore the consequences of their actions. Content unlocking refers to access to new content or pathways based on the learner's performance or decisions. This mechanism could motivate participants to perform actions correctly to unlock more content.

Overall, participants appreciated direct feedback after making a wrong decision. However, they suggested that feedback should provide hints [5] of the correct answer rather than giving it away directly to challenge [4] the participants and maintain their interest:

> *...Progress bars would be good to see where you are, for example, in your way. One can have one progress bar with yellow, green, and red for tactical decisions, and one can have another one yellow, green, and red for medical decisions.* [Participant 4]

> *...Say, for example, if you get a certain point, you would then be able to unlock something...You can unlock something for the avatar. It could be anything, it doesn't have to be anything that really matters, but if I succeed in this [task], then a can of Präriegryta [a stew traditionally served in the military] appears—Congratulations! You have now unlocked Präriegryta! A reward, which may have no impact whatsoever on my ability to solve upcoming tasks. I just think it would be fun, and it would make me wonder, what will happen next? The next thing could be: Here is a package of the Swedish Armed Forces' skin ointment, it's silly, but I think the human brain is wired to seek these kinds of rewards, and they don't have to have any real value...* [Participant 5]

Theme 2: Agency in Interaction With VPs

Agency in interaction with the VPs was discussed in terms of how participants navigate and interact with the VP environment. Participants highlighted the importance of having a sense of control over their actions and decisions within the VP environment. They mentioned that, in certain cases, they felt limited in their decision-making and movement within the environment.

Participants suggested that they would appreciate more freedom in decision-making and how they could move around in the environment. However, answering with just free-text answers was perceived as more challenging; participants suggested that such an option should be combined with listed decision alternatives, known as imposed choice [4]. Imposed choice refers to decisions made from a set of given options, which can help guide the learning process.

The participants also discussed the importance of narrative [4], which is the structured sequence of events and choices that shape the learner's experience. They believed that being able to navigate within the VP by making choices listed in the VPs, combined with an interactive free-text menu, would give them a feeling of freedom and help unfold the scenario.

In addition, the participants mentioned the importance of the agency provided by avatars. They believed that avatars could enhance the feeling of control over the environment, which could influence motivation positively:

> *Yes, like it's "high alert," okay? But there is always much more information available. How do we assess the threat level right now? Do we have an external barrier in the form of surveillance? Have we set up an alarm system? Is the terrain mined, or can they [the enemy] just walk in? I need to know that! Because I will behave differently based on this information...So, yes, the medical and tactical [scoring], it's the tactical I am talking about; that's where it's very difficult if I can't make entirely tactical decisions because the information isn't there. Or the options I would have liked are not there...* [Participant 1]

> *When it comes to graphics, I don't think it is so important in a way, I think that I value the autonomy of the player the most. For instance, I would like to be able to create something by myself if I can give a name to my character or give it my name.* [Participant 5]

> *[The interactive menu] gives you a bit more freedom if there is something else you would like to do, but the option is not there.* [Participant 9]

Theme 3: Realistic Tactical Experience

The significance of delivering authentic, realistic experiences through VPs was discussed by the study participants. They highlighted the importance of the tactical environment, which may influence their decision-making. This theme is closely related to theme 2 and how immersion in the virtual environment representation can support the users' sense of agency. Participants frequently mentioned the visual detection of potential threats or enemies within the VP simulation as an example of realism. Multimedia elements, such as sound, pictures, and video, can enhance the feeling of presence and improve the perception of the virtual environment. Sensation [4], which involves visual and auditory stimuli, was highlighted as crucial for improving the experience. This includes virtual worlds, virtual reality, or augmented reality to engage the senses.

Participants noted that the tactical environment is constantly changing and unpredictable, often intersecting with medical decisions. This interplay makes military medical professionals distinct from civilian ones. Therefore, participants implied that the characteristics of the tactical environment should be highly visible and emphasized in the virtual environment. They suggested using multimedia elements to enhance realism and aid in understanding the tactical environment. Randomness as a game element [4] refers to the unpredictable elements introduced into the simulation, which can create a more realistic and challenging environment. This can simulate the unpredictability of real-life scenarios, making the experience more immersive and engaging. Realistic cases were perceived to enhance engagement and influence the sense of agency in the VP interaction:

> *I know what I would do, and I know what the game [VPs] would like me to do. The more the war environment and the situation is clarified, the tighter these two become. For instance [in the last case], I didn't know if they were still shooting, and that's why I decided to run into the building. If I knew, if there were people screaming, if I knew something was happening...Now my intuition told me that the shooting had stopped, so I wanted to get into the building...It is important with sound; to be able to listen the sounds of a war, someone can hear things before they come.* [Participant 3]

> *One gets already a lot of information from the picture...[referring to the second VP]. Someone can embed videos for 5-10 seconds so one gets a better idea of the patient.* [Participant 7]

Theme 4: Confidence, "I Know That the Knowledge I Have Works"

Participants felt that high scoring could increase their confidence, encourage them to continue interacting with cases, and thus enhance learning. They reflected that their previous combat casualty care training was applicable when managing the VPs and that this was reflected in their scores. Alternatively, participants commented that they needed more practice if they did not achieve the highest scores.

Adjusting the difficulty of presented educational content based on the learners' performance scores is known as difficulty adaptation [5] of the designed cases. Participants suggested that if the tasks are too challenging, they may become discouraged by low scores. Conversely, very easy cases might not challenge learners enough. Participants thought the cases were relevant to the course content, quizzing knowledge they had received, with the first and second cases being medically easier than the third one. Participants suggested that feeling unable to solve cases due to high difficulty levels could negatively influence their motivation:

> *I learned [from the VPs] that the knowledge I got from the course works [after looking at his final score].* [Participant 5]

> *Yes, it's good [the scoring]! And it is also good to get a patient that isn't possible to save. It's good to know...that you don't [always] have a chance...that you are done. But the percentage of such cases cannot be too high, because then you will lose motivation...It's the same thing when you talk about game development: if you have too much, too difficult, then people quit. They will lose interest; you have to have some "wins" too.* [Participant 4]

Theme 5: Social Influence on Motivation

Participants mentioned that scoring could be fun and enjoyable, particularly that higher scores increased their enjoyment of interacting with the VPs. They believed that social elements, such as group competition, could be motivating. Competition [4] involves encouraging learners to compete with each other. Leaderboards [5] are rankings that display the top scores of participants, motivating learners through visible performance comparisons. Participants suggested that displaying the top 3 players' scores to everyone could act as a goal for others to achieve. However, they also thought that low scores should not be visible, as this could embarrass individuals and cause demotivation. Social pressure [4] refers to the influence from peers or the community that can either motivate or demotivate learners. In this context, participants mentioned that visible scores and rankings could exert social pressure, influencing their motivation positively or negatively. While high scores and top rankings can be motivating, low scores can create negative feelings and reduce motivation:

> *I don't think that one should have everyone's score in a list, but maybe one could have a list with the ones having the highest score, let's say Gustafsson, that has 77 points. Then I want to beat Gustafsson and get a higher score and I will keep on trying. But then on the other hand, the one who is lower on the list, if I had 26 points for example out of 100, then I wouldn't like it to be visible. Then everyone would see that I have difficulties [with the game] and then I would lose motivation...I think that it would be good if the list just showed the top 3 or 5.* [Participant 5]

Theme 6: Personalized Learning

Participants highlighted the importance of personalized learning, where feedback and progression are tailored to the individual's performance. Scoring can provide personalized feedback on the number of correct decisions made during interactions with VPs. This feedback helps learners understand their current knowledge level and identify areas needing more practice.

Progression [4] refers to tracking and visualizing a learner's development and achievements over time. Participants suggested that users could log in and interact with VPs, with the system tracking their progression. This tracking allows users to see their current knowledge level and identify areas needing more practice. Participants also suggested receiving personalized material to practice based on previous performance. This feature aligns with the game elements of scoring and progression, where specific feedback on progress influences participants' confidence levels and personalized learning paths.

Furthermore, the system could enable users to repeat specific decisions after receiving negative feedback, promoting mastery in learning by allowing users to retry and improve their scores. Renovation [4,31] allows users to face the consequences of their actions and try again once their previous actions have failed. The participants perceived positively the fact that they could have a second chance to make a new decision, after failing with the first one, because it enabled them to realize wrong decisions and improve them.

Finally, participants suggested that they would be willing to practice decision-making with a library of VPs that offer a mix of varying levels of difficulty, allowing for continuous and adaptive learning experiences:

> *I learned [from the VPs] that I need to practice more...One can complete the VP and get some form of assessment about the performance and then receive tasks to do.* [Participant 1]

> *The scoring could be a way to see how well you are doing and someone could perhaps log in and see next time maybe that there are new things and tasks to do. The scoring can provide information about whether one has improved or is at the same level, so that I think it can be a good thing and that one might get a personal login or something so that one can follow the statistics on.* [Participant 9]

> *It is very good that one is given the option to try things again after one gets to know what the problem is. Because sometimes one has been making the wrong decision and one continues in the same direction [without noticing the mistake]. Here one knows, ok this is wrong, I am going to try again and make another decision…. I think it is good to get feedback immediately after making a mistake, and then provide them again with different decision alternatives. But it should not give away the right answer, like stopping the bleeding, directly, because then it becomes too easy.* [Participant 11]

Discussion

Principal Findings

This study investigated how to design VPs with game elements for Swedish military medics. The interaction analysis of the think-aloud sessions, structured with the uPEA framework, and the thematic analysis of interviews provided insights about the participants' interactions and design preferences. The insights gained may encourage designing VPs with game elements, as well as integrating errors that may support military medics' reflections and decision-making.

Overall, the VPs were positively received and effectively highlighted problems, leading to reflection and discussion. In the first VP case, the medics were confused; they lacked awareness of possible threats posed by the tactical environment and injured themselves by rushing to help patients. A shift was noted where most medics became tactically more aware, resulting in improved tactical decisions in the second and third VP cases. It is likely that the medics realized the consequences of their wrong tactical decisions in the first VP and applied this experience in the second and third VP. This observation shows the potential of VPs as an educational aid to support learning of situational awareness. This pattern was not observed in medical decision-making due to the varied medical challenges in each case, preventing the transfer of insights from one VP to another.

While analyzing how the participants explained their mistakes, we observed that some participants showed reluctance to accept their mistakes, attributing errors to the limitations of the VPs, such as the absence of 3D pictures and audio. They emphasized the importance of authentic and realistic experiences, particularly in detecting potential threats. For instance, being able to visually detect potential threats and enemies in the VPs was frequently mentioned as an example of "realism." While the 2D pictures and lack of sound do limit the presentation of the environment, this expectation may diverge from the reality of actual military operations, where threats often may remain concealed or unexpected. This raises the question about the role of realism in educational simulations and how fidelity to real-world conditions is realized in simulations [32].

Realism, although highlighted as important by the participants, is not the overall goal of simulation. Using various techniques that enable reflection to support decision-making may be more important. A study by Massoth et al [33] observed that high-fidelity simulation led to equal or even worse performance as compared with low-fidelity simulation. A previous systematic review noted that "presenting realistic patient scenarios with a great degree of freedom cannot be an excuse for neglecting guidance in relation to learning objectives" [17]. In some cases, "departing from realism" can enhance the quality and effectiveness of training by allowing participants to have a second chance, repeat scenarios after debriefings, or slow down the deterioration of a patient to provide time for reflection [34].

Although feedback and prompts for corrective actions may reduce the "realism" of field conditions—where consequences of wrong decision-making may not come with a possibility of correction—it was perceived positively by the participants. Feedback helped them recognize poor decision-making and personalized the learning experience. Unlike dialogue-oriented VPs, which primarily simulate clinical encounters through scripted interactions, our study's VPs enabled reflection by allowing participants to analyze their decisions and learn from their mistakes.

Purposefully selected multimedia elements, such as pictures, audio, and videos, can enhance the perceived realism and participants' understanding of the tactical environment. In our study, we used the generative AI application DALL-E [24] to generate realistic patient images in various simulated contexts. Such educational use of the new generation of artificial intelligence tools in VP design has been suggested also elsewhere [35] and seems to be a viable solution to increase realism, situation awareness of the simulation, and learner emotional engagement. Emphasizing the characteristics of the military environment in simulations was perceived as important in our study, as it allows participants to practice dealing with consequences that might be challenging to face in a live simulation.

Some of the participants mentioned feeling constrained while navigating in the VPs because of the limited number of listed decision alternatives they had to choose from. An interactive menu, where one can type in keywords, or a complementary free-text box could enable participants' wish for agency in decision-making, combined with several listed decision alternatives. In addition, participants mentioned that they value the agency in the interactions with the VPs, for instance, to be able to navigate freely in a virtual environment. Avatars were discussed as a positive feature to enhance motivation and the feeling of controlling the environment.

VPs effectively highlighted errors by showing the negative consequences of suboptimal choices, scoring performance, and providing textual feedback after decisions. What initially seemed like a limitation—such as the lack of sound and videos, which might lead to an incomplete understanding of the tactical environment—actually promoted participants' imagination and creativity. Many participants progressed from recognizing errors to reflecting on explanations and considering alternative solutions. Similarly, a previous study found that facilitators helped learners shift from focusing on explanations to generating alternative solutions by dramatizing scenarios [26].

In our study, the think-aloud sessions supported this reflection. To more effectively raise users' levels of reflection, VPs could be designed to prompt learners to consider why a mistake was made and think about next steps before presenting optional paths. Reflective practice and feedback, crucial for developing expertise, have been emphasized in several studies and align well with our findings [36,37]. VPs allow users to make mistakes and receive feedback, providing valuable opportunities for reflection on decision-making. Studies suggest that feedback in VPs can positively influence learning outcomes [17,38].

The participants also argued for various game elements that they believed would enhance the learning experience by making it more enjoyable, increasing their confidence, and providing reassurance that they were making the right decisions. These elements included scoring, badges, virtual goods, progress bars, performance tables, content unlocking, hints, challenge, control, imposed choice, narrative, avatars, sensation, randomness, difficulty adaptation, competition, leaderboards, social pressure, progression, and renovation.

Feedback was suggested to be realized in the way of providing hints for the correct answer and not giving away the right answer directly. Scoring and feedback can add a personalized dimension to learning; participants suggested that they would like to have access to several short cases to be able to log in to the VP platform and receive challenges in the form of new cases to solve. Personalized feedback would allow them to train competencies based on their performance, while it would allow them to monitor their progress and understand which competencies they need to train and improve. Visual progress bars depicting progression and achievements in medical and tactical decision-making were suggested as an alternative to visualize scoring by the participant.

Scoring of the users' performance was viewed as an element that can motivate users to keep on trying to improve. Although all participants agreed that scoring is a useful feature, many of the participants had not noticed the scores in the VP cases of this study as the scores were not placed in a prominent place on the screen. Scoring has been previously used in studies in various forms, to inform the participants about their progress. In a previous study, scoring was presented in the form of a trauma score [39], tracking the management of the trauma patient. Achatz et al [40] used a similar concept, presenting scores in the form of "health points" to depict how well the patient responded to the decision-making and events taking place in the game. Participants suggested that obtaining good scores makes the learning experience enjoyable and adds confidence, as it indicates that the knowledge they acquired during their training is useful. Similarly, previous studies have demonstrated that scoring-based game mechanics effectively motivated players and enhanced their learning engagement in a Swedish emergency department game [41].

According to the study's participants, displaying the third to fifth top scores to everyone can motivate learners to compete with each other, reflecting the concept of leaderboards. However, participants also noted that displaying less successful performance in the form of low scores could undermine the learners' reputation and lead to negative feelings. Social aspects and gaming preferences may vary across different cultures and professions. Some participants suggested integrating competition into VPs to increase motivation. However, other studies have reported that stakeholders in emergency medicine believe competition is not aligned with the culture of health care professionals [41]. Future studies should therefore carefully consider the target end users when designing and implementing competition as a game element.

Participants suggested that future efforts could focus on developing a VP repository with short cases and supporting the creation of VPs tailored to different levels, roles, and experiences of Swedish Home Guard members. Such VP repositories already exist for general medicine [42], and the idea could be transferred to military medicine education. This approach could further explore the potential of accelerating the design of VPs. In addition, future research could examine whether the design principles and game elements are applicable to other military medicine professionals and across different cultural and operational contexts.

According to the literature, serious games may improve performance in life-saving interventions on the battlefield [43,44]. Future research should investigate how gamification and specific game elements can enhance learning outcomes and the long-term retention of skills and knowledge acquired through gamified VP interactions.

We acknowledge that the context and culture of Swedish military medics may not fully generalize to other military or medical training environments. Moreover, the number of participants was limited (14) and therefore caution should be taken when considering how findings may transfer to other contexts. However, it is important to note that this study is qualitative in nature, where the depth and richness of the data are prioritized over the sample size.

While the uPEA framework might suggest that mistakes were the primary focus of the interaction analysis, we used it to understand how participants became more creative and reconsidered decisions when reflecting on mistakes. These insights could inform future VP designs.

The reflexive thematic analysis followed the 6-step approach suggested by Braun and Clarke [30]. While this is an established and flexible approach, it is highly interpretive, encouraging the experiences of the researchers to influence and inform the themes. Although this subjectivity can affect the consistency and replicability of the findings, it also serves as a strength in the context of this study. Our experience and familiarity with gamification strategies allowed coding game elements and informing the developed themes within the semistructured interviews. This capability was crucial, given that the participants were not well familiar with gamification terminology and thus unable to explicitly name these elements themselves. Through the reflexive thematic analysis, our study draws out implicit references to game elements and interprets discussions around participant experiences that are aligned with gamification concepts. This approach leverages the inherent subjectivity of our analysis to produce deeper insights that might otherwise be overlooked by the use of other methodologies.

Conclusions

Gamification has the potential to enhance VP design and offer flexible learning opportunities to support the education and training of Swedish military medics. The uPEA framework provides a useful lens for understanding decision-making during interactions with VPs, helping to inform how mistakes followed by feedback in VPs can promote reflection and creativity.

Game elements may effectively leverage the potential of VPs by fostering motivation, enabling personalized learning, supporting agency, and enhancing the confidence of military medics. Emphasis may be placed on representing austere aspects and consequences of the tactical environment. Incorporating opportunities to make errors followed by direct, constructive feedback mechanisms in the VPs may further enhance learning outcomes, encouraging critical reflection and continuous improvement.

Acknowledgments

The authors would like to express their gratitude to the participants of the TCCC-CLS course for their invaluable contributions during the think-aloud sessions and interviews. Special thanks are extended to Charlotte Silén and Karl Hybinette for their efforts in testing the virtual patient cases and providing essential feedback. The authors also wish to acknowledge the support of the Home Guard of Sweden (Hemvärnet) in facilitating this study. They extend their sincere appreciation to Lukas Berg, a certified instructor in combat casualty care, for his valuable role in developing the content of the VPs. In addition, they are grateful to Martin Olsson, a course leader in the Home Guard, for his pivotal contributions to the development of the VPs and for granting us access to the course participants. This study is financially supported by the Swedish Armed Forces. The ToC image was generated with the assistance of artificial intelligence and is one of the images used within the virtual patients. The authors have reviewed and approved the final output.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Visualization of virtual patient (VP) VP 1, VP 2 and VP 3 structure (Figures S1-S3); visualization of learners' interaction with decision nodes in VP 1, V2, and VP3 (Figures S4-S6); tactical key decision in VP 1 and decision alternatives (Table S1), and semistructured interviews guide (Table S2).

[PDF File (Adobe PDF File), 439 KB - games_v12i1e63390_app1.pdf]

References

1. Stathakarou N, Sonesson L, Lundberg L, Boffard KD, Kononowicz AA, Karlgren K. Teams managing civilian and military complex trauma: what are the competencies required in austere environments and the potential of simulation technology to address them? Health Informatics J 2021;27(4):14604582211052253 [FREE Full text] [doi: 10.1177/14604582211052253] [Medline: 34821149]
2. Ellaway R, Poulton T, Fors U, McGee JB, Albright S. Building a virtual patient commons. Med Teach 2008;30(2):170-174. [doi: 10.1080/01421590701874074] [Medline: 18464142]
3. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
4. Toda AM, Klock AC, Oliveira W, Palomino PT, Rodrigues L, Shi L, et al. Analysing gamification elements in educational environments using an existing gamification taxonomy. Smart Learn Environ 2019 Dec 04;6:16. [doi: 10.1186/s40561-019-0106-1]

5. Maheu-Cadotte MA, Cossette S, Dubé V, Fontaine G, Mailhot T, Lavoie P, et al. Effectiveness of serious games and impact of design elements on engagement and educational outcomes in healthcare professionals and students: a systematic review and meta-analysis protocol. BMJ Open 2018 Mar 16;8(3):e019871 [FREE Full text] [doi: 10.1136/bmjopen-2017-019871] [Medline: 29549206]
6. van Gaalen AE, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma AD, Georgiadis JR. Gamification of health professions education: a systematic review. Adv Health Sci Educ Theory Pract 2021 May;26(2):683-711 [FREE Full text] [doi: 10.1007/s10459-020-10000-3] [Medline: 33128662]
7. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
8. Lampropoulos G, Kinshuk. Virtual reality and gamification in education: a systematic review. Educ Technol Res Dev 2024 Mar 19;72:1691-1785. [doi: 10.1007/s11423-024-10351-3]
9. Stathakarou N, Kononowicz AA, Swain C, Karlgren K. Game elements in the design of simulations in military trauma management training: protocol for a systematic review. JMIR Res Protoc 2023 Sep 08;12:e45969 [FREE Full text] [doi: 10.2196/45969] [Medline: 37682596]
10. Hemvärnet. Försvarsmakten. URL: https://www.forsvarsmakten.se/sv/var-verksamhet/det-har-gor-forsvarsmakten/hemvarnet/ [accessed 2024-10-07]
11. Andersson SO, Lundberg L, Jonsson A, Tingström P, Abrandt Dahlgren M. Doctors' and nurses' perceptions of military pre-hospital emergency care - when training becomes reality. Int Emerg Nurs 2017 May;32:70-77. [doi: 10.1016/j.ienj.2017.01.001] [Medline: 28228341]
12. Ignatiadis IA, Mavrogenis AF, Igoumenou VG, Polyzois VD, Tsiampa VA, Arapoglou DK, et al. Gunshot and blast injuries of the extremities: a review of 45 cases. Eur J Orthop Surg Traumatol 2019 Feb;29(2):295-305. [doi: 10.1007/s00590-018-2350-4] [Medline: 30498906]
13. Penn-Barwell JG, Brown KV, Fries CA. High velocity gunshot injuries to the extremities: management on and off the battlefield. Curr Rev Musculoskelet Med 2015 Sep;8(3):312-317 [FREE Full text] [doi: 10.1007/s12178-015-9289-4] [Medline: 26108862]
14. Andersson SO, Lundberg L, Jonsson A, Tingström P, Dahlgren MA. Interaction, action, and reflection: how medics learn medical care in the Swedish armed forces. Mil Med 2013 Aug;178(8):861-866. [doi: 10.7205/milmed-d-13-00048]
15. Butler FK, Bennett B, Wedmore CI. Tactical combat casualty care and wilderness medicine: advancing trauma care in austere environments. Emerg Med Clin North Am 2017 May;35(2):391-407. [doi: 10.1016/j.emc.2016.12.005] [Medline: 28411934]
16. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development. Upper Saddle River, NJ: Prentice Hall; 1984.
17. Kononowicz AA, Woodham LA, Edelbring S, Stathakarou N, Davies D, Saxena N, et al. Virtual patient simulations in health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Jul 02;21(7):e14676 [FREE Full text] [doi: 10.2196/14676] [Medline: 31267981]
18. Posel N, Mcgee JB, Fleiszer DM. Twelve tips to support the development of clinical reasoning skills using virtual patient cases. Med Teach 2015;37(9):813-818. [doi: 10.3109/0142159X.2014.993951] [Medline: 25523009]
19. Hege I, Kononowicz AA, Tolks D, Edelbring S, Kuehlmeyer K. A qualitative analysis of virtual patient descriptions in healthcare education based on a systematic literature review. BMC Med Educ 2016 May 13;16:146 [FREE Full text] [doi: 10.1186/s12909-016-0655-8] [Medline: 27177766]
20. Deployed Medicine. URL: https://deployedmedicine.com/ [accessed 2024-10-07]
21. Butler FKJ. Tactical combat casualty care: beginnings. Wilderness Environ Med 2017 Jun;28(2S):S12-S17. [doi: 10.1016/j.wem.2016.12.004] [Medline: 28284483]
22. Patton MQ. Qualitative Research & Evaluation Methods: Integrating Theory and Practice. Thousand Oaks, CA: SAGE Publications; 2014.
23. OpenLabyrinth. URL: http://openlabyrinth.ca [accessed 2024-10-07]
24. DALL·E 2. OpenAI. URL: https://openai.com/index/dall-e-2/ [accessed 2024-10-07]
25. Huwendiek S, De leng BA, Zary N, Fischer MR, Ruiz JG, Ellaway R. Towards a typology of virtual patients. Med Teach 2009 Aug;31(8):743-748. [doi: 10.1080/01421590903124708] [Medline: 19811212]
26. van Someren MW, Barnard YF, Sandberg J. The Think Aloud Method A Practical Guide to Modelling Cognitive Processes · Volume 12. Cambridge, MA: Academic Press; 1994.
27. Jordan B, Henderson A. Interaction analysis: foundations and practice. J Learn Sci 1995;4(1):39-103. [doi: 10.1207/s15327809jls0401_2]
28. Derry SJ, Pea RD, Barron B, Engle RA, Erickson F, Goldman R, et al. Conducting video research in the learning sciences: guidance on selection, analysis, technology, and ethics. J Learn Sci 2010;19(1):3-53. [doi: 10.1080/10508400903452884]
29. Karlgren K, Larsson F, Dahlström A. Eye-opening facilitator behaviours: an interaction analysis of facilitator behaviours that advance debriefings. BMJ Simul Technol Enhanc Learn 2020 Jul;6(4):220-228 [FREE Full text] [doi: 10.1136/bmjstel-2018-000374] [Medline: 32832101]

30. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: SAGE Publications; 2021.
31. Lee J, Hammer J. Gamification in education: what, how, why bother? Acad Exchange Q 2011;15(2):1-5 [FREE Full text]
32. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. Acad Med 2014 Mar;89(3):387-392 [FREE Full text] [doi: 10.1097/ACM.0000000000000130] [Medline: 24448038]
33. Massoth C, Röder H, Ohlenburg H, Hessler M, Zarbock A, Pöpping DM, et al. High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students. BMC Med Educ 2019 Jan 21;19(1):29 [FREE Full text] [doi: 10.1186/s12909-019-1464-7] [Medline: 30665397]
34. Dieckmann P, Gaba D, Rall M. Deepening the theoretical foundations of patient simulation as social practice. Simul Healthc 2007;2(3):183-193. [doi: 10.1097/SIH.0b013e3180f637f5] [Medline: 19088622]
35. Bakkum MJ, Hartjes MG, Piët JD, Donker EM, Likic R, Sanz E, et al. Using artificial intelligence to create diverse and inclusive medical case vignettes for education. Br J Clin Pharmacol 2024 Mar;90(3):640-648. [doi: 10.1111/bcp.15977] [Medline: 38016816]
36. Prakash S, Sladek RM, Schuwirth L. Interventions to improve diagnostic decision making: a systematic review and meta-analysis on reflective strategies. Med Teach 2019 May;41(5):517-524. [doi: 10.1080/0142159X.2018.1497786] [Medline: 30244625]
37. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. Psychol Rev 1993;100(3):363-406. [doi: 10.1037//0033-295x.100.3.363]
38. Huwendiek S, Reichert F, Bosse HM, de Leng BA, van der Vleuten CP, Haag M, et al. Design principles for virtual patients: a focus group study among students. Med Educ 2009 Jun;43(6):580-588. [doi: 10.1111/j.1365-2923.2009.03369.x] [Medline: 19493183]
39. Chi DM, Kokkevis E, Ogunyemi O, Bindiganavale R, Hollick MJ, Clarke JR, et al. Simulated casualties and medics for emergency training. Stud Health Technol Inform 1997;39:486-494. [Medline: 10168943]
40. Achatz G, Friemert B, Trentzsch H, Hofmann M, Blätzinger M, Hinz-Bauer S, et al. Terror and disaster surgical care: training experienced trauma surgeons in decision making for a MASCAL situation with a tabletop simulation game. Eur J Trauma Emerg Surg 2020 Aug;46(4):717-724. [doi: 10.1007/s00068-020-01441-x] [Medline: 32725275]
41. Zhang C, Härenstam KP, Meijer S, Darwich AS. Serious gaming of logistics management in pediatric emergency medicine. Int J Serious Games 2020 Mar 02;7(1):47-77 [FREE Full text] [doi: 10.17083/ijsg.v7i1.334]
42. Mayer A, Da Silva Domingues V, Hege I, Kononowicz AA, Larrosa M, Martínez-Jarreta B, et al. Planning a collection of virtual patients to train clinical reasoning: a blueprint representative of the European population. Int J Environ Res Public Health 2022 May 19;19(10):6175 [FREE Full text] [doi: 10.3390/ijerph19106175] [Medline: 35627711]
43. Planchon J, Vacher A, Comblet J, Rabatel E, Darses F, Mignon A, et al. Serious game training improves performance in combat life-saving interventions. Injury 2018 Jan;49(1):86-92. [doi: 10.1016/j.injury.2017.10.025] [Medline: 29042033]
44. de Lesquen H, Paris R, Fournier M, Cotte J, Vacher A, Schlienger D, et al. Toward a serious game to help future military doctors face mass casualty incidents. J Spec Oper Med 2023 Jun 23;23(2):88-93. [doi: 10.55460/IJCP-BLY6] [Medline: 37126778]

Abbreviations

MARCH: massive hemorrhage, airway, respirations, circulation, head injury, and hypothermia
TCCC: Tactical Combat Casualty Care
TCCC-CLS: Tactical Combat Casualty Care Combat Lifesavers course
uPEA: unawareness, problem identification, explanation, and alternative strategies or solutions
VP: virtual patient

Edited by T Leung, N Ahmadpour; submitted 18.06.24; peer-reviewed by MT Chaudhry, L Woodham; comments to author 26.08.24; revised version received 10.09.24; accepted 22.09.24; published 22.10.24.

Please cite as:
Stathakarou N, Kononowicz AA, Mattsson E, Karlgren K
Gamification in the Design of Virtual Patients for Swedish Military Medics to Support Trauma Training: Interaction Analysis and Semistructured Interview Study
JMIR Serious Games 2024;12:e63390
URL: https://games.jmir.org/2024/1/e63390
doi:10.2196/63390
PMID:39436692

Original Paper

Empowering Social Growth Through Virtual Reality–Based Intervention for Children With Attention-Deficit/Hyperactivity Disorder: 3-Arm Randomized Controlled Trial

Ka Po Wong[1], PhD; Bohan Zhang[2], MSc; Cynthia Yuen Yi Lai[3], PhD; Yao Jie Xie[4], PhD; Yan Li[4], PhD; Chen Li[1,5], PhD; Jing Qin[2], PhD

[1]Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)
[2]Centre for Smart Health, School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)
[3]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)
[4]School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)
[5]Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Ka Po Wong, PhD
Department of Applied Social Sciences
The Hong Kong Polytechnic University
Room GH329, 3rd Floor, Block G
11 Yuk Choi Rd, Hung Hom
Hong Kong, 000000
China (Hong Kong)
Phone: 852 64860214
Email: portia.wong@polyu.edu.hk

Abstract

Background: Attention-deficit/hyperactivity disorder (ADHD) usually begins in childhood and is often accompanied by impairments in social functioning. Virtual reality (VR) has emerged as an adjunctive tool to embed in social skills training to enhance the social skills of children with ADHD, but its effectiveness requires further investigation.

Objective: This study aims to enhance the social skills of children with ADHD by examining the feasibility and effectiveness of VR-based training in comparison to traditional social skills training.

Methods: A 3-arm randomized controlled trial was conducted with 90 children with ADHD aged 6-12 years. Participants were randomly assigned to 3 weeks of 12-session VR-based social skills training, traditional social skills training, or a waitlist control group of equivalent duration. Outcome measures included assessments by a clinical psychologist who was blinded to group assignments, the Social Skills Improvement System Rating Scale, the Behavior Rating Inventory of Executive Function, and the Simulator Sickness Questionnaire, conducted at baseline and after the intervention.

Results: The preliminary results support the feasibility and acceptability of VR training for children with ADHD aged 6-12 years. Analysis showed that the VR and traditional social skills training groups experienced a statistically significant improvement in the clinical psychologist assessment of social skills and parent-rated self-control, initiative, and emotional control after the intervention compared with baseline. The VR group performed significantly better than the traditional social skills group on social skills assessed by clinical psychologists ($F_{2,85}$=76.77; P<.001) and on parent-rated self-control ($F_{2,85}$=18.77; P<.001), initiative ($F_{2,85}$=11.93; P<.001), and emotional control ($F_{2,85}$=17.27; P<.001). No significant between-group differences were found for parent-rated cooperation and inhibition (all P>.05).

Conclusions: The findings provide preliminary evidence supporting the feasibility and superior effectiveness of VR-based social skills training compared to traditional approaches for enhancing social skills and related executive functions in children with ADHD. These results suggest that VR may be a valuable tool to embed within social skills interventions for this population. Further research is warranted to explore the long-term impacts and generalizability of these benefits.

Trial Registration: ClinicalTrials.gov NCT05778526; https://clinicaltrials.gov/study/NCT05778526

International Registered Report Identifier (IRRID): RR2-10.2196/48208

(JMIR Serious Games 2024;12:e58963) doi:10.2196/58963

KEYWORDS

attention deficit and hyperactivity disorder; virtual reality; social skills; social skills training; emotional control; social growth; digital world; social learning theory

Introduction

Overview

Attention-deficit/hyperactivity disorder (ADHD) is one of the children's most common neurobehavioral disorders [1], with a global prevalence of 5.3% [2] and a prevalence of 7.6% among children aged 3-12 years [3]. Inattention, impulsivity, and hyperactivity are the 3 main categories of ADHD symptoms, which are associated with nontypical cognitive and behavioral functioning due to their influence on the regulation of emotions and brain networks. Persistent patterns of inattention, impulsivity, and hyperactivity that interfere with daily functioning are considered ADHD [4]. Meanwhile, these symptoms may contribute to social difficulties and peer problems due to a deficit in social executive function skills [5].

Children with ADHD often experience specific social skills difficulties that can affect their interactions and relationships with others [6]. These children may struggle to think before acting, leading to impulsive behaviors that can disrupt conversations or be seen as socially inappropriate [7]. They may have difficulty considering the consequences of their actions, leading to difficulty in maintaining self-control during social interactions. They also have trouble sustaining attention during conversations and missing important details or cues [8]. They often move excessively, are restless, or have difficulty staying in a particular place, making it difficult to actively participate in group activities, take turns, or engage in cooperative play [9]. Furthermore, difficulty regulating emotions can affect social interactions in children with ADHD [10], leading to outbursts of anger, frustration, or impatience. These emotional challenges can strain relationships with peers and make it difficult to respond effectively to social situations. This persistent social impairment can lead to peer rejection and exacerbate academic failure, depression, and anxiety [11,12]; thus, addressing social impairment during childhood is critical.

Social skills training is one of the ubiquitous approaches to ameliorating social behaviors in people with ADHD. The training usually involves a combination of didactic instruction and role playing activities [13]. Effective social skills interventions offer strategies and support to help children with ADHD improve their social competence, navigate social situations better, and foster positive relationships [14]. However, this method may be restricted by time and space, and imagining actual situations during training is difficult for children with ADHD.

Virtual reality (VR)–based interventions are considered an effective alternative to traditional social skills training, providing the immersion of a digital environment and engaging users in the learning process [15]. VR can produce various scenarios that are not easily and safely operated in contrast to equivalent real-life training programs [16,17]. Typical VR-based interventions for ADHD use immersive digital environments that simulate real-world social and academic situations. For example, these interventions may place the child in a classroom setting with various auditory (eg, classmates chatting and teacher lecturing) and visual (eg, posters on walls and other students moving around) distractions [18,19]. By requiring the child to maintain focus and inhibit responses to these distractions within the VR environment, the intervention aims to improve their inhibitory control and attention regulation skills [20]. The VR setting allows the child to practice these skills in a safe, controlled space without the consequences of real-world mistakes.

Additionally, VR interventions may present the child with digital social scenarios, such as interacting with digital peers or authority figures. These scenarios allow the child to observe and practice appropriate social behaviors, like taking turns, making eye contact, and responding to social cues, which can be challenging for children with ADHD [21]. Immediate feedback and the ability to repeat scenarios within VR can help consolidate these social skills.

By incorporating these types of immersive VR environments, the interventions leverage the unique affordances of VR to target the core cognitive, behavioral, and social deficits associated with ADHD engagingly and safely. Social skills training using VR technology exists, and the efficiency of such applications has been demonstrated; however, beneficiaries are mainly those with autism spectrum disorders, schizophrenia, and intellectual disabilities [22-25]. A VR-based intervention to enhance the social interaction skills of children with ADHD is limited, and its effectiveness has not been empirically studied.

Theoretical Background

The VR-based social skills intervention is rooted in social learning theory, as developed by Bandura [26]. Social learning theory posits that individuals can acquire new behaviors, attitudes, and emotional responses by observing the actions and consequences experienced by others [27]. This observational learning process is a key mechanism by which children develop social competence and navigate complex interpersonal situations. For children with ADHD, deficits in social skills are often linked to challenges with attention, impulse control, and emotional regulation—factors that can hinder their ability to effectively observe, process, and learn from social cues and interactions [28]. The immersive nature of VR provides an ideal platform to address these challenges and leverage the principles of social learning theory [15].

Within the VR environment, children with ADHD can observe modeled social behaviors and their consequences in a safe, controlled setting. This allows them to focus on the relevant social cues and responses without the distractions or anxiety that may arise in real-world interactions [29]. The VR-based intervention can facilitate observational learning and cognitive processing of adaptive social behaviors by presenting diverse social scenarios and modeling appropriate social skills.

The VR-based intervention incorporates opportunities for active practice and role playing of these observed social skills. This aligns with Bandura's emphasis on the importance of engaging in enactive learning, where individuals have the chance to rehearse and refine new behaviors [26]. The immediate feedback and reinforcement provided within the VR environment can further support the acquisition and consolidation of social skills. Importantly, social learning theory also highlights the role of self-efficacy—an individual's belief in their ability to successfully execute a behavior [29]. By providing a supportive, scaffolded learning environment in VR, the intervention aims to enhance children's confidence and self-efficacy in navigating social situations, ultimately promoting the generalization of social skills to real-world contexts.

By incorporating social learning theory [26] into VR-based social skills training, an effective and engaging learning environment that aligns with the unique needs and challenges of children with ADHD can be provided. The VR-based approach leverages the power of observation, modeling, reinforcement, and cognitive processes to promote the acquisition and generalization of adaptive social skills, empowering children with ADHD to navigate social interactions with greater confidence and success.

Rationales of VR-Based Social Skills Training for ADHD

While traditional social skills training has been used to help children with ADHD, they may be limited by logistical constraints and difficulty simulating real-world social scenarios [13,14]. VR technology offers a promising alternative, as it can immerse participants in customizable social environments that are difficult to replicate in physical settings [15-17]. VR-based interventions have been shown to effectively target social skills deficits in populations such as autism spectrum disorder and schizophrenia [22-25].

However, research on the use of VR-based social skills training for children with ADHD is limited. A few existing studies have focused on using VR to address ADHD symptoms like inattention and impulsivity but have not specifically examined the impact on social skills [30]. Given the unique social challenges faced by this population and the potential benefits of VR-based training, there is a clear need for rigorous investigation into the feasibility and efficacy of VR-based social skills interventions for children with ADHD.

This study aims to address this gap by examining the feasibility and effectiveness of VR-based social skills training for children with ADHD, in comparison to a traditional social skills training approach and a waitlist control group. Drawing on social learning theory, the VR-based intervention is designed to leverage the immersive, engaging features of VR to enhance observational learning, skill practice, and self-efficacy in navigating social situations. Findings from this randomized controlled trial (RCT) will provide valuable insights into the potential of VR-based interventions to improve social competence in children with ADHD.

Methods

Study Design

In this study conducted at The Hong Kong Polytechnic University (ClinicalTrials.gov NCT05778526), a 3-arm RCT was implemented from November 2023 to February 2024. The participants were assigned randomly to 3 groups: the VR training group, the traditional social skills training group, and the waitlist control group, maintaining a 1:1:1 ratio. The VR training group received 12 sessions for 3 weeks, with each session lasting 20 minutes. Similarly, the traditional social skills training group also underwent 12 sessions in 3 weeks, with each session lasting 20 minutes. These intervention durations and frequencies align with the findings of a study by Willis et al [14], which reported intervention durations ranging from 2 to 12 weeks and weekly time commitments of 60-90 minutes. Therefore, the chosen intervention duration and frequency in our study are considered appropriate based on existing research. Participants randomized to the waitlist control group were offered VR training after 1 month when data were collected. The guardians of the participants were reminded to attend the training via phone in the first 4 sessions and then via WhatsApp (Meta Platforms) or WeChat (Tencent Holdings Ltd) in the fifth and eighth sessions, respectively. There were no follow-up reminders issued in the following 4 sessions. The protocol of the RCT was published [31]. The study followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Multimedia Appendix 1).

Ethical Considerations

This study received approval from the institutional review board at The Hong Kong Polytechnic University (HSEARS20221221003) and was conducted from November 2023 to February 2024. It adhered to the principles outlined in the Declaration of Helsinki, which included ensuring participant anonymity and obtaining informed consent from the guardians of the participants. Participation was voluntary. Informed consent was obtained through a clear process, ensuring that guardians fully understood the study's purpose, procedures, and potential risks. Participants were also given the option to withdraw from the study at any time without penalty. Data collected during the study were anonymized to protect participant confidentiality.

Recruitment

Participants were recruited from the children and youth community centers through extensive advertising, including posters and advertisements on social media. A web-based registration form was created for the interested guardian to register. Interested guardians who provided written informed consent were further evaluated for inclusion and exclusion criteria.

Eligibility criteria for participants were as follows: (1) aged between 6 and 12 years; (2) Chinese ethnicity; (3) residing in Hong Kong; (4) having received a diagnosis of ADHD by Child Assessment Service in Hong Kong or via private practice; (5) stable on pharmacological and psychological treatment for ADHD 8 weeks before baseline (determined by health care professionals based on medication data and behavioral observation); (6) no initiation or change of pharmacological treatment for ADHD during the intervention period; (7) ability to read Chinese and speak and listen to Cantonese by the child and by at least one of their legal guardian; and (8) willing to provide informed consent by the participants' legal guardians.

Exclusion criteria were as follows: (1) comorbid autism; (2) mental health disorders; (3) an estimated IQ lower than 85 (using the Wechsler Intelligence Scale for Children—Fourth Edition); (4) autism spectrum disorder (previously diagnosed by health care professionals), (5) comorbid acute psychiatric disorder (previously diagnosed by health care professionals); and (6) severe physical disability (eg, blindness, deafness) or learning disability (eg, dyslexia).

Procedure and Interventions

Potential participants' guardians completed a telephone prescreening. Following consent, the guardian showed the diagnosis of ADHD of the participants and completed the baseline assessment for the participants. The randomization was conducted using a computer-generator randomizer to generate the random allocation list. The randomization was undertaken by another research assistant not directly involved in the study. A number generated by the computer will be assigned to each eligible subject who will be randomly allocated to the 3 different groups by using the number.

VR Training Group

The VR training aimed to enhance the social interaction skills of children with ADHD. Participants in this group wore a head-mounted display and 2 controllers during the training. The intervention included three real-life scenarios, including (1) a classroom and playground, (2) a mass transit railway station and carriages, and (3) a supermarket and restaurant (Multimedia Appendix 2). These scenarios were purposefully designed to simulate real-life social situations, allowing participants to practice and refine their social skills in a controlled and supportive digital environment. Gamification elements were integrated into the immersive VR training, including dynamic simulations and real-time feedback to provide immediate correction of participants' behaviors.

Each scenario contained approximately 20 or more tasks for the participants to complete. Participants underwent 1 scenario in 1 session. The sequences of the scenarios used in each session were the same for all participants. Each session lasted for a maximum of 20 minutes to ensure the concentration of the participants and prevent any physical effects [32]. The duration was adjusted subject to the emotions of the participants during the training. During the intervention, a research assistant guided the participants in completing the tasks. Multimedia Appendix 3 shows the participants conducting the intervention.

Traditional Social Skills Training Group

Participants in this group were educated in social skills through traditional methods, including role-playing activities and didactic instruction. Four modules, including (1) how to introduce yourself and basic social skills, (2) how to listen to others, (3) how to share with others, and (4) learning to know how people feel and how to empathize, were covered in the training, which has been adopted in previous studies [33]. Each participant was taught by 1 instructor. The content of the training and the duration of each session were kept similar to the VR training group.

Waitlist Control Group

Participants in this group received no training and were not allowed to change or initiate their medical treatment during the study period.

Outcome Measures

Data were collected at 2 time intervals, including baseline (T1) and immediately after the last sessions (T2) to determine the feasibility and effectiveness of the RCT.

Primary Outcome: Social Skills

The social skills of the participants were evaluated by the guardians of the participants and a clinical psychologist. The guardians rated using the Social Skills Improvement System-Rating Scales, which has been validated in children with ADHD [34]. A total of 31 items were selected from the self-control, initiative, and cooperation subscales of the Social Skills Improvement System–Parents. The clinical psychologist used a modified version of Riggio basic social skills assessment [35] to evaluate the participants. This involved a 20-minute conversation in which the clinical psychologist observed and assessed their social skills. To minimize unblinding bias during the assessment, each participant will be identified by a case number.

Secondary Outcome

Executive Functioning

The executive functioning of the participants was measured by the subscale of inhibition (16 items) and emotional control (10 items) of the Behavior Rating Inventory of Executive Function-Parents, which is useful in Chinese children with ADHD [36]. The executive functioning of the participants was evaluated by the guardians of the participants.

Motion Sickness

The motion sickness or physical discomfort of subjects using the VR was measured by the Simulator Sickness Questionnaire [37], which was administered at T2 to the participants in the VR training group only.

Satisfaction

Satisfaction with training was measured by a 7-point Likert scale administered at T2 to the VR training group.

Feasibility and Acceptability Outcomes

Feasibility was assessed through participants' retention of the intervention. The attendance of the participants during the interventions was recorded, and absence from any training session was considered nonadherence.

Justification of Sample Size

A sample size of 20 subjects per group can attain at least 80% power and 90% confidence according to the study of Whitehead et al [38]. We intended to recruit 90 participants (30 per group), assuming a conservative attrition rate of 25% to 30% [19], to reliably determine these outcomes.

Statistical Analysis

All statistical analyses were performed in the SPSS (version 28, IBM Corp) software and were 2-sided with a level of significance of $P<.05$. The data analysis complied with the principles of intention-to-treat analysis. The subjects' demographic information was summarized into categorical variables with frequency and percentage. The change between T1 and T2 of each group was tested by a 2-tailed *t* test. Analyses of covariance (adjusted for possible confounding factors) were conducted to evaluate the within-group effects and between-group effects in terms of outcomes. Effect sizes were measured by Cohen *d* and 95% CI. To assess the improvement during the intervention period among the 3 groups, *F* tests were performed on primary and secondary outcome measures at T1 and T2.

Results

Participant Characteristics

Between November 2023 and February 2024, we enrolled 108 participants in our study. However, 10 participants were excluded before randomization due to not meeting the inclusion or exclusion criteria, and 8 declined to provide written consent. Consequently, 90 participants, along with their guardians, agreed to participate in the study. These participants and their guardians agreed to join the study and were randomly assigned to the VR training (n=30), traditional social skills training (n=30), or waitlist control group (n=30). Baseline characteristics were balanced among the 3 groups without statistically significant differences (Tables 1 and 2). The mean age for the VR training, traditional social skills training, and waitlist control groups were 8.63 (SD 1.90), 8.30 (SD 1.70), and 8.67 (SD 1.45) years, respectively. All participants attended mainstream schools. Table 1 presents the participants' baseline demographic characteristics and outcome measure scores, respectively.

Table 1. Baseline data for the VRT[a], TSST[b], and WLC[c] groups.

Variables	VRT (n=30)	TSST (n=30)	WLC (n=30)	*P* value[d]
Sex, n (%)				.60
Male	23 (77)	26 (87)	24 (80)	
Female	7 (23)	4 (13)	6 (20)	
Age (years), mean (SD)	8.63 (1.90)	8.30 (1.70)	8.67 (1.45)	.20
ADHD[e] subtype, n (%)				.29
Combined	19 (63)	24 (80)	26 (87)	
Inattention	7 (23)	4 (13)	3 (10)	
Hyperactivity or impulsivity	4 (13)	2 (7)	1 (3)	
Medication, n (%)				.73
Yes	18 (60)	17 (57)	15 (50)	
No	12 (40)	13 (43)	15 (50)	
Outcomes variables, mean (SD)				
Clinical psychological assessment of social skills	26.23 (6.40)	25.60 (7.16)	24.23 (6.27)	.49
Social Skills Improvement System–Parents				
Self-control	31.57 (4.10)	31.02 (3.79)	31.60 (5.00)	.84
Initiative	20.43 (4.91)	19.78 (4.30)	19.65 (4.08)	.77
Cooperation	12.07 (1.82)	11.00 (2.68)	11.98 (2.67)	.17
Behavior Rating Inventory of Executive Function–Parents				
Inhibition	31.20 (5.46)	32.20 (4.76)	33.05 (6.45)	.44
Emotional control	18.53 (4.19)	18.92 (3.35)	19.67 (5.42)	.60

[a]VRT: virtual reality training.

[b]TSST: traditional social skills training.

[c]WLC: waitlist control.

[d]Chi-square test or *F* test.

[e]ADHD: attention-deficit/hyperactivity disorder.

Table 2. Social skills and executive functioning according to the group before and after the intervention

Outcome variables	VRT[a] (n=30), mean (SD)		TSST[b] (n=30), mean (SD)		WLC[c] (n=30), mean (SD)		Within-group analysis, *P* value			Between-group analysis		Effect sizes, Cohen *d* (95% CI)	
	T1	T2	T1	T2	T1	T2	VRT	TSST	WLC	*F* test (*df*)	*P* value	VRT vs TSST	VRT vs WLC
Clinical psychological assessment of social skills[d]	26.23 (6.40)	38.20 (2.83)	25.60 (7.16)	30.63 (6.57)	24.23 (6.27)	24.50 (5.17)	<.001	<.001	.73	76.77 (2, 85)	<.001	1.50 (0.91-2.07)	3.29 (2.50-4.06)
Social Skills Improvement System–Parents[d]													
Self-control	31.57 (4.10)	33.33 (3.56)	31.02 (3.79)	31.93 (2.78)	31.60 (5.00)	31.50 (4.70)	<.001	0.003	.50	18.77 (2, 85)	<.001	0.44 (–0.08 to 0.95)	0.44 (–0.08 to 0.95)
Initiative	20.43 (4.91)	21.13 (3.80)	19.78 (4.30)	20.20 (4.02)	19.65 (4.08)	19.60 (3.92)	.01	.001	.59	11.93 (2, 85)	<.001	0.24 (–0.27 to 0.74)	0.40 (–0.11 to 0.91)
Cooperation	12.07 (1.82)	12.17 (1.58)	11.00 (2.68)	11.37 (2.16)	11.98 (2.67)	12.03 (2.51)	.33	.05	.61	0.74 (2, 85)	.48	0.42 (–0.09 to 0.93)	0.07 (–0.44 to 0.57)
Behavior Rating Inventory of Executive Function–Parents[e]													
Inhibition	31.20 (5.46)	32.00 (4.45)	32.20 (4.76)	33.00 (4.05)	33.05 (6.45)	34.97 (6.75)	.02	.003	<.001	7.10 (2, 85)	.001[f]	–0.24 (–0.74 to 0.27)	–0.52 (–1.03 to –0.00)
Emotional control	18.53 (4.19)	16.83 (3.08)	18.92 (3.35)	17.98 (2.68)	19.67 (5.42)	19.63 (4.91)	<.001	<.001	.89	17.27 (2, 85)	<.001	–0.40 (–0.91 to 0.11)	–0.69 (–1.20 to –0.16)

[a]VRT: virtual reality training.

[b]TSST: traditional social skills training

[c]WLC: waitlist control group.

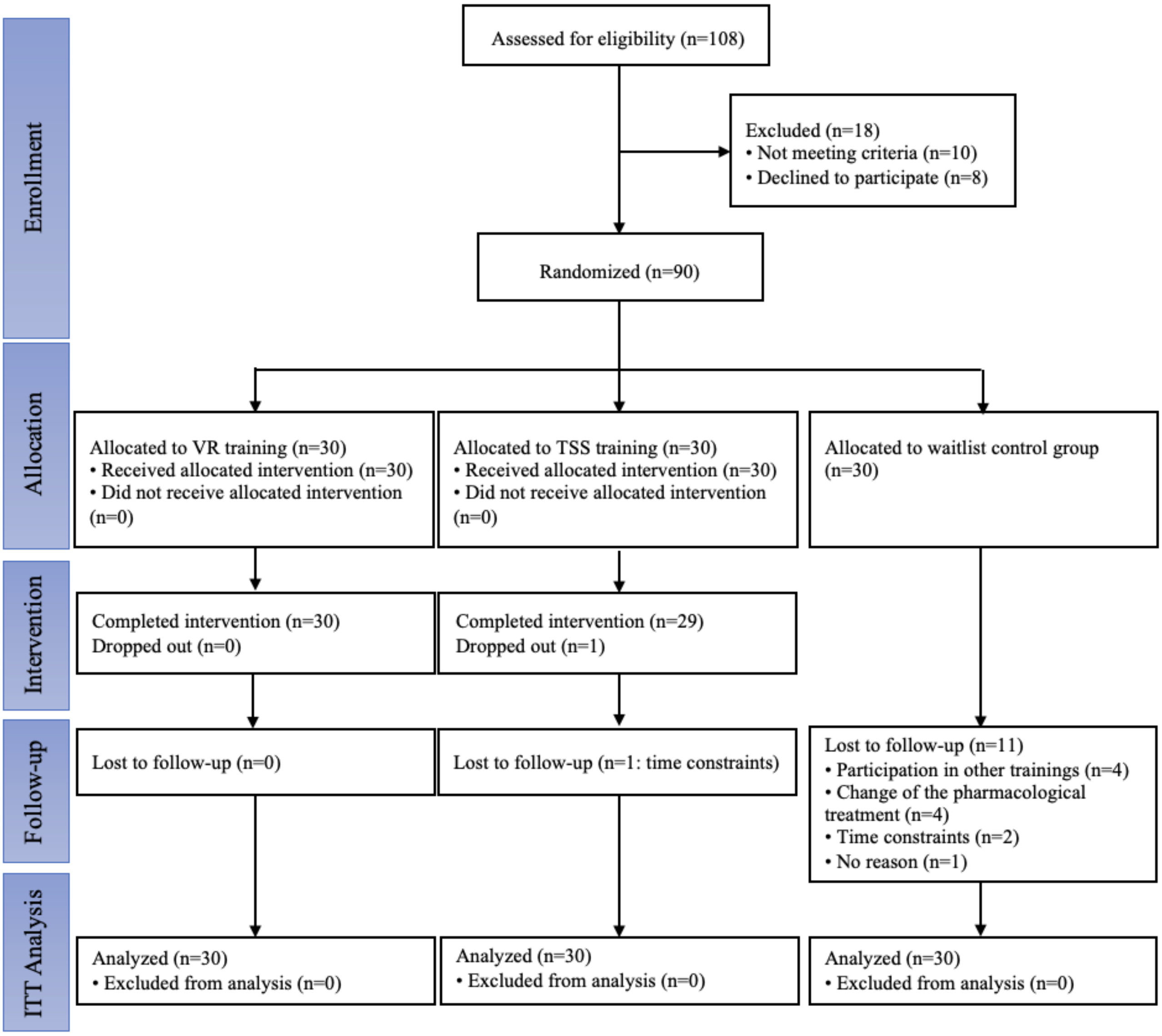
[d]Higher scores indicate better social skills.

[e]Higher scores indicate poor executive functions.

[f]Heterogeneity variance between independent groups was assessed using Levene test (*P*>.05 indicates no significant difference in variances).

During the training, all participants in the VR training group attended 12 sessions, 1 participant in the traditional social skills training group withdrew after 3 sessions due to time constraints, and 11 participants in the waitlist control group withdrew due to participation in other similar trainings, change of the pharmacological treatment, time constraints and loss to follow (Figure 1). Findings provided promising evidence of participants' acceptance of the VR training.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. ITT: intention-to-treat; TSS: traditional social skills; VR: virtual reality.



Effects of the Interventions on Outcomes

The mean scores for the clinical psychologist assessment of social skills and parent-rated executive functioning and social skills of the 3 groups at T1 and T2 are illustrated in Table 2. Significant improvements between T1 and T2 in the VR training group and the traditional social skills group in terms of clinical psychologist assessment of social skills and parent-rated self-control, initiative, and emotional control were found. Significant deterioration of parent-rated inhibition was also found. No statistically significant difference between T1 and T2 was found in the 3 groups in terms of parent-rated cooperation. In consideration of confounding variables, the assumption of normality, that is, homogeneity of variance between independent groups, was met in clinical psychologist assessments of social skills and parent-rated self-control, initiative, cooperation, and emotional control.

Clinical psychologist assessments of social skills had a statistically significant difference between groups ($F_{2,85}$=76.77; P<.001), in which the VR training group was 7.57 higher than the traditional social skills training group (d=1.50, 95% CI 0.91-2.07) and 13.70 higher than the waitlist control group (d=3.29; 95% CI 2.50-4.06).

A statistically significant difference was found between groups ($F_{2,85}$=18.77; P<.001) regarding parental ratings of self-control, with the VR training group being 1.40 higher than the traditional social skills training group (d=0.44; 95% CI –0.08 to 0.95) and 1.83 higher than the waitlist control group (d=0.44; 95% CI –0.08 to 0.95). Parent-rated initiative had a statistically significant difference between groups ($F_{2,85}$=11.93; P<.001), in which the VR training group was 0.93 higher than the traditional social skills training group (d=0.24; 95% CI –0.27 to 0.74) and 1.53 higher than the waitlist control group (d=0.40; 95% CI –0.11 to 0.91). Parent-rated cooperation had no statistically significant difference between groups.

For parent-rated emotional control, the VR training group was 1.15 lower than the traditional social skills training group (d=–0.40; 95% CI –0.91 to 0.11) and 2.80 lower than the waitlist control group (d=–0.69; 95% CI –1.20 to 0.16), and this difference was statistically significant ($F_{2,85}$=17.27; P<.001). The determination of parent-rated inhibition was not feasible due to significant heterogeneity in variance between the independent groups.

Regarding cybersickness in participants in the VR training group, no participant reported discomfort, fatigue, headache, eyestrain, nausea, difficulty concentrating, blurred vision, and dizziness during and after the training. A total of 16 participants were sweating during the VR training. All participants were satisfied with the VR training (mean 6.77, SD 0.43).

Discussion

Principal Findings

To our knowledge, this is the first RCT to evaluate the feasibility and acceptability of social skills training for children with ADHD developed through VR technology and clinically assessed by an independent clinical psychologist. Acceptance and satisfaction with the VR training by participants were high. Our findings showed that participants in the VR training and traditional social skills groups experienced significant improvements in social skills, self-control, initiative, and emotional control. Compared with the traditional social skills training group, the VR training group improved more in social skills, self-control, initiative, and emotional control. These results are preliminary; however, the results may be considered useful given their positive impact on social skills and executive function in children with ADHD.

Comparison With Prior Work

The findings match those observed in earlier studies that VR-based social skills training and traditional social skills training could improve the social skills of children, and VR-based intervention had a better performance than traditional social skills training [25,39]. The results suggest that VR-based interventions may be a promising approach for helping children with neurobiological disabilities, such as ADHD and autism spectrum disorder, acquire social skills due to the immersive and engaging nature of VR. Previous meta-analyses have shown that VR-based interventions are effective in improving cognitive function (ie, learning and attention) in children with ADHD [19,40,41], although these comparisons require caution due to the different purposes of the training and intervention protocols. The focus of our VR training program was to improve children's social skills, rather than improving their cognitive function as in previous studies.

The result of VR intervention containing significant improvement in the self-control of children with ADHD seems to accord with the study of Witowska et al [42], which suggested that VR interventions can promote sustained focus and attention, which are essential components of self-control, by presenting engaging and visually stimulating tasks. The immediate feedback loop generated by the VR system can help children with ADHD understand the relationship between their behaviors and outcomes, facilitating self-monitoring and self-regulation. VR interventions can promote self-control by reinforcing desired behaviors and highlighting the impact of impulsive or impulsive actions.

This study found that VR-based intervention could improve the initiative of children with ADHD. This may be attributable to the sense of autonomy and self-determination that VR provides, allowing children to make choices and decisions within the digital environment [43]. The opportunity to experience success and skill development through VR-based feedback and reinforcement can also boost children's confidence and willingness to initiate social interactions [39,44]. Additionally, interaction with digital characters or avatars may foster a sense of relatedness and social connectedness, encouraging children to practice and refine their social skills within the VR environment [45].

Contrary to expectations, no significant differences were found between VR training and traditional social skills training in terms of inhibition and cooperation. This differs from previous suggestions that digital classroom remediation could improve the inhibition of children with ADHD [18]. The absence of peer involvement in this study, as opposed to the successful peer interactions emphasized in prior research [46], maybe a factor contributing to the lack of differentiation between the 2 training approaches in enhancing cooperation. Further research is needed to directly compare the effects of VR-based interventions on self-control, initiative, inhibition, and cooperation across diverse populations, including children with ADHD, individuals with autism spectrum disorder, and neurotypical adults, to elucidate the specific mechanisms and contextual factors that influence the efficacy of VR-based approaches for different groups.

While the results of this study appear to be consistent with the previous studies on the positive effects of VR-based intervention on emotional control [47,48], it is important to note that the current sample of children with ADHD may respond differently to the VR intervention compared to other populations, such as individuals with autism spectrum disorder or neurotypical adults. Differences in cognitive profiles, developmental stages, and specific challenges faced by these groups could lead to variations in the effectiveness of the VR-based approach.

It was suggested that VR can evoke emotional experiences through a sense of "being there" in the digital environment [49]. VR alters individual experiences by motivating users to try new things and allowing them to alter habitual emotional responses [50]. The sample groups of previous studies were different from ours in that their participants were adults and older adults; thus, the analyses must be interpreted with caution. Further research is needed to directly compare the effects of VR-based interventions on self-control and related outcomes across diverse populations, including children with ADHD, individuals with autism spectrum disorder, and neurotypical adults. Such comparative studies would help elucidate the specific mechanisms and contextual factors that contribute to the efficacy of VR-based approaches for different groups.

An important consideration for this study is the cost-effectiveness of the VR-based social skills intervention compared to the traditional in-person approach. While a full cost-effectiveness analysis was beyond the scope of this study, we were able to gather some preliminary data on the relative costs of the 2 interventions. The VR-based intervention required 1 part-time research assistant at a rate of HK $70 (US $9) per hour to guide each participant through the digital scenarios. In contrast, the traditional in-person training involved 1 teacher with a special education background at a rate of HK $200 (US $25.74) per hour for each participant. This suggests the VR-based approach may be significantly more cost-effective, requiring less specialized personnel time per participant. Additionally, the VR equipment and software can be reused across multiple participants, whereas the in-person training requires the teacher's time for each session. This further implies the VR intervention could have lower marginal costs as the number of participants increases. While these initial cost comparisons are promising, a more rigorous cost-effectiveness analysis would be needed to fully evaluate the economic feasibility of implementing VR-based social skills training on a larger scale. Factors such as the upfront investment in VR hardware, software licensing fees, and training for research assistants would also need to be considered. Future research should explore these cost implications in more depth to comprehensively assess the economic viability of VR-based social skills interventions compared to traditional approaches.

In summary, our findings suggest that creating VR environments with different social interaction scenarios for children with ADHD may have some practical implications. VR can be particularly beneficial in improving basic social skills, self-control, initiative, and emotional control, although it was less effective in enhancing inhibition and cooperation in children with ADHD. Thus, VR may be appropriate for providing training in social skills, self-control, initiative, and emotional control. Additionally, children with ADHD reported high levels of satisfaction and motivation during short-term VR-based interventions, suggesting that VR training could be a useful first phase in assisting these children in acquiring social skills in a community setting.

Limitations

Although the effects of VR training are evident, potential limitations should be considered when interpreting the results. First, the small sample size recruited for this study may have limited the statistical power to determine differences between groups; nevertheless, this study provides important insights. Second, the design of this RCT had no follow-up period, and thus, the long-term effects after the intervention are unknown. Third, the trial was conducted with a highly specific type of population, and the results may not generalize to more complex ADHD populations. Fourth, a key limitation was that the satisfaction of the participants in the traditional social skills training group was not measured. Evaluating the participants' satisfaction, engagement, and perceptions of the traditional training approach would have provided valuable comparative insights. Without this information, it is difficult to fully assess the relative benefits and drawbacks of VR training compared to the standard intervention from the participants' perspectives. Future studies should consider incorporating participant satisfaction measures for all intervention groups to enable a more comprehensive evaluation of the different training methods. Fifth, although age differences across groups were found to be insignificant, the sample size within multiple age groups was limited, with some age groups having only 1 participant. This uneven distribution restricts our ability to conduct robust subgroup analyses to understand the differential effects of VR-based intervention versus traditional social skills training on social skills across various age groups. Given that social skills in children with ADHD are known to improve with age, this limitation highlights a significant knowledge gap in our study. Future research should aim to recruit a larger and more evenly distributed sample across different age groups to enable more detailed age-specific analyses. Finally, the sample size of this study may not provide sufficient statistical power for conducting subgroup analyses. This limitation means that our findings primarily reflect general comparisons between the intervention groups rather than nuanced insights into how different age groups respond to VR-based intervention versus traditional social skills training. Researchers should consider this limitation when interpreting our results and should design future studies with larger sample sizes to ensure adequate power for subgroup analyses.

Conclusions

This RCT investigated the feasibility and effectiveness of using VR in improving the social skills of children aged 6-12 years with ADHD. The results showed that VR training was more effective than traditional social skills training in terms of enhancing basic social skills, self-control, initiative, and emotional control. VR and traditional social skills training can enhance basic social skills, self-control, initiative, and emotional control in children with ADHD. These encouraging findings require further confirmation with large-scale interventions to determine the sustainability of the effects. Continued research in this area will contribute to the development of evidence-based practices and interventions that can ultimately improve the social outcomes and quality of life for this vulnerable population.

Acknowledgments

The authors thank all the participants and their guardians who participated in this study. This study is supported by the Health and Medical Research Fund (project 10211516).

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2354 KB - games_v12i1e58963_app1.pdf]

Multimedia Appendix 2
Description of the 3 scenarios in the virtual reality intervention.
[DOCX File , 7933 KB - games_v12i1e58963_app2.docx]

Multimedia Appendix 3
Participants conducting the virtual reality–based intervention.
[DOCX File , 798 KB - games_v12i1e58963_app3.docx]

References

1. Wilens TE, Spencer TJ. Understanding attention-deficit/hyperactivity disorder from childhood to adulthood. Postgrad Med 2010;122(5):97-109 [FREE Full text] [doi: 10.3810/pgm.2010.09.2206] [Medline: 20861593]
2. Espinet SD, Graziosi G, Toplak ME, Hesson J, Minhas P. A review of Canadian diagnosed ADHD prevalence and incidence estimates published in the past decade. Brain Sci 2022;12(8):1051 [FREE Full text] [doi: 10.3390/brainsci12081051] [Medline: 36009114]
3. Salari N, Ghasemi H, Abdoli N, Rahmani A, Shiri MH, Hashemian AH, et al. The global prevalence of ADHD in children and adolescents: a systematic review and meta-analysis. Ital J Pediatr 2023;49(1):48 [FREE Full text] [doi: 10.1186/s13052-023-01456-1] [Medline: 37081447]
4. Jangmo A, Stålhandske A, Chang Z, Chen Q, Almqvist C, Feldman I, et al. Attention-deficit/hyperactivity disorder, school performance, and effect of medication. J Am Acad Child Adolesc Psychiatry 2019;58(4):423-432 [FREE Full text] [doi: 10.1016/j.jaac.2018.11.014] [Medline: 30768391]
5. Clark C, Prior M, Kinsella G. The relationship between executive function abilities, adaptive behaviour, and academic achievement in children with externalising behaviour problems. J Child Psychol Psychiatry 2002;43(6):785-796. [doi: 10.1111/1469-7610.00084] [Medline: 12236613]
6. Powell V, Riglin L, Hammerton G, Eyre O, Martin J, Anney R, et al. What explains the link between childhood ADHD and adolescent depression? Investigating the role of peer relationships and academic attainment. Eur Child Adolesc Psychiatry 2020;29(11):1581-1591 [FREE Full text] [doi: 10.1007/s00787-019-01463-w] [Medline: 31932968]
7. Sedgwick JA, Merwood A, Asherson P. The positive aspects of attention deficit hyperactivity disorder: a qualitative investigation of successful adults with ADHD. Atten Defic Hyperact Disord 2019;11(3):241-253. [doi: 10.1007/s12402-018-0277-6] [Medline: 30374709]
8. Ronk MJ, Hund AM, Landau S. Assessment of social competence of boys with attention-deficit/hyperactivity disorder: problematic peer entry, host responses, and evaluations. J Abnorm Child Psychol 2011;39(6):829-840. [doi: 10.1007/s10802-011-9497-3] [Medline: 21468667]
9. Kofler MJ, Groves NB, Singh LJ, Soto EF, Chan ESM, Irwin LN, et al. Rethinking hyperactivity in pediatric ADHD: preliminary evidence for a reconceptualization of hyperactivity/impulsivity from the perspective of informant perceptual processes. Psychol Assess 2020;32(8):752-767 [FREE Full text] [doi: 10.1037/pas0000856] [Medline: 32478528]
10. Bodalski EA, Knouse LE, Kovalev D. Adult ADHD, emotion dysregulation, and functional outcomes: examining the role of emotion regulation strategies. J Psychopathol Behav Assess 2018;41(1):81-92. [doi: 10.1007/s10862-018-9695-1]
11. Erskine HE, Norman RE, Ferrari AJ, Chan GC, Copeland WE, Whiteford HA, et al. Long-term outcomes of attention-deficit/hyperactivity disorder and conduct disorder: a systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2016;55(10):841-850. [doi: 10.1016/j.jaac.2016.06.016] [Medline: 27663939]
12. Evans SW, Owens JS, Bunford N. Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. J Clin Child Adolesc Psychol 2014;43(4):527-551 [FREE Full text] [doi: 10.1080/15374416.2013.850700] [Medline: 24245813]
13. Mikami AY, Smit S, Khalis A. Social skills training and ADHD—What works? Curr Psychiatry Rep 2017;19(12):93. [doi: 10.1007/s11920-017-0850-2] [Medline: 29082428]
14. Willis D, Siceloff ER, Morse M, Neger E, Flory K. Stand-alone social skills training for youth with ADHD: a systematic review. Clin Child Fam Psychol Rev 2019;22(3):348-366. [doi: 10.1007/s10567-019-00291-3] [Medline: 30796673]
15. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017;47(14):2393-2400 [FREE Full text] [doi: 10.1017/S003329171700040X] [Medline: 28325167]
16. Manca D, Brambilla S, Colombo S. Bridging between virtual reality and accident simulation for training of process-industry operators. Adv Eng Softw 2013;55:1-9. [doi: 10.1016/j.advengsoft.2012.09.002]
17. Wong KP, Lai CYY, Qin J. Systematic review and meta-analysis of randomised controlled trials for evaluating the effectiveness of virtual reality therapy for social anxiety disorder. J Affect Disord 2023;333:353-364. [doi: 10.1016/j.jad.2023.04.043] [Medline: 37084968]
18. Bioulac S, Micoulaud-Franchi J, Maire J, Bouvard MP, Rizzo AA, Sagaspe P, et al. Virtual remediation versus methylphenidate to improve distractibility in children with ADHD: a controlled randomized clinical trial study. J Atten Disord 2020;24(2):326-335. [doi: 10.1177/1087054718759751] [Medline: 29562853]

19. Parsons TD, Duffield T, Asbee J. A comparison of virtual reality classroom continuous performance tests to traditional continuous performance tests in delineating ADHD: a meta-analysis. Neuropsychol Rev 2019;29(3):338-356. [doi: 10.1007/s11065-019-09407-6] [Medline: 31161465]
20. Rizzo AA, Bowerly T, Buckwalter JG, Klimchuk D, Mitura R, Parsons TD. A virtual reality scenario for all seasons: the virtual classroom. CNS Spectr 2006;11(1):35-44. [doi: 10.1017/s1092852900024196] [Medline: 16400254]
21. Parsons TD, Silva TM, Pair J, Rizzo AA. Virtual environment for assessment of neurocognitive functioning: virtual reality cognitive performance assessment test. Stud Health Technol Inform 2008;132:351-356. [Medline: 18391320]
22. Howard MC, Gutworth MB. A meta-analysis of virtual reality training programs for social skill development. Comput Educ 2020;144:103707. [doi: 10.1016/j.compedu.2019.103707]
23. Park KM, Ku J, Choi SH, Jang HJ, Park JY, Kim SI, et al. A virtual reality application in role-plays of social skills training for schizophrenia: a randomized, controlled trial. Psychiatry Res 2011;189(2):166-172. [doi: 10.1016/j.psychres.2011.04.003] [Medline: 21529970]
24. Parsons S, Mitchell P. The potential of virtual reality in social skills training for people with autistic spectrum disorders. J Intellect Disabil Res 2002;46:430-443. [doi: 10.1046/j.1365-2788.2002.00425.x] [Medline: 12031025]
25. Montoya-Rodríguez MM, de Souza Franco V, Tomás Llerena C, Molina Cobos FJ, Pizzarossa S, García AC, et al. Virtual reality and augmented reality as strategies for teaching social skills to individuals with intellectual disability: a systematic review. J Intellect Disabil 2023;27(4):1062-1084. [doi: 10.1177/17446295221089147] [Medline: 35420906]
26. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
27. Horsburgh J, Ippolito K. A skill to be worked at: using social learning theory to explore the process of learning from role models in clinical settings. BMC Med Educ 2018;18(1):156 [FREE Full text] [doi: 10.1186/s12909-018-1251-x] [Medline: 29970052]
28. Matthys W, Schutter DJLG. Improving our understanding of impaired social problem-solving in children and adolescents with conduct problems: implications for cognitive behavioral therapy. Clin Child Fam Psychol Rev 2022;25(3):552-572. [doi: 10.1007/s10567-021-00376-y] [Medline: 35165840]
29. Eom H, Kim KK, Lee S, Hong Y, Heo J, Kim J, et al. Development of virtual reality continuous performance test utilizing social cues for children and adolescents with attention-deficit/hyperactivity disorder. Cyberpsychol Behav Soc Netw 2019;22(3):198-204. [doi: 10.1089/cyber.2018.0377] [Medline: 30672714]
30. Wong KP, Qin J, Xie YJ, Zhang B. Effectiveness of technology-based interventions for school-age children with attention-deficit/hyperactivity disorder: systematic review and meta-analysis of randomized controlled trials. JMIR Ment Health 2023;10:e51459 [FREE Full text] [doi: 10.2196/51459] [Medline: 37988139]
31. Wong KP, Qin J. Effectiveness of social virtual reality training in enhancing social interaction skills in children with attention-deficit/hyperactivity disorder: protocol for a three-arm pilot randomized controlled trial. JMIR Res Protoc 2023;12:e48208 [FREE Full text] [doi: 10.2196/48208] [Medline: 37721790]
32. Risi D, Palmisano S. Effects of postural stability, active control, exposure duration and repeated exposures on HMD induced cybersickness. Displays 2019;60:9-17. [doi: 10.1016/j.displa.2019.08.003]
33. Huang YH, Chung CY, Ou HY, Tzang RF, Huang KY, Liu HC, et al. Treatment effects of combining social skill training and parent training in Taiwanese children with attention deficit hyperactivity disorder. J Formos Med Assoc 2015;114(3):260-267 [FREE Full text] [doi: 10.1016/j.jfma.2012.10.019] [Medline: 25777975]
34. Cheung PPP, Siu AMH, Brown T. Measuring social skills of children and adolescents in a Chinese population: preliminary evidence on the reliability and validity of the translated Chinese version of the social skills improvement system-rating scales (SSIS-RS-C). Res Dev Disabil 2017;60:187-197 [FREE Full text] [doi: 10.1016/j.ridd.2016.11.019] [Medline: 28012420]
35. Riggio RE. Assessment of basic social skills. J Pers Soc Psychol 1986;51(3):649-660. [doi: 10.1037//0022-3514.51.3.649]
36. Huang F, Sun L, Qian Y, Liu L, Ma QG, Yang L, et al. Cognitive function of children and adolescents with attention deficit hyperactivity disorder and learning difficulties: a developmental perspective. Chin Med J 2016;129(16):1922-1928 [FREE Full text] [doi: 10.4103/0366-6999.187861] [Medline: 27503016]
37. Bouchard S, Berthiaume M, Robillard G, Forget H, Daudelin-Peltier C, Renaud P, et al. Arguing in favor of revising the Simulator Sickness Questionnaire factor structure when assessing side effects induced by immersions in virtual reality. Front Psychiatry 2021;12:739742 [FREE Full text] [doi: 10.3389/fpsyt.2021.739742] [Medline: 34803762]
38. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
39. Frolli A, Savarese G, Di Carmine F, Bosco A, Saviano E, Rega A, et al. Children on the autism spectrum and the use of virtual reality for supporting social skills. Children 2022;9(2):181 [FREE Full text] [doi: 10.3390/children9020181] [Medline: 35204903]
40. Adams R, Finn P, Moes E, Flannery K, Rizzo AS. Distractibility in attention/deficit/ hyperactivity disorder (ADHD): the virtual reality classroom. Child Neuropsychol 2009;15(2):120-135. [doi: 10.1080/09297040802169077] [Medline: 18608217]

41. Corrigan N, Păsărelu CR, Voinescu A. Immersive virtual reality for improving cognitive deficits in children with ADHD: a systematic review and meta-analysis. Virtual Real 2023:1-20 [FREE Full text] [doi: 10.1007/s10055-023-00768-1] [Medline: 36845650]
42. Witowska J, Schmidt S, Wittmann M. What happens while waiting? How self-regulation affects boredom and subjective time during a real waiting situation. Acta Psychol 2020;205:103061 [FREE Full text] [doi: 10.1016/j.actpsy.2020.103061] [Medline: 32203734]
43. Huang YC, Backman SJ, Backman KF, McGuire FA, Moore D. An investigation of motivation and experience in virtual learning environments: a self-determination theory. Educ Inf Technol 2018;24(1):591-611 [FREE Full text] [doi: 10.1007/s10639-018-9784-5]
44. Skalski S, Konaszewski K, Pochwatko G, Balas R, Surzykiewicz J. Effects of hemoencephalographic biofeedback with virtual reality on selected aspects of attention in children with ADHD. Int J Psychophysiol 2021;170:59-66. [doi: 10.1016/j.ijpsycho.2021.10.001] [Medline: 34653532]
45. Thabrew H, Chubb LA, Kumar H, Fouché C. Immersive reality experience technology for reducing social isolation and improving social connectedness and well-being of children and young people who are hospitalized: open trial. JMIR Pediatr Parent 2022;5(1):e29164 [FREE Full text] [doi: 10.2196/29164] [Medline: 35333184]
46. Ramani GB. Influence of a playful, child-directed context on preschool children‘s peer cooperation. Merrill Palmer Q 2012;58(2):159-190. [doi: 10.1353/mpq.2012.0011]
47. Colombo D, Díaz-García A, Fernandez-Álvarez J, Botella C. Virtual reality for the enhancement of emotion regulation. Clin Psychol Psychother 2021;28(3):519-537. [doi: 10.1002/cpp.2618] [Medline: 34048621]
48. Montana JI, Matamala-Gomez M, Maisto M, Mavrodiev PA, Cavalera CM, Diana B, et al. The benefits of emotion regulation interventions in virtual reality for the improvement of wellbeing in adults and older adults: a systematic review. J Clin Med 2020;9(2):500 [FREE Full text] [doi: 10.3390/jcm9020500] [Medline: 32059514]
49. Gross JJ. Emotion regulation: current status and future prospects. Psychol Inq 2015;26(1):1-26. [doi: 10.1080/1047840x.2014.940781]
50. Valmaggia LR, Latif L, Kempton MJ, Rus-Calafell M. Virtual reality in the psychological treatment for mental health problems: an systematic review of recent evidence. Psychiatry Res 2016;236:189-195. [doi: 10.1016/j.psychres.2016.01.015] [Medline: 26795129]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
CONSORT: Consolidated Standards of Reporting Trials
RCT: randomized controlled trial
VR: virtual reality

Edited by A Coristine; submitted 29.03.24; peer-reviewed by H Bhurawala, WH Dong; comments to author 26.06.24; revised version received 14.07.24; accepted 17.09.24; published 28.10.24.

Please cite as:
Wong KP, Zhang B, Lai CYY, Xie YJ, Li Y, Li C, Qin J
Empowering Social Growth Through Virtual Reality–Based Intervention for Children With Attention-Deficit/Hyperactivity Disorder: 3-Arm Randomized Controlled Trial
JMIR Serious Games 2024;12:e58963
URL: https://games.jmir.org/2024/1/e58963
doi:10.2196/58963
PMID:

Original Paper

Risk Perception and Knowledge Following a Social Game–Based Tobacco Prevention Program for Adolescents: Pilot Randomized Comparative Trial

Georges Khalil[1], MPH, PhD; Erica Ramirez[1], BS, RN; Meerah Khan[1], MPH; Bairu Zhao[1], BS, MS; Nuno Ribeiro[2], PhD; Patrick Balian[3], DBA

[1]Department of Health Outcomes and Biomedical Informatics, University of Florida, Gainesville, FL, United States

[2]Cancer Prevention Unit, Institute of Pathology and Molecular Immunology of University of Porto and Institute for Research and Innovation in Health, University of Porto, Porto, Portugal

[3]Warrington College of Business, University of Florida, Gainesville, FL, United States

Corresponding Author:
Georges Khalil, MPH, PhD
Department of Health Outcomes and Biomedical Informatics
University of Florida
Malachowsky Hall for Data Science and Information Technology
1889 Museum Rd, Suite 7000
Gainesville, FL, 32603
United States
Phone: 1 3526279467
Email: gkhalil@ufl.edu

Abstract

Background: Adolescence is a critical developmental stage that is particularly vulnerable to the initiation of tobacco use. Despite the well-documented health risks associated with tobacco use, it remains prevalent among adolescents. Games for health are a promising strategy for tobacco prevention, using experiential and social learning theories to enhance engagement and improve behavior change.

Objective: This pilot study aims to (1) compare the social game–based program Storm-Heroes to a nonsocial program regarding adolescents' personal and social experiences and (2) examine how these experiences predict higher tobacco knowledge and perceived risks of vaping and conventional tobacco use.

Methods: In a cluster-randomized comparative design, 4 after-school sites (N=79 adolescents) were recruited in person and randomized in a single-blinded format to 1 of 2 interventions: the social game Storm-Heroes (44/79, 56%) or the nonsocial program A Smoking Prevention Interactive Experience (ASPIRE; 35/79, 44%). A study team member supervised both interventions. Data were collected at baseline, immediate follow-up, and a 1.5-month follow-up (45/74, 61% retained). Repeated measures mixed effects models were conducted.

Results: A total of 45 participants continued until the 1.5-month follow-up. Participants in the Strom-Heroes group were more likely to increase their perceived risk of vaping (B=0.40; *P*<.001), perceived risk of conventional tobacco use (B=0.35; *P*=.046), and tobacco knowledge (B=1.63; *P*<.001) than those in the control condition. The usability level of the program was related to a higher perceived risk of vaping (B=0.16; *P*=.003) and conventional tobacco use (B=0.16; *P*=.02) by follow-up. Attention to the program was also related to higher perceived risk of vaping (B=0.12; *P*=.002) and conventional tobacco use (B=0.14; *P*<.001). Distraction was not related to either perceived risk of vaping (*P*=.15) or perceived risk of conventional tobacco use (*P*=.71). In contrast, both more attention (B=0.60; *P*<.001) and less distraction (B=–0.37; *P*<.001) were related to higher tobacco knowledge.

Conclusions: The increased perceived risk of vaping and conventional tobacco among Storm-Heroes participants aligns with the program's goals of improving participants' awareness of the risks associated with tobacco use and their tobacco knowledge. However, distraction weakened the effect of the program on tobacco knowledge, indicating that emphasis needs to be placed on minimizing distraction for better outcomes. With the results of this study, researchers can work to advance the current version of Storm-Heroes and amplify engagement in the program to improve its potential for preventing adolescents' initiation of tobacco use.

Trial Registration: ClinicalTrials.gov NCT02703597; https://clinicaltrials.gov/study/NCT02703597

(JMIR Serious Games 2024;12:e63296) doi:10.2196/63296

KEYWORDS

tobacco prevention; vaping; combustible tobacco; risk perception; adolescent; games; social interaction

Introduction

Background

Adolescence is a critical developmental stage that is particularly vulnerable to the initiation of tobacco use [1-3]. Research indicates that exposure to nicotine during this period is associated with substantial impairments in brain growth, psychological harm, and long-term physical health outcomes [4,5]. Despite the well-known risks associated with adolescent tobacco use, the rates of use among this age group remain a concern. In 2023, approximately 28% of high-school students and 14.7% of middle-school students reported ever using a tobacco product [6,7]. Of the youth who reported ever using tobacco products, approximately 50% of them are current tobacco users [6,7]. As a result, the need for tobacco prevention is evident.

One particularly promising strategy for tobacco prevention is the application of games for health. Gameplay can include an amalgam of entertainment and education strategies to drive health behavior change [8]. On the basis of the experiential learning theory and the social learning theory, the immersive nature of gameplay facilitates a successful learning process through a playful and entertaining environment [9,10]. Game-based interventions can increase motivation, engagement, and overall sustainability of health behaviors [11]. In addition to our work, researchers have shown the success of games for health through randomized controlled trials [12-15].

Engaging gameplay has proven to be a promising avenue for tobacco prevention [16]. On the basis of a systematic review [17], most games for combustible tobacco prevention and cessation have leveraged the use of rewards and interactive activities to drive behavior change. Among these interventions, success was primarily observed in smoking cessation games rather than in prevention efforts [17,18]. In contrast, games meant for vaping or e-cigarette prevention have recently shown success. One example is a game called “Invite Only VR,” which showed improvement in e-cigarette knowledge, nicotine addiction knowledge, perceived addictiveness of e-cigarettes, and perceptions of harm [19]. In addition, 1 comprehensive game covering vaping and combustible tobacco, smokeSCREEN, improved antitobacco beliefs and tobacco knowledge [20]. These results highlight the potential success of game-based interventions.

By including various gaming elements (eg, competition, collaboration, reward, goal setting, and storytelling), games can provide flexibility in addressing different issues pertaining to tobacco use. One qualitative study for the design of tobacco prevention games examined adolescents’ gaming preferences and showcased the elements of cooperation, storytelling, and physical performance as key experiential learning elements for tobacco prevention [21]. The findings suggest that gaming elements can be combined to design an effective and engaging tool that covers the complexities of different tobacco products and addresses unique topics pertaining to this risky behavior.

A Social Game–Based Intervention

This line of research on gaming elements for tobacco prevention led to the design of a social game–based intervention, called Storm-Heroes, which is ideal for education systems (eg, schools and after-school programs). As a social game, Storm-Heroes offers adolescents the opportunity to witness and model healthy behaviors, such as rejecting tobacco, thereby promoting tobacco risk education [21]. With social gaming, Storm-Heroes relies on the social learning theory by promoting interpersonal discussions and boosting self-efficacy through the practice of skills to stay tobacco free. In addition, Storm-Heroes conveys normative feedback, influencing adolescents’ risk perceptions regarding tobacco use [21]. This aligns with the health beliefs model, which posits that psychosocial factors such as social interaction and peer pressure can promote risk perception and ultimately encourage behavior change [22]. Through these mechanisms, Storm-Heroes serves as a tool for tobacco prevention, leveraging peer influence and normative feedback to positively impact adolescents’ perceptions of tobacco use risks and improve knowledge.

Study Objectives

The purpose of this pilot study is to (1) compare the social game–based program Storm-Heroes to a nonsocial program with respect to adolescents’ personal and social experience with the program and (2) examine the role of adolescents’ experience with the program in predicting higher perceived risk of vaping, perceived risk of conventional tobacco use, and knowledge by follow-up. Table 1 clarifies the hypotheses tested in this study.

Table 1. List of hypotheses.

Types of hypothesis	Hypothesis statements
Personal experience	• Engagement with Storm-Heroes will result in higher attention, lower distraction, and higher recognition of program imagery than engagement with a nonsocial equivalent program. • Engagement with Storm-Heroes will result in perceptions of better usability, higher level of fun, better narrative quality, more enjoyment, and more creative freedom than engagement with a nonsocial equivalent program.
Social experience	• Engagement with Storm-Heroes will result in higher engagement in peer-to-peer discussions and better quality of discussions than engagement with a nonsocial equivalent program.
Tobacco-related outcomes	• Engagement with Storm-Heroes will result in improved perceived risk of vaping and higher perceived risk of conventional tobacco use (cigarettes, cigars, and little cigars) than engagement with a nonsocial equivalent program. • Engagement with Storm-Heroes will result in improved tobacco knowledge than engagement with a nonsocial equivalent program.
User experience mechanisms of change	• Personal experience factors and social interactivity will predict higher perceived risk of vaping, perceived risk of conventional tobacco use, and tobacco knowledge by follow-up while controlling for program allocation.

Methods

Study Design

To pilot-test adolescents' experience with Storm-Heroes, this study involved a 2-arm single-blinded cluster-randomized comparative trial. The pilot study was conducted in June 2021 at 4 after-school sites in Florida, and it was registered at the ClinicalTrials.gov registry, as part of a larger study (identifier: NCT02703597). Its components adhere to the CONSORT (Consolidated Standards of Reporting Trials) and CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines [23,24]. We assessed demographic information at baseline and program experience at immediate follow-up. We measured perceived risk of vaping, perceived risk of combustible tobacco use, and tobacco knowledge at baseline and 1.5-month follow-up.

A Brief Description of the Interventions

We compared Storm-Heroes with A Smoking Prevention Interactive Experience (ASPIRE), a nonsocial program that is similar to Storm-Heroes in terms of session structure and type of health content. Table 2 describes the differences and similarities in design elements for Storm-Heroes and ASPIRE. A detailed description of the intervention is provided in Multimedia Appendix 1 [25-32], which follows the Template for Intervention Description and Replication (TIDieR) checklist [33]. Multimedia Appendix 2 presents the CONSORT-EHEALTH items.

Table 2. Differences and similarities in elements for Storm-Heroes and A Smoking Prevention Interactive Experience (ASPIRE).

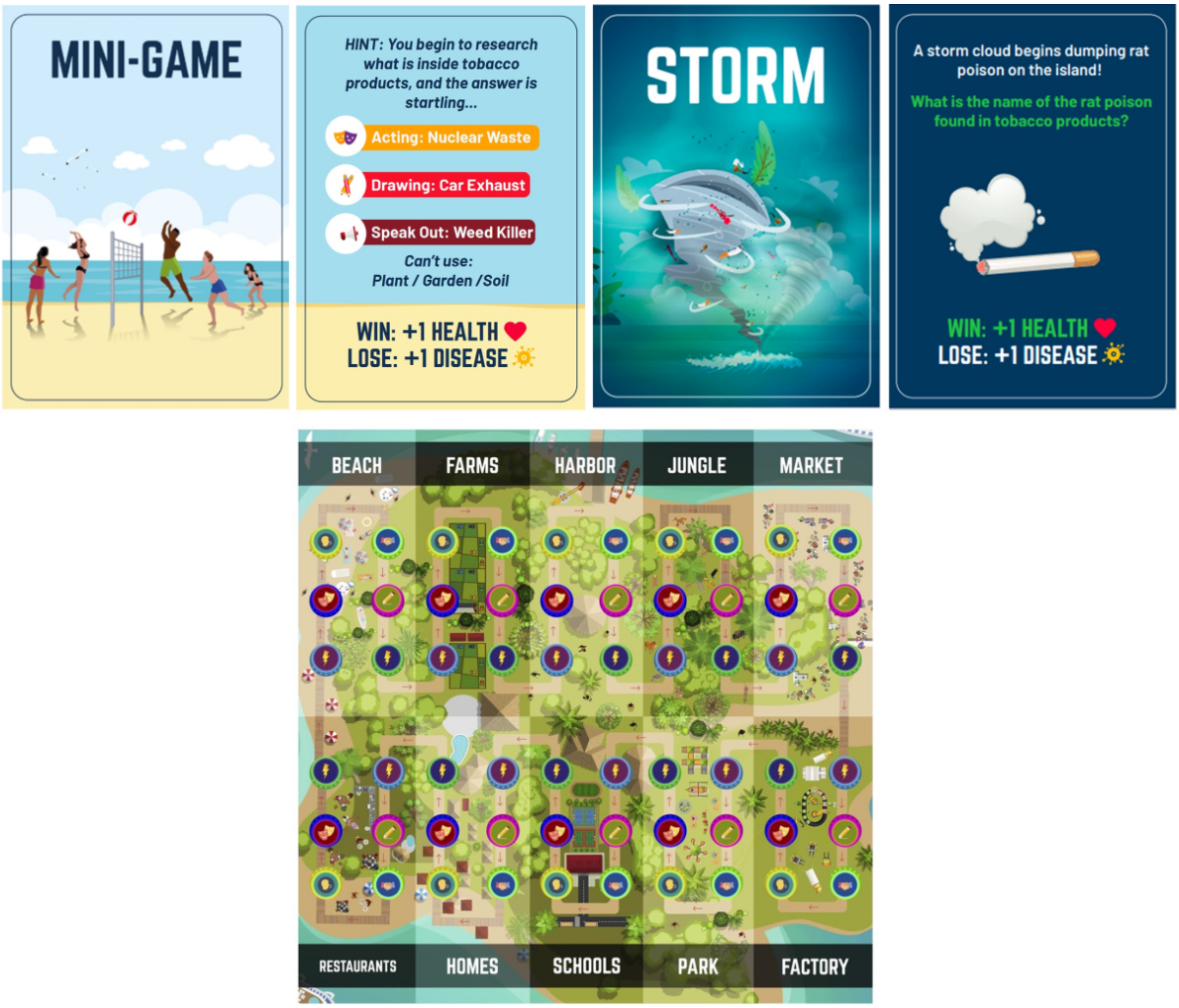
Elements	Storm-Heroes	ASPIRE	Mechanism of action
Health content based on TTM[a]	Yes	Yes	Presenting both programs with the TTM ensures that adolescents receive consistent health information and persuasive messaging based on the processes of change.
Entertainment-education videos	Yes	Yes	By leveraging narratives, these videos deliver messages that correct misconceptions about the prevalence and acceptability of tobacco use, shaping adolescents' perceptions to align with healthy norms.
Human-computer interaction	Yes	Yes	This interaction engages adolescents through a digital platform, making learning about tobacco risks engaging and interactive.
Gameplay	Yes	No	Gameplay improves engagement and provides a simulated environment where adolescents can learn through experience and build confidence.
Social interaction	Yes	No	Social interaction encourages adolescents to discuss tobacco risks and refusal skills with peers, enhancing their understanding and commitment to staying tobacco free.
Hybrid format	Yes	No	Combines online and offline activities to keep adolescents engaged and reinforce learning in various contexts.
Dosage and frequency	5 weekly sessions for 45 min each	5 weekly sessions for 45 min each	The dosage and frequency ensure that adolescents receive similar and consistent exposure for both programs.

[a]TTM: transtheoretical model.

ASPIRE is a computer-guided intervention that uses engaging videos and interactive activities across 5 sessions over 5 weeks, with each session lasting approximately 45 minutes. It aims to enhance information retention and guide adolescents toward a tobacco-free lifestyle by engaging users through text, animations, videos, and activities. ASPIRE is evidence based and tested for tobacco prevention. The intervention program is freely accessible over the web [34,35].

The Storm-Heroes intervention was developed collaboratively, involving a game designer, a research team with tobacco education expertise, and a youth design committee. Messages were designed using scientific evidence and message-framing strategies to impact tobacco risk perception, knowledge, and intention to use. The intervention aims to educate adolescents about tobacco risks, environmental consequences, and impacts on social and mental well-being, incorporating 367 unique antitobacco messages based on the transtheoretical model [36] and empowerment theory. The design process resulted in a dynamic and socially engaging educational program, Storm-Heroes, combining digital and in-person elements. It seamlessly integrates web-based components with game-based tabletop activities, including ASPIRE-derived videos and game-based social activities for group interaction (Figure 1).

Figure 1. Depiction of Storm-Heroes activities and game board.



In Storm-Heroes, adolescents engage within a narrative. They play the role of friends on an island struck by a storm bringing tobacco products, harmful chemicals, and disease. To combat the storm's effects and save their island, teams embark on quests, participating in entertainment-education videos and activities. Before engaging in the program, adolescents are grouped using a validated social network algorithm. The grouping process ensures that each participant with high intentions to use tobacco is grouped with close friends who do not intend to use tobacco, facilitating constructive support during activities. Storm-Heroes offers adolescents 5 main activities delivered on validated board game material. These include trivia with multiple-choice questions, acting where one member silently acts as others guess, drawing for guessing from sketches, speaking out for verbal clues, and teamwork scenarios presenting group dilemmas. The activities aim to engage teams in collaborative problem-solving around tobacco-related topics. Multimedia Appendix 1 describes the activities and how they are presented to players. The materials of Storm-Heroes include informative background information in game-based social activities, such as scripts and task instructions, a tabletop game board, decks for board game cards, dice, tokens, and pons. The materials can be accessed by reaching out to the researchers.

Both ASPIRE and Storm-Heroes cover a comprehensive list of health topics related to tobacco, including its composition, effects on the body and brain, environmental impact, and strategies for tobacco prevention and advocacy (Multimedia

Appendix 1). The content is structured consistently across both programs, covering aspects from understanding tobacco to building skills for a tobacco-free lifestyle and community activism.

Ethical Considerations

The institutional review board for human subject research at the University of Florida approved this study (IRB201903082). Adolescents and their parents were informed of the study's purpose and procedure. They provided written informed parental consent and written informed child assent. During data collection sessions, participants were reminded of the study purpose, procedures, risks, and that they could withdraw from the study at any time. Participants completed the surveys in a private classroom with supervision. Participants were asked to maintain the confidentiality of their own identity and the identity of other participants. Participants received up to US $50 as compensation. Data were deidentified before analysis.

Study Procedures

A total of 4 after-school sites located in Florida were randomly selected for recruitment. After approval from the program directors, a verbal announcement reached adolescents at these sites, and interested adolescents completed child assent and parental permission. For participation, adolescents needed to be aged 11 to 18 years (inclusive) and enrolled in a middle school or high school. Adolescents also needed to be comfortable using a computer and the internet.

During recruitment, adolescents and their parents were informed offline that the study aimed to improve adolescent health through an interactive program in Florida after-school sites; that the study may take approximately 2 months and 1 week; and that they will engage in activities, surveys, and interviews. The incentive was described to potential participants, and they were informed that participation is voluntary and confidential and that the data would be securely stored at the University of Florida. Recruitment and data collection took 3 months to complete.

Participants started their experience with the intervention 3 to 7 days after they completed the baseline survey. The statistical team generated the random allocation of sites to each condition. In ASPIRE and Storm-Heroes, participants used similar computers and had private classrooms for participation. A study team member was available for technical assistance and supervision. A volunteer site staff trained in youth engagement was present to ensure that participants did not deviate from the requested data collection procedures. Participants completed surveys in a supervised classroom setting immediately after the intervention and again 1.5 months later. Participants completed other survey assessments at follow-up (data not included in this paper).

Implementation of Each Intervention

Our study staff were trained to implement the program at the after-school sites. They traveled to each study site location to administer the ASPIRE and the Storm-Heroes programs. Participants were not told which intervention was of interest to the researchers. During the site visits, study staff recorded attendance, ensured that the appropriate regimen was implemented, and addressed any questions or concerns participants had during the sessions.

At each site designated to receive ASPIRE, adolescents engaged in five 1-hour sessions exclusively focused on the full ASPIRE program. This regimen was conducted similarly to previous work on ASPIRE. At each site designated to receive Storm-Heroes, participants were first organized into groups comprising 3 to 6 individuals, determined by the outcomes of the social network survey conducted at baseline. With the aid of the social network algorithm, the study team grouped each at-risk adolescent (those indicating the highest intention to use tobacco) with 2 to 5 of their closest peers exhibiting lower intention to use tobacco. Unexpectedly, it was observed that some participants were absent during certain sessions. As a result, the grouping was re-evaluated using the algorithm for the sessions when participants were absent. Within their groups, participants were instructed to engage in ASPIRE activities, followed by game-based social activities within the board game. The duration of board game play varied for each session depending on the length of the assigned ASPIRE activity.

Measures

We assessed the measures through web-based closed surveys in a classroom setting, and a study staff was available for assistance. Survey measures have been previously tested and validated. Multimedia Appendix 3 [34,37-47] includes a detailed description of the main measures, measure references, and Cronbach α values when applicable.

At baseline only, we included survey questions pertaining to potential confounders and demographic characteristics, including age, sex at birth, ethnicity, race, average grade at school, the number of detentions at school, parents' highest level of education, and perceived skills in playing board games. At baseline, we also measured the status of using vaping products, cigarettes, and cigars or little cigars using the Minnesota smoking index [37].

At both baseline and 1.5-month follow-up, we measured perceived risk of vaping, perceived risk of using conventional tobacco products (cigarettes, cigars, and little cigars), and tobacco knowledge [38,39] (Multimedia Appendix 3).

At immediate follow-up, we collected data regarding participants' experience with each of the programs using validated measures. First, to check for expected differences and similarities between the ASPIRE and the Storm-Heroes conditions, we assessed measures pertaining to key program features. We expected group differences with respect to perceived social interactivity [40] and group similarities with respect to attitude toward the program, visual esthetics, and emotional involvement [41-44]. Next, to assess engagement, we measured recognition of program imagery, attention to the program, and distraction from the program [45-47]. To capture user experience, we assessed participants' perceptions regarding the usability of the program [42,43], level of fun [41], narrative quality [42,43], program enjoyment [34], and creative freedom [42,43]. Considering the role of social interaction in the success of Storm-Heroes, we asked participants to indicate if they engaged in any discussions with their peers after the program. If they confirmed that they engaged in discussions, they were then asked to share the content of their discussion through an open-ended qualitative question. With a mixed methods approach, the qualitative responses were analyzed and coded to identify if participants discussed the program or tobacco (coded 1) or not (coded 0).

Statistical Analysis

We conducted statistical analyses using Stata (version 14; StataCorp LP). Considering cluster-randomization, we used multilevel generalized linear mixed effects models (GLMMs). For all GLMMs, we identified demographic characteristics that may need to be included in the models. In every GLMM, an after-school site was modeled as a random effect nested within the intervention condition, and the intervention condition and time (and their interaction) were modeled as fixed effects. GLMMs use maximum likelihood estimation, producing unbiased estimates when data are assumed to be not missing completely at random.

To conduct GLMMs using a target power of 0.85 and an effect size of 0.23 to perceived risk of vaping with an α value of .05, the estimated sample size was 45 participants [39]. We estimated that 75 adolescents would be needed to test the hypotheses, with an anticipated completion rate of approximately 60% (45/75). Considering the pilot nature of this study, this sample size was considered sufficient for the study of short-term secondary outcomes.

First, with GLMMs, we tested for any baseline differences between the 2 conditions with respect to demographic characteristics (eg, age, sex at birth, gender identity, grades at school, number of detentions, parental education level, and perceived skills playing board games). Second, with one-way ANOVA and chi-square tests, we examined attrition by testing differences between those retained and those lost to follow-up with respect to the outcome variables at baseline and other potential confounding factors.

Next, with GLMMs, we examined group differences for outcomes of interest. We conducted GLMMs predicting 5 types of outcomes. Manipulation check outcomes included perceived social interactivity, attitude toward the program, visual esthetics, and emotional involvement. Outcomes pertaining to participants' attention to the program included general attention, distraction from the program, and recognition of imagery from the program. Personal experience with the program included perceptions regarding program usability, level of fun, narrative quality in the program, program enjoyment, and creative freedom. Communication outcomes included engagement in discussions and quality of discussions. Tobacco-related outcomes included perceived risk of vaping, perceived risk of conventional tobacco use, and tobacco knowledge. Models predicting tobacco-related outcomes included group assignment, time, and the group-by-time interaction term as predictors. Following these models, we examined the role of attention to the program and personal experience factors in predicting tobacco-related outcomes.

For qualitative data, we conducted a thematic analysis of participants' responses to the open-ended question on engagement in discussion. We aimed to look for themes pertaining to tobacco, the program, or both. Next, we generated a binary variable that indicates if participants positively discussed tobacco or the program.

Results

Participants

In terms of demographics, the average age of the participants was 13.55 (SD 1.65) years, with 5% (40/74) of the participants aged ≤13 years, 58% (43/74) being female at birth, and most (56/72, 78%) being Black or African American. Approximately 37% (26/70) of the participants reported having at least 1 friend who vapes, and approximately 13% (9/69) of the participants reported having at least 1 friend who smokes a combustible product. Table 3 presents the demographic characteristics by group.

Table 3. Baseline participants' characteristics.

Characteristics	Total sample (N=74)	Storm-Heroes (n=39)	ASPIRE[a] (n=35)	*P* value
Age (y), n (%)				.96[b]
≤13	40 (56)	19 (49)	21 (64)	
>13	32 (44)	20 (51)	12 (36)	
Sex at birth, n (%)				.83[b]
Male	31 (42)	21 (54)	10 (29)	
Female	43 (58)	18 (46)	25 (71)	
Race , n (%)				.17[b]
Black or African American	56 (78)	35 (90)	21 (64)	
Not Black or African American	16 (22)	4 (10)	12 (36)	
Ethnicity, n (%)				.96[b]
Hispanic or Latinx	16 (22)	6 (15)	10 (30)	
Not Hispanic or Latinx	56 (78)	33 (85)	23 (69)	
Grades at school, n (%)				.31[b]
Mostly A	35 (48)	16 (41)	19 (56)	
Mostly B or C	38 (52)	23 (59)	15 (44)	
Parents' level of education, n (%)				.50[b]
Received a college degree	53 (73)	27 (69)	26 (76)	
Did not receive a college degree	20 (27)	12 (31)	8 (23)	
Number of detentions at school, n (%)				.45[b]
None	58 (78)	26 (67)	32 (91)	
≥1	10 (13)	9 (23)	1 (3)	
≥2	6 (8)	4 (10)	2 (6)	
Perceived board game skills, mean (SD)	3.45 (0.91)	3.49 (0.84)	3.41 (0.99)	.82[c]
Number of friends who vape, mean (SD)	3.14 (12.32)	1.29 (2.83)	5.34 (17.86)	.11[c]
Number of friends who smoke, mean (SD)	0.68 (2.54)	0.55 (2.390)	0.84 (2.76)	.52[c]
Perceived risk of vaping, mean (SD)	3.15 (0.95)	3.07 (0.97)	3.25 (0.93)	.49[c]
Perceived risk of conventional tobacco use, mean (SD)	3.31 (0.82)	3.26 (0.81)	3.37 (0.84)	.41[c]
Tobacco knowledge, mean (SD)	10.30 (3.07)	10.45 (3.19)	10.14 (2.99)	.58[c]

[a]ASPIRE: A Smoking Prevention Interactive Experience.

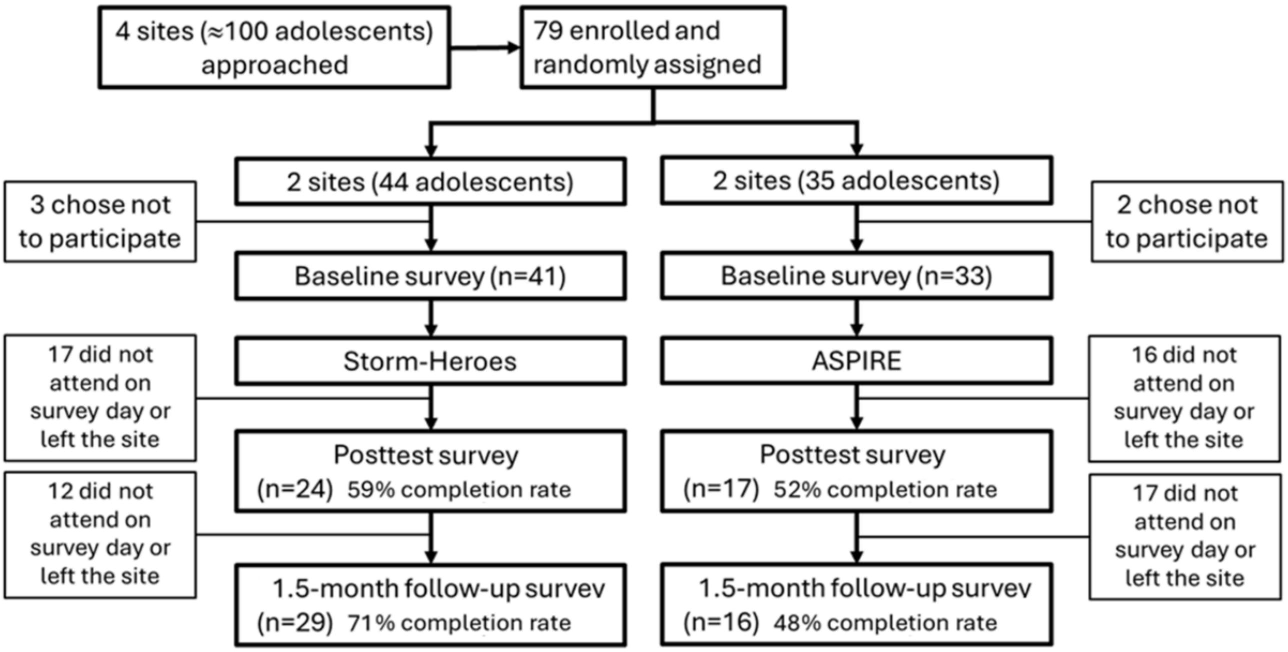
[b]Significance testing with mixed effect logistic regression to adjust for group randomization (categorical variables).

[c]Significance testing with mixed effect regression to adjust for group randomization (continuous variables).

Attrition

No harm or unintended effects occurred in this study. In this study, >100 adolescents expressed interest, and 79 enrolled, with the 4 sites randomly assigned to either Storm-Heroes or ASPIRE. In total, 94% (74/79) of the adolescents completed the baseline survey. Among baseline participants, 55% (41/74) participated in the posttest experience survey and 61% (45/74) participated in the 1.5-month follow-up survey (Figure 2). Participants who did not complete surveys had left the after-school site or did not attend the site on the day of data collection.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. Participants were allowed to participate in any survey assessment over time. ASPIRE: A Smoking Prevention Interactive Experience.



Participants in Storm-Heroes were as likely to continue to follow-up assessment as those in ASPIRE (χ^2_1=3.2; *P*=.07). There were no significant differences between participants who did and those who did not continue to the 1.5-month follow-up with respect to baseline perceived risk of vaping ($F_{1,73}$=3.74; *P*=.06), perceived risk of conventional tobacco use ($F_{1,73}$=2.43; *P*=.12), tobacco knowledge ($F_{1,58}$=2.40; *P*=.13), sex at birth (χ^2_1=0.5; *P*=.50), age ($F_{1,76}$=0.94; *P*=.34), being Black or African American (χ^2_1=2.2; *P*=.14), being Hispanic or Latinx (χ^2_1=0.1; *P*=.75), grades at school (χ^2_1=0.6; *P*=.44), parents' level of education (χ^2_1=0.2; *P*=.34), the number of friends who vape ($F_{1,68}$=0.59; *P*=.44), the number of friends who smoke ($F_{1,67}$=2.92; *P*=.09), or perceived skills in playing board games ($F_{1,70}$=0.57; *P*=.45). These data are further detailed in Multimedia Appendix 4.

Manipulation Checks

We checked to ensure that participants expressed positive attitudes toward both interventions equally. There was no significant difference between the 2 conditions with respect to attitude scores (B=0.76; *P*=.07) or visual esthetics of the program (B=1.28; *P*=.14). With both interventions being entertainment based, there was no significant difference between the conditions with respect to being emotionally involved in the content (B=0.31; *P*=.56). As expected, participants who received Storm-Heroes perceived the program to be more socially interactive than those who received ASPIRE (B=5.33; *P*=.02).

Checking for Confounders

We tested potential confounding effects of demographic characteristics. There were no significant differential effects on perceived risk of vaping as a function of perceived board game skills, race, or the number of detentions. Being younger (*P*<.001), being male (*P*=.001), being non-Hispanic (*P*=.02), having friends who vape (*P*=.001), having friends who smoke (*P*<.001), and having parents with lower education level (*P*=.04) moderated the effect of Storm-Heroes on perceived risk of vaping.

There were no significant differential effects on perceived risk of conventional tobacco use as a function of age, perceived board game skills, ethnicity, race, and number of detentions. Being male (*P*<.001), having friends who vape (*P*<.001), having friends who smoke (*P*<.001), and having parents with lower education level (*P*=.04) moderated the effect of Storm-Heroes on perceived risk of conventional tobacco use.

There were no significant differential effects on tobacco knowledge as a function of age, ethnicity, race, or the number of detentions. Having lower boardgame skills (*P*<.001), being female (*P*=.001), being non-Hispanic (*P*<.001), having friends who vape (*P*<.001), having friends who smoke (*P*=.01), and having parents with lower education level (*P*<.001) moderated the effect of Storm-Heroes on tobacco knowledge.

Personal Experience

Mixed effects models controlling for confounders showed that participants who received Storm-Heroes were significantly more likely to be distracted during the program (B=1.36; *P*=.002) and less likely to recognize images from the program (B=1.68; *P*<.001). However, they were more likely to pay attention to the program than those who received ASPIRE (B=1.30; *P*=.02). By examining the interaction between intervention groups and distraction, we found that distraction weakened the effect of Storm-Heroes on recognition of program imagery (B=–0.49; *P*=.005).

Participants who received Storm-Heroes found the program to have significantly better usability (B=0.88; *P*=.001), higher level of fun (B=4.14; *P*=.001), better narrative quality (B=2.66; *P*=.001), more enjoyment (B=2.16; *P*=.047), and more creative freedom (B=1.90; *P*=.047) than participants who received ASPIRE.

Communication Outcomes

Participants who received Storm-Heroes were significantly more likely to talk to others during the program (odds ratio [OR] 4.99, 95% CI 1.04-23.85; *P*=.04). They also experienced a better quality of peer-to-peer discussions (B=2.16; *P*=.047). According to the open-ended questions about the content of their discussions, participants who received Storm-Heroes were significantly more likely to discuss the program and the negative effects of tobacco with their peers than those who received ASPIRE (OR 5.63, 95% CI 1.25-25.29; *P*=.02). By examining the role of social interactivity, it was found that participants who found the program to be socially interactive were approximately twice as likely to talk about the program and the negative effects of tobacco (OR 1.98, 95% CI 1.28-3.07; *P*=.002).

Tobacco-Related Outcomes

Mixed effects models indicated that group allocation by time was significantly related to perceived risk of vaping (group-by-time: B=0.35; *P*=.001; Figure 3A). Participants who received Storm-Heroes were significantly more likely to exhibit a higher perceived risk of vaping at follow-up than participants who received ASPIRE, controlling for perceived risk of vaping at baseline (B=0.40; *P*=.02; Table 4).

Figure 3. Adjusted predictions of condition-by-time. Coefficients and *P* values show significance of the group-by-time interaction effect. Perceived risk measures can range between 1 and 4, while tobacco knowledge can range between 0 and 22. ASPIRE: A Smoking Prevention Interactive Experience.

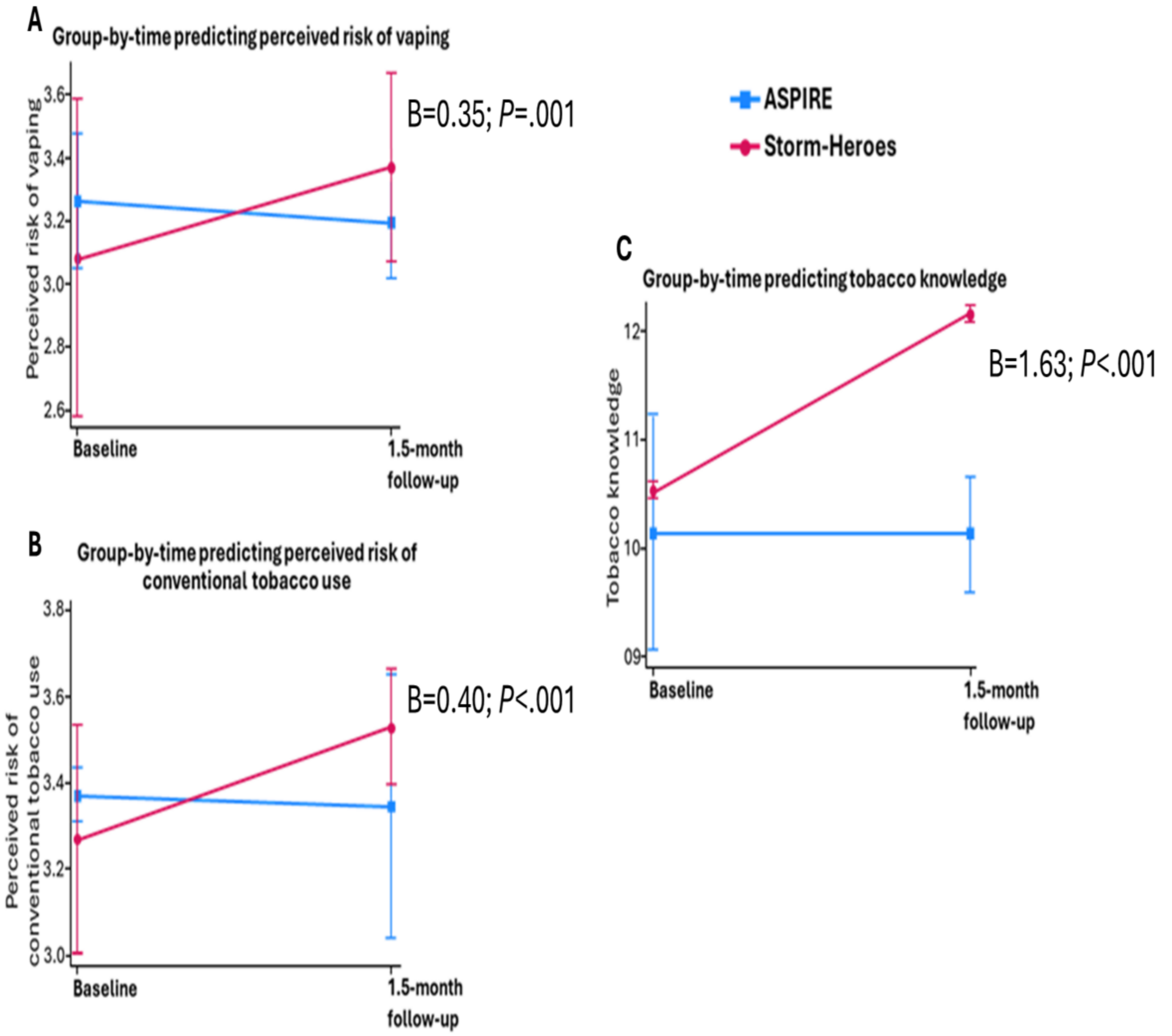


Table 4. Multilevel models predicting the perceived risk and tobacco knowledge.

	Model 1: predicting the perceived risk of vaping at follow-up (n=42)		Model 2: predicting the perceived risk of conventional tobacco use at follow-up (n=42)		Model 3: predicting tobacco knowledge at follow-up (n=19)	
	B (SE)	*P* value	B (SE)	*P* value	B (SE)	*P* value
Receiving Storm-Heroes	0.40 (0.17)	.02	0.35 (0.18)	.046	1.75 (0.56)	.002
Perceived risk of vaping at baseline	0.60 (0.17)	<.001	—[a]	—	—	—
Perceived risk of conventional tobacco use at baseline	—	—	0.68 (0.19)	<.001	—	—
Tobacco knowledge at baseline	—	—	—	—	0.53 (0.21)	.01
Number of detentions	–0.44 (0.31)	.16	–0.28 (0.15)	.06	—	—
Average grades at school	—	—	—	—	0.93 (0.21)	<.001
Parents' level of education	—	—	—	—	0.76 (1.84)	.68

[a]Not available; the models were fitted based on identified key covariates.

Participants in the Storm-Heroes condition were significantly more likely to exhibit an increase in perceived risk of conventional tobacco than participants in the ASPIRE condition (group-by-time: B=0.40; *P*<.001; Figure 3B). Participants in the Storm-Heroes condition were significantly more likely to exhibit higher perceived risk of conventional tobacco use at follow-up than participants in the ASPIRE condition, controlling for perceived risk of conventional tobacco at baseline (B=0.35; *P*=.046; Table 4).

Participants in the Storm-Heroes condition were significantly more likely to exhibit an increase in tobacco knowledge than participants in the ASPIRE condition (group-by-time: B=1.63; *P*<.001; Figure 3C). Participants in the Storm-Heroes condition were significantly more likely to exhibit higher tobacco knowledge at follow-up than participants in the ASPIRE condition, controlling for tobacco knowledge at baseline (B=0.53; *P*=.01; Table 4).

Experience Factors Predicting Tobacco-Related Outcomes

Controlling for group allocation, the results showed that the usability level of the program was related to a higher perceived risk of vaping (B=0.16; *P*=.003) and conventional tobacco use (B=0.16; *P*=.02) by follow-up. Attention to the program was also related to higher perceived risk of vaping (B=0.12; *P*=.002) and conventional tobacco (B=0.14; *P*<.001). Distraction was not related to either perceived risk of vaping (*P*=.15) or perceived risk of conventional tobacco use (*P*=.71). In contrast, both more attention (B=0.60; *P*<.001) and less distraction (B=–0.37; *P*<.001) were related to higher tobacco knowledge. A follow-up exploratory analysis of moderation indicated that distraction weakened the effect of receiving Storm-Heroes on tobacco knowledge by follow-up (group-by-distraction: B=–6.67; *P*<.001).

Discussion

Principal Findings and Comparison to Prior Work

This paper describes a pilot cluster-randomized comparative trial examining the short-term effectiveness of Storm-Heroes, a social game–based intervention, in improving secondary tobacco-related outcomes, including perceived risk of tobacco use and tobacco knowledge. The paper also presents results from adolescents' experience with the intervention and its prediction of such outcomes. We hypothesized that adolescents' engagement with Storm-Heroes would result in (1) better quality of program experience; (2) improved perceived risk of vaping and conventional tobacco use; and (3) improved tobacco knowledge compared with the engagement in ASPIRE, a nonsocial, non–game-based equivalent program.

The increased perceived risk of vaping and conventional tobacco among Storm-Heroes participants aligns with the program's goals of improving participants' awareness of the risks associated with tobacco use. With antitobacco messages designed to communicate tobacco risk, Storm-Heroes may have effectively presented the severity of tobacco-related harm. The comprehensive content in Storm-Heroes is designed with key risk communication strategies, including emotionally involving gain-framed and loss-framed messages that cover the psychological, physiological, medical, and environmental consequences of tobacco use [25]. In addition, along with other theoretical frameworks, the program design is grounded in the health belief model and empowerment theory, promoting perceived susceptibility [48] and self-efficacy [49,50]. The game-based social activities in Storm-Heroes allow adolescents to engage in interactive learning experiences that empower them to recognize and internalize the harms of tobacco, motivating them toward tobacco-free lifestyles [21].

The Storm-Heroes group showed a significant increase in tobacco knowledge scores from baseline to 1.5-month follow-up. Knowledge gained among Storm-Heroes participants compared to ASPIRE participants may be the outcome of both exposure to information within the program and increased motivation to seek information elsewhere (eg, from school teachers, the internet, etc). First, by integrating multimedia resources and a proactive learning approach, Storm-Heroes aimed to equip adolescents with comprehensive knowledge. As supported by previous research [25], the program's tobacco education content was carefully designed to cover several key topics (Multimedia

Appendix 1) and promote a holistic understanding of information from a wide array of tobacco products [25]. In addition, the gameplay aspect of Storm-Heroes encourages structured information retention that can support knowledge gain. Second, gameplay and other entertainment-based programming have been shown to stimulate interest in understanding health issues and ultimately promote health information seeking beyond the content of a program [12,51,52]. This information-seeking behavior may ultimately contribute to increased knowledge. In addition to the tobacco-related outcomes, we identified user experience differences between the 2 programs.

In the context of program experience, our results indicate that participants expressed similar positive attitudes toward both programs, with no significant difference in attitude scores or perceived visual esthetics. Emotional involvement in the content was also similar between the 2 groups. However, Storm-Heroes was perceived as more socially interactive than ASPIRE. Supportive of previous research, the similar attitude toward both interventions suggests that entertainment-based approaches, regardless of social interactivity, can effectively engage adolescents. However, Storm-Heroes was perceived as more socially interactive than ASPIRE, which aligns with previous findings indicating that interactive elements enhance program appeal [21,53,54]. As supported by the social learning theory and a systematic review of tobacco education programs, incorporating social features into interventions can promote health behavior change by fostering a sense of peer support for adolescents [55]. Our results suggest that while entertainment-based approaches effectively engage adolescents, perceived social interactivity of Storm-Heroes may play a unique role in its success.

Our results further indicated that participants in the Storm-Heroes program were more likely to engage in conversations with others during the program and experienced better-quality peer-to-peer discussions compared to those in the ASPIRE program. They were more likely to discuss the program and the harm of tobacco use. This suggests that Storm-Heroes may have included strategies that successfully encourage healthy dialogues among participants. Theoretical frameworks such as the extended elaboration likelihood model [56] support the ability of entertainment-based programming to promote healthy interpersonal discussions. This has been particularly evident when it comes to sensitive health topics such as contraceptive use, organ donation, and underage tobacco use [52,57-59]. Our results show that participants who found the program to be socially interactive were more likely to engage in healthy discussions. In line with the social learning theory, social interaction can facilitate social modeling and promote healthy learning [60,61]. Future research could further investigate the mechanisms that allow Storm-Heroes to drive these communication outcomes.

Participants who received Storm-Heroes were more likely to be distracted during the program and less likely to recall images from it. Nevertheless, they were more attentive overall compared to ASPIRE recipients. While Storm-Heroes led to more distractions and lower image recall, its higher attention levels imply deeper engagement despite potential distractions. However, the challenge lies in balancing engagement with lower distractions, as distractions may undermine the program's effectiveness. While distraction did not significantly impact perceived risk of vaping or conventional tobacco use, it was negatively associated with tobacco knowledge, thereby hindering adolescents' learning and retention of information. Future research should focus on implementation strategies to minimize distraction for Storm-Heroes. For example, using a flipped classroom approach can allow adolescents to receive tobacco-related information at home through entertaining videos and engage in social activities in class to practice what they learned [62]. This implementation strategy can reduce cognitive load, thereby optimizing engagement without compromising health education [63].

The findings support the original mechanisms of action outlined in Table 2. The significant increase in attention among participants who received Storm-Heroes highlights the effectiveness of using engaging and interactive elements such as gameplay and social interaction. Despite an increase in distraction and a lower recognition of program imagery, participants who received Storm-Heroes reported higher levels of usability, fun, narrative quality, enjoyment, and creative freedom compared to those who received ASPIRE. These factors likely contributed to enhanced perceived risk and knowledge, as suggested by the mechanisms of action. The enjoyment and narrative quality could have facilitated social interactions and discussions about tobacco, while the creative freedom and program interactivity bolstered participants' engagement and practice of tobacco-free skills. Although distraction diminished the program's impact on recognizing program imagery, the overall roles of positive reception and attention underscore the potential of Storm-Heroes to effectively leverage the social learning theory and the health belief model to promote tobacco prevention among adolescents. Considering the pilot nature of this study, we invite researchers in games for health to further explore these mechanisms.

This study advances our understanding of how game-based approaches that leverage social elements can be strategically applied to address adolescent tobacco use. While existing research has highlighted the potential role of personal engagement in games in driving health outcomes [20,64], this study distinguishes itself by focusing on social gameplay. The study specifically examines the comparative effectiveness of a socially interactive game versus a nonsocial program. By integrating social interactivity with tobacco prevention strategies, the research provides new insights into how interactive elements can enhance engagement and improve outcomes such as tobacco knowledge and perceived risks.

Unlike previous studies [26,65], which have broadly addressed game-based learning, our study delves into the unique role of multiplayer gameplay and its impact on adolescent tobacco prevention. The findings underscore the unique role of social games in fostering meaningful peer-to-peer discussions and elevating tobacco risk awareness among adolescents. Moreover, while previous research highlighted the need to boost engagement by lowering the negative effect of social influence [26,65], this study demonstrates that multiplayer gameplay can be affected by distractions, and minimizing these distractions can optimize the educational impact of social games.

Limitations

First, this study ended with a relatively low retention rate (45/74, 61%). By the time this study reached 1.5-month follow-up, adolescents were at a transition out of the after-school summer period, entering the fall semester, and ultimately, several of them were not available to continue in the study. In addition, several participants had to leave early during the data collection at the 1.5-month follow-up, which led to a low sample of participants who completed the knowledge index. This low retention rate could have influenced the study's results by reducing the generalizability and statistical power. While our use of repeated measures mixed effects modeling allowed us to account for missing data, future studies could mitigate this limitation by planning data collection at more stable periods, such as during school class sessions, and by enrolling a larger initial sample size to account for potential dropouts.

Second, although this study showed a change in short-term outcomes (ie, tobacco risk perception and knowledge), we did not examine a long-term change in tobacco use behavior. It must be noted, though, that this early pilot trial was meant to test the potential for adolescents' experience with Strom-Heroes to drive risk perception and knowledge. The lack of long-term behavioral data may limit the ability to draw conclusions about the program's effectiveness in reducing actual tobacco use over time. To address this, future research should include longer follow-up periods and assess behavioral outcomes to provide a more comprehensive evaluation of the program's impact.

Third, this study did not examine the specific types of discussions adolescents engaged in during their interactions. Without a detailed analysis of the content of these discussions, it is not possible to investigate how the program influences communication behaviors or the outcomes of these discussions. Future studies could enhance the research by incorporating qualitative methods, such as focus groups or interviews, to explore the content and context of these discussions, thereby providing deeper insights into the program's impact on social interactions.

Fourth, from an implementation perspective, the study required staff members to deliver the program to each classroom and moderate the sessions. This approach may limit the program's scalability and dissemination potential. The reliance on staff for program delivery could also introduce variability in implementation quality. To improve this, future efforts should consider adapting the procedures to allow teachers to deliver the program. This would not only facilitate broader reach but also create a more sustainable implementation model. Using Proctor's Framework for Implementation Outcomes, future studies could evaluate the program by assessing teachers' adherence to key steps, the quality of their engagement, their satisfaction, and the perceived feasibility of the program.

Implications

The results of this study suggest that Storm-Heroes can be a promising intervention for tobacco prevention. Nevertheless, we must further examine strategies that may allow us to minimize distractions while maximizing engagement to boost the success of this intervention. Once the design of this program is clear, it becomes possible to further investigate its success by examining its long-term effects on actual tobacco use. In addition, promoting peer-to-peer interactions can improve the impact of such interventions by facilitating knowledge dissemination and perceived tobacco risks. In the long run, going beyond these short-term outcomes, randomized trials with longitudinal data collection can provide valuable insights into the success of Storm-Heroes in preventing actual initiation of tobacco use and identify the factors that may promote long-term prevention outcomes. Future researchers can work to identify the specific program components and delivery methods that contribute to enhancing adolescents' experience and improving tobacco-related outcomes. In addition, by identifying effective components responsible for an improved program experience, we can design novel interventions that can be tailored to target specific groups of adolescents and address their unique needs concerning different tobacco products.

Acknowledgments

This study was supported by the National Institute on Drug Abuse of the National Institutes of Health (R00DA044277; principal investigator: GK). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We used generative artificial intelligence (ChatGPT, OpenAI) for proofreading the readability of the Limitation section, ensuring human oversight and control throughout the process. The authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to security purposes requested by the academic institution but are available from the corresponding author on reasonable request.

Authors' Contributions

GK is responsible for the design of the study, and NR and PB provided guidance on the design of the study. GK was responsible for data curation, validation, visualization, preparing the original draft, and reviewing and editing the draft. MK and ER participated in data curation, validation, and implementation of the study. GK, ER, MK, BZ, NR, and PB were responsible for preparing the original draft and reviewing and editing the draft. GK, ER, MK, BZ, NR, and PB participated in reviewing and editing the draft. All authors read and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A description of the interventions based on the Template for Intervention Description and Replication (TIDieR) checklist.
[DOCX File , 38 KB - games_v12i1e63296_app1.docx]

Multimedia Appendix 2
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist.
[PDF File (Adobe PDF File), 1512 KB - games_v12i1e63296_app2.pdf]

Multimedia Appendix 3
Main study measures.
[DOCX File , 18 KB - games_v12i1e63296_app3.docx]

Multimedia Appendix 4
Participant characteristics based on attrition.
[DOCX File , 20 KB - games_v12i1e63296_app4.docx]

References

1. Birdsey J, Cornelius M, Jamal A, Park-Lee E, Cooper MR, Wang J, et al. Tobacco product use among U.S. middle and high school students - national youth tobacco survey, 2023. MMWR Morb Mortal Wkly Rep 2023 Nov 03;72(44):1173-1182 [FREE Full text] [doi: 10.15585/mmwr.mm7244a1] [Medline: 37917558]
2. Pérez A, N'hpang RS, Callahan E, Bluestein M, Kuk AE, Chen B, et al. Age at initiation of cigarette use in a nationally representative sample of US youth, 2013-2017. JAMA Netw Open 2021 Feb 01;4(2):e210218. [doi: 10.1001/jamanetworkopen.2021.0218] [Medline: 33635326]
3. Richmond-Rakerd LS, Slutske WS, Wood PK. Age of initiation and substance use progression: a multivariate latent growth analysis. Psychol Addict Behav 2017 Sep;31(6):664-675 [FREE Full text] [doi: 10.1037/adb0000304] [Medline: 28805408]
4. Albaugh MD, Ottino-Gonzalez J, Sidwell A, Lepage C, Juliano A, Owens MM, et al. Association of cannabis use during adolescence with neurodevelopment. JAMA Psychiatry 2021 Jun 16;78(9):1-11 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.1258] [Medline: 34132750]
5. Hansen K, Lindström M, Rosvall M. Age at smoking initiation and self-rated health among second grade high school boys and girls in Scania, Sweden, a cross-sectional study. BMC Public Health 2015 Nov 18;15:1143 [FREE Full text] [doi: 10.1186/s12889-015-2457-z] [Medline: 26581335]
6. State fact sheets. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/tobacco/stateandcommunity/state-fact-sheets/ [accessed 2024-10-21]
7. Youth and tobacco use. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/tobacco/php/data-statistics/youth-data-tobacco/?CDC_AAref_Val=https://www.cdc.gov/tobacco/data_statistics/fact_sheets/youth_data/tobacco_use/index.htm [accessed 2024-10-21]
8. Singhal A, Cody MJ, Rogers EM, Sabido M. Entertainment-Education and Social Change: History, Research, and Practice. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2004.
9. Khalil GE. When losing means winning: the impact of conflict in a digital game on young adults' intentions to get protected from cancer. Games Health J 2012 Aug;1(4):279-286. [doi: 10.1089/g4h.2012.0017] [Medline: 26191631]
10. Khalil GE. Fear and happiness in “re-mission”: teasing out emotional gaming events responsible for cancer risk perception. In: Proceedings of the 3rd European Conference on Gaming and Playful Interaction in Health Care. 2013 Presented at: GFHEU 2013; October 28-29, 2013; Amsterdam, The Netherlands. [doi: 10.1007/978-3-658-02897-8_3]
11. Watson AM, Alber JM, Barnett TE, Mercado R, Bernhardt JM. Content analysis of anti-tobacco videogames: characteristics, content, and qualities. Games Health J 2016 Jun;5(3):216-223. [doi: 10.1089/g4h.2015.0096] [Medline: 27096576]
12. Khalil GE, Beale IL, Chen M, Prokhorov AV. A video game promoting cancer risk perception and information seeking behavior among young-adult college students: a randomized controlled trial. JMIR Serious Games 2016 Jul 28;4(2):e13 [FREE Full text] [doi: 10.2196/games.5793] [Medline: 27470927]
13. Bohr Y, Litwin L, Hankey JR, McCague H, Singoorie C, Lucassen MF, et al. Evaluating the utility of a psychoeducational serious game (SPARX) in protecting Inuit youth from depression: pilot randomized controlled trial. JMIR Serious Games 2023 Mar 09;11:e38493 [FREE Full text] [doi: 10.2196/38493] [Medline: 36892940]

14. Fijačko N, Masterson Creber R, Metličar Š, Strnad M, Greif R, Štiglic G, et al. Effects of a serious smartphone game on nursing students' theoretical knowledge and practical skills in adult basic life support: randomized wait list-controlled trial. JMIR Serious Games 2024 Apr 05;12:e56037 [FREE Full text] [doi: 10.2196/56037] [Medline: 38578690]
15. Liu B, Wan X, Li X, Zhu D, Liu Z. An augmented reality serious game for children's optical science education: randomized controlled trial. JMIR Serious Games 2024 Feb 01;12:e47807 [FREE Full text] [doi: 10.2196/47807] [Medline: 38329284]
16. Pentz MA, Hieftje KD, Pendergrass TM, Brito SA, Liu M, Arora T, et al. A videogame intervention for tobacco product use prevention in adolescents. Addict Behav 2019 Apr;91:188-192 [FREE Full text] [doi: 10.1016/j.addbeh.2018.11.016] [Medline: 30477819]
17. Derksen ME, van Strijp S, Kunst AE, Daams JG, Jaspers MW, Fransen MP. Serious games for smoking prevention and cessation: a systematic review of game elements and game effects. J Am Med Inform Assoc 2020 May 01;27(5):818-833 [FREE Full text] [doi: 10.1093/jamia/ocaa013] [Medline: 32330255]
18. Guo JL, Hsu HP, Lai TM, Lin ML, Chung CM, Huang CM. Acceptability evaluation of the use of virtual reality games in smoking-prevention education for high school students: prospective observational study. J Med Internet Res 2021 Sep 28;23(9):e28037 [FREE Full text] [doi: 10.2196/28037] [Medline: 34581679]
19. Weser VU, Duncan LR, Sands BE, Schartmann A, Jacobo S, François B, et al. Evaluation of a virtual reality E-cigarette prevention game for adolescents. Addict Behav 2021 Nov;122:107027. [doi: 10.1016/j.addbeh.2021.107027] [Medline: 34225030]
20. Hieftje KD, Fernandes CS, Lin IH, Fiellin LE. Effectiveness of a web-based tobacco product use prevention videogame intervention on young adolescents' beliefs and knowledge. Subst Abus 2021;42(1):47-53. [doi: 10.1080/08897077.2019.1691128] [Medline: 31825759]
21. Khalil GE, Kim J, McLean D, Ramirez E, Zhao B, Salloum RG. Identifying adolescents' gaming preferences for a tobacco prevention social game: a qualitative study. PLoS One 2023 Jul 28;18(7):e0289319 [FREE Full text] [doi: 10.1371/journal.pone.0289319] [Medline: 37506112]
22. Orji R, Vassileva J, Mandryk R. Towards an effective health interventions design: an extension of the health belief model. Online J Public Health Inform 2012;4(3):ojphi.v4i3.4321 [FREE Full text] [doi: 10.5210/ojphi.v4i3.4321] [Medline: 23569653]
23. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
25. Khalil GE, McLean D, Ramirez E, Mihaj PP, Zhao B, Dhar B, et al. Developing a text-message library for tobacco prevention among adolescents: a qualitative study. PLoS One 2024;19(1):e0296503 [FREE Full text] [doi: 10.1371/journal.pone.0296503] [Medline: 38175835]
26. Tudor-Sfetea C, Rabee R, Najim M, Amin N, Chadha M, Jain M, et al. Evaluation of two mobile health apps in the context of smoking cessation: qualitative study of cognitive behavioral therapy (CBT) versus non-CBT-based digital solutions. JMIR Mhealth Uhealth 2018 Apr 18;6(4):e98. [doi: 10.2196/mhealth.9405] [Medline: 29669708]
27. Kim Y, Kang S, Vongjaturapat N. Application of transtheoretical model to explain adolescents' smoking behavior. J Subst Use 2021 Dec 22;28(2):200-205. [doi: 10.1080/14659891.2021.2019328]
28. Brown EC, Sumichrast RT. Evaluating performance advantages of grouping genetic algorithms. Eng Appl Artif Intell 2005 Feb;18(1):1-12. [doi: 10.1016/j.engappai.2004.08.024]
29. Krouska A, Virvou M. An enhanced genetic algorithm for heterogeneous group formation based on multi-characteristics in social-networking-based learning. IEEE Trans Learning Technol 2020;13(3):465-476. [doi: 10.1109/tlt.2019.2927914]
30. Prokhorov AV, Kelder SH, Shegog R, Conroy JL, Murray N, Peters R, et al. Project ASPIRE: an interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high school students. Subst Use Misuse 2010 May;45(6):983-1006. [doi: 10.3109/10826080903038050] [Medline: 20397881]
31. McCarthy M. Experiential learning theory: from theory to practice. J Bus Econ Res 2010 May 01;8(5):131-140. [doi: 10.19030/jber.v8i5.725]
32. Bandura A. Social Learning Theory. New York, NY: General Learning Press; 1977.
33. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]
34. Khalil GE, Wang H, Calabro KS, Mitra N, Shegog R, Prokhorov AV. From the experience of interactivity and entertainment to lower intention to smoke: a randomized controlled trial and path analysis of a web-based smoking prevention program for adolescents. J Med Internet Res 2017 Feb 16;19(2):e44 [FREE Full text] [doi: 10.2196/jmir.7174] [Medline: 28209560]
35. Prokhorov AV, Kelder SH, Shegog R, Murray N, Peters RJ, Agurcia-Parker C, et al. Impact of a smoking prevention interactive experience (ASPIRE), an interactive, multimedia smoking prevention and cessation curriculum for culturally diverse high-school students. Nicotine Tob Res 2008 Sep;10(9):1477-1485. [doi: 10.1080/14622200802323183] [Medline: 19023839]

36. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
37. Shete SS, Wilkinson AV. Identifying demographic and psychosocial factors related to the escalation of smoking behavior among Mexican American adolescents. Prev Med 2017 Jun;99:146-151 [FREE Full text] [doi: 10.1016/j.ypmed.2017.02.018] [Medline: 28235542]
38. Johnston LD, O'Malley PM, Bachman JG. Monitoring the future: national results on adolescent drug use: overview of key findings. Focus 2003 Apr;1(2):213-234. [doi: 10.1176/foc.1.2.213]
39. Prokhorov AV, Calabro KS, Arya A, Russell S, Czerniak KW, Botello GC, et al. Mobile text messaging for tobacco risk communication among young adult community college students: randomized trial of project debunk. JMIR Mhealth Uhealth 2021 Nov 24;9(11):e25618 [FREE Full text] [doi: 10.2196/25618] [Medline: 34822339]
40. Coursaris CK, Sung J. Antecedents and consequents of a mobile website's interactivity. New Media Soc 2012 Apr 26;14(7):1128-1146. [doi: 10.1177/1461444812439552]
41. Brunet J, Forest-Boucher A, Fredette M. Brand placement in virtual environments: impact of flow on brand recognition and recall. Les Cahiers du GERAD. 2008 Nov. URL: https://www.researchgate.net/publication/237746158_Brand_Placement_in_Virtual_Environments_Impact_of_Flow_on_Brand_Recognition_and_Recall [accessed 2024-09-21]
42. Phan MH, Chaparro BS, Keebler JR, Shelstad WJ, Smith DC. Validation of the GUESS-18: a short version of the Game User Experience Satisfaction Scale (GUESS). J User Exp 2020;16(1):49-62. [doi: 10.1037/t58113-000]
43. Phan MH, Keebler JR, Chaparro BS. The development and validation of the Game User Experience Satisfaction Scale (GUESS). Hum Factors 2016 Dec 27;58(8):1217-1247. [doi: 10.1177/0018720816669646] [Medline: 27647156]
44. Hess U, Kleck RE. Differentiating emotion elicited and deliberate emotional facial expressions. Eur J Soc Psychol 1990;20(5):369-385. [doi: 10.1002/ejsp.2420200502]
45. Collins RL, Eberhart NK, Estrada-Darley I, Roth E. Evaluation of Los Angeles County's fall 2020 WhyWeRise mental health campaign and WeRise 2021. RAND. 2022 Mar 10. URL: https://www.rand.org/pubs/research_reports/RRA875-2.html [accessed 2024-09-21]
46. Davis K, Shafer PR, Rodes R, Kim A, Hansen H, Patel D, et al. Does digital video advertising increase population-level reach of multimedia campaigns? Evidence from the 2013 tips from former smokers campaign. J Med Internet Res 2016 Sep 14;18(9):e235 [FREE Full text] [doi: 10.2196/jmir.5683] [Medline: 27627853]
47. Duke JC, Alexander TN, Zhao X, Delahanty JC, Allen JA, MacMonegle AJ, et al. Youth's awareness of and reactions to the real cost national tobacco public education campaign. PLoS One 2015;10(12):e0144827 [FREE Full text] [doi: 10.1371/journal.pone.0144827] [Medline: 26679504]
48. Mohammadi S, Ghajari H, Valizade R, Ghaderi N, Yousefi F, Taymoori P, et al. Predictors of smoking among the secondary high school boy students based on the health belief model. Int J Prev Med 2017;8(1):24. [doi: 10.4103/ijpvm.ijpvm_264_16]
49. Park E, Chang YP. Using digital media to empower adolescents in smoking prevention: mixed methods study. JMIR Pediatr Parent 2020 Mar 31;3(1):e13031 [FREE Full text] [doi: 10.2196/13031] [Medline: 32229463]
50. Trisnowati H, Ismail D, Padmawati RS. Health promotion through youth empowerment to prevent and control smoking behavior: a conceptual paper. Health Educ 2021 Feb 09;121(3):275-294. [doi: 10.1108/he-09-2020-0092]
51. Beale IL, Kato PM, Marin-Bowling VM, Guthrie N, Cole SW. Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. J Adolesc Health 2007 Sep;41(3):263-270. [doi: 10.1016/j.jadohealth.2007.04.006] [Medline: 17707296]
52. Khalil GE, Rintamaki LS. A televised entertainment-education drama to promote positive discussion about organ donation. Health Educ Res 2014 Apr 07;29(2):284-296 [FREE Full text] [doi: 10.1093/her/cyt106] [Medline: 24399264]
53. Hickey AM. An adolescents peer-led motivational interviewing intervention conducted in low socio-economic community youth organisations to change health risk behaviours (smoking, alcohol and sedentary behaviours). Trinity College Dublin. 2024. URL: https://www.tara.tcd.ie/handle/2262/108280?show=full [accessed 2024-09-21]
54. Khalil GE, Prokhorov AV. Friendship influence moderating the effect of a web-based smoking prevention program on intention to smoke and knowledge among adolescents. Addict Behav Rep 2021 Jun;13:100335 [FREE Full text] [doi: 10.1016/j.abrep.2020.100335] [Medline: 33521230]
55. Peirson L, Ali MU, Kenny M, Raina P, Sherifali D. Interventions for prevention and treatment of tobacco smoking in school-aged children and adolescents: a systematic review and meta-analysis. Prev Med 2016 Apr;85:20-31. [doi: 10.1016/j.ypmed.2015.12.004] [Medline: 26743631]
56. Moyer-Gusé E, Chung AH, Jain P. Identification with characters and discussion of taboo topics after exposure to an entertainment narrative about sexual health. J Commun 2011 Jun;61(3):387-406. [doi: 10.1111/j.1460-2466.2011.01551.x]
57. Boulay M, Storey JD, Sood S. Indirect exposure to a family planning mass media campaign in Nepal. J Health Commun 2002;7(5):379-399. [doi: 10.1080/10810730290001774] [Medline: 12455760]
58. Blitstein JL, Evans WD, Davis KC, Kamyab K. Repeated exposure to media messages encouraging parent-child communication about sex: differential trajectories for mothers and fathers. Am J Health Promot 2012;27(1):43-51. [doi: 10.4278/ajhp.110302-QUAN-95] [Medline: 22950925]

59. Papa M, Singhal A, Law S, Pant S. Entertainment-education and social change: an analysis of parasocial interaction, social learning, collective efficacy, and paradoxical communication. J Commun 2000;50(4):31-55. [doi: 10.1111/j.1460-2466.2000.tb02862.x]
60. Hill JR, Song L, West RE. Social learning theory and web-based learning environments: a review of research and discussion of implications. Am J Distance Educ 2009 May 19;23(2):88-103. [doi: 10.1080/08923640902857713]
61. Murray FB. Learning and development through social interaction and conflict: a challenge to social learning theory. In: Piaget and the Foundations of Knowledge. London, UK: Psychology Press; 1983.
62. Hew KF, Bai S, Dawson P, Lo CK. Meta-analyses of flipped classroom studies: a review of methodology. Educ Res Rev 2021 Jun;33:100393. [doi: 10.1016/j.edurev.2021.100393]
63. Amorim MB, Moraes AR, Souza DK, Coêlho EC, Vieira GB, Oliveira GP, et al. Benefits of the flippedclassroom in health education -a systematic review. Braz J Health Rev 2020 Dec 02;3(6):17428-17437. [doi: 10.34119/bjhrv3n6-160]
64. Damaševičius R, Maskeliūnas R, Blažauskas T. Serious games and gamification in healthcare: a meta-review. Information 2023 Feb 07;14(2):105. [doi: 10.3390/info14020105]
65. El-Hilly A, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game on? Smoking cessation through the gamification of mHealth: a longitudinal qualitative study. JMIR Serious Games 2016 Oct 24;4(2):e18. [doi: 10.2196/games.5678] [Medline: 27777216]

Abbreviations

ASPIRE: A Smoking Prevention Interactive Experience
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
GLMM: generalized linear mixed effects model
OR: odds ratio
TIDieR: Template for Intervention Description and Replication

Edited by T Leung; submitted 15.06.24; peer-reviewed by A Eisingerich, T Leonard; comments to author 01.08.24; revised version received 08.08.24; accepted 13.08.24; published 05.11.24.

Please cite as:
Khalil G, Ramirez E, Khan M, Zhao B, Ribeiro N, Balian P
Risk Perception and Knowledge Following a Social Game–Based Tobacco Prevention Program for Adolescents: Pilot Randomized Comparative Trial
JMIR Serious Games 2024;12:e63296
URL: https://games.jmir.org/2024/1/e63296
doi:10.2196/63296
PMID:

The Impact and Acceptance of Gamification by Learners in a Digital Literacy Course at the Undergraduate Level: Randomized Controlled Trial

Abeer Alnuaim, BSc, MSc, PhD

Department of Computer Science and Engineering, College of Applied Studies and Community Service, King Saud University, P.O. Box 22459, Riyadh, Saudi Arabia

Corresponding Author:
Abeer Alnuaim, BSc, MSc, PhD

Abstract

Background: In recent years, the integration of technology in education has revolutionized traditional learning paradigms. Digital literacy, a crucial skill in the 21st century, has become a vital aspect of modern education, enabling students to navigate, critically assess, and effectively use digital tools. As educators strive to boost engagement and learning outcomes, gamification has appeared as an auspicious pedagogical approach. By applying game mechanics to nongame contexts, gamification seeks to create a more immersive and digital learning experience.

Objective: This research paper aims to investigate the impact and acceptance of gamification by learners in a digital literacy course at the undergraduate level.

Methods: In a pre-post intervention study, 168 undergraduate students were randomly assigned either to the experimental group (gamification based) or control group (conventional) learning condition. Both groups of participants learned the same topics in digital literacy.

Results: Empirical findings showed that participants from the experimental group had better academic performance in digital literacy than those who were not exposed to the game-based learning environment. The participants' prior experience with gamification was not found to be a significant predictor of their acceptance of gamification in a digital literacy course.

Conclusions: The study provides evidence supporting the potential benefits of gamification in enhancing digital literacy education and opens the door for further exploration and implementation of gamified learning approaches in higher education settings.

(JMIR Serious Games 2024;12:e52017) doi:10.2196/52017

KEYWORDS

gamification; games; technology integration; information literacy; technology acceptance

Introduction

Background

In today's interconnected world, the demand for digital literacy has intensified across academic, professional, and personal domains. Digital literacy is a set of skills essential for 21st century individuals to use digital tools to support the achievement of goals in their life situations [1,2]. It has become not only a key factor in enabling participation in education and employment but also a means of interacting with the world. It encompasses the ability to access, analyze, evaluate, and create information using various digital applications. In the context of undergraduate education, cultivating digital literacy is essential for students to excel in their studies, conduct research, and adapt to the demands of the workforce in the information era. As universities strive to equip their students with these vital skills, innovative teaching approaches that enhance engagement and knowledge retention are warranted [3].

At the same time, incorporating gamification into pedagogical approaches has gained considerable attention [4,5]. The term gamification first appeared in 2008 and received growing significance since the 2010s [6]. In general, gamification refers to a process of augmenting services with (motivational) affordances to raise gameful experiences and promote behavioral outcomes [7]. In contrast to games, gamification is described by its serious purpose. It involves applying game design elements and mechanics in nongame contexts to enhance user engagement, motivation, and learning outcomes. According to Trinidad et al [8], gamification has swiftly appeared as one of the preferred persuasive technologies widely used with the aim of stimulating a positive change in the user's behavior through game-like elements in nongame contexts. Likewise, Krath et al [6] argued that gamification is a great way to demonstrate goals and their germaneness, push users through directed tracks, provide users instant feedback, strengthen good performance, and streamline content to manageable tasks. It leverages the intrinsic motivational elements found in games to create a positive and engaging learning environment. Hamari et al [7] suggest that gamification has beneficial impacts; however, these effects greatly rely on the context in which it is used as well as on the individuals who practice it. Similar to this, Huang et al [9] argue that it can be difficult for educational academics and practitioners to decide when and how to apply gamification design elements.

Given the significant traction gained by gamification in recent years, several state-of-the-art gamified solutions and approaches have been developed and tested [10]. One widely used example of gamification for education is Kahoot. Kahoot! is a game-based learning platform that allows educators to create and deliver quizzes in a game format [11]. It incorporates leaderboards, points, and real-time feedback to create a competitive and engaging learning environment. Research indicates that Kahoot! enhances student participation and knowledge retention [12]. Similarly, Duolingo is a language-learning platform that uses gamification to teach foreign languages. It uses elements such as skill trees, streaks, and in-game currency to motivate learners. Studies have found that Duolingo is effective in improving language proficiency and maintaining learner interest [13].

Gamification is significantly correlated with game-based learning. Game-based learning is defined as the achievement of distinct learning objectives through game content and play and augmenting learning by including problem-solving spaces and challenges that offer learners, who are also players, with a feeling of achievement [6]. By incorporating game design elements such as points, badges, levels, leaderboards, and immediate feedback, gamified learning experiences can stimulate students' curiosity and foster a sense of accomplishment [5,14]. Through gamification, educators aim to increase student motivation, participation, and knowledge retention by transforming learning from a passive experience into an active and enjoyable process [15,16]. Hamari and Homner have also recognized it as a promising method for instructional contexts for its motivational power [7]. The application of gamification has been successful in various educational contexts [9], including language learning [17,18], mathematics [19,20], and computer programming [21]. Its use in nongame contexts is associated with impacts on motivation, behavior (eg, academic achievement and engagement), and cognitive learning [6]. All of this demonstrates the potential of gamification to augment a digital literacy learning experience as well.

Even though recent research has made significant strides in this area, additional information about the integration of game aspects into educational materials is still required. Particularly, there is a dearth of coherent understanding on its use in the subject area of digital literacy. Particularly, there is a need to explore how adult learners perceive and accept gamification in the context of a digital literacy course at the undergraduate level in Saudi Arabia. By investigating the impact and acceptance of gamification in a digital literacy course at the undergraduate level, this research aims to shed light on its efficacy and potential for cultivating essential digital literacy. Addressing this gap in knowledge will provide valuable insights into the factors that influence Saudi students' acceptance of gamification and the impact it has on their learning experience, thereby enabling educators and course designers to make informed decisions regarding the effective implementation of gamification strategies.

In the next section, we discuss the theoretical foundations for this study that helped us formulate our research questions and hypotheses.

Theoretical Framework

In recent years, scientific papers have progressively investigated the use of different theoretical foundations such as motivation, behavior, and learning theories to explain the effects of certain gamification elements [6]. The theoretical footings of this study align with 2 pronounced frameworks in the field of education and technology: the self-determination theory (SDT) and the technology acceptance model (TAM). These theoretical foundations provide a comprehensive lens through which to understand the impact and acceptance of gamification in the context of digital literacy education at the undergraduate level. A brief overview of these theories and justification how they are aligned with this study is provided in the following sections.

Self-Determination Theory

Developed by Deci and Ryan, the SDT posits that individuals have innate psychological needs for autonomy, competence, and relatedness, which serve as essential motivators for behavior and engagement [22]. In the context of education, SDT suggests that learners are more likely to be intrinsically motivated and experience greater satisfaction and well-being when their psychological needs for autonomy, competence, and relatedness are supported [22].

The integration of gamification in educational settings can be viewed through the lens of the SDT, as it has the potential to fulfill learners' psychological needs helping them experience intrinsic motivation, which drives them to engage in activities for the sheer enjoyment and interest in the task itself [22]. By providing learners with autonomy through choice and control over their learning paths, gamified experiences empower students to take ownership of their learning. Further, the challenge-based nature of game-based experiences foster feelings of competence as learners strive to achieve goals and overcome obstacles within the game environment. In addition, the social elements inherent in many gamified platforms facilitate a sense of relatedness by promoting collaboration, competition, and community among learners [6].

SDT is one of the well-known theories in the context of gamification. In the context of this study, the impact of gamification on learning outcomes can be better understood in terms of its alignment with the principles of SDT. Accordingly, by leveraging game mechanics to enhance autonomy, competence, and relatedness, gamification has the potential to promote intrinsic motivation and engagement [7] among undergraduate students in a digital literacy course.

Technology Acceptance Model

A popular framework for evaluating people's attitudes and behavioral intentions toward embracing new technologies is Davis's TAM [23]. According to TAM, perceived usefulness and perceived ease of use are key determinants of any individual's intention to use a technology, which eventually influences his or her actual usage behavior. Perceived usefulness refers to the extent to which a person believes that using a particular technology will enhance their performance or productivity, while perceived ease of use pertains to the degree of effort required to use the technology effectively [24].

TAM has been extensively used in the context of educational technology to examine learners' attitudes and behaviors toward various digital tools and platforms. Researchers can learn more about the elements influencing students' acceptance and engagement with these technologies by evaluating how useful and simple they believe gamified learning environments to be [24,25]. In this study, TAM provides a theoretical framework for understanding learners' acceptance of gamification in a digital literacy course. By investigating the perceived usefulness and ease of use of gamified learning experiences, the study seeks to elucidate the factors that contribute to students' willingness to engage with and embrace gamification as a pedagogical approach.

By integrating the SDT and TAM frameworks into the research design and analysis, this study aims to provide a comprehensive understanding of the impact and acceptance of gamification by learners in a digital literacy course at the undergraduate level. In aligning our research design with the constructs of SDT, we focused on measuring students' intrinsic motivation and perceived autonomy in the learning process. Similarly, the TAM informed our analysis of students' acceptance of gamified learning tools, focusing on perceived ease of use and usefulness as key determinants. Furthermore, in interpreting our findings, we draw on both frameworks to discuss theoretical and practical implications of the study.

Research Gap and Rationale

The growing significance of digital literacy in the modern world necessitates effective and engaging teaching strategies. Gamification has arisen as an encouraging approach that leverages the power of games to enhance learning experiences. While gamification shows promise as an innovative pedagogical approach, its impact and acceptance specifically in the context of digital literacy courses at the undergraduate level remain relatively unexplored. Existing research primarily focuses on the effects of gamification in K-12 education or specialized domains [6,26,27]. For example, Dehghanzadeh et al [25] concentrated on gamification-supported learning in K-12 settings, and Tan et al [27] described several mathematics gamification instances to enrich algebra teaching at school levels. Thus, there is a need to delve into the unique challenges and opportunities of implementing gamification in higher education, particularly in digital literacy courses. By investigating the impact of gamification on learning outcomes, the study can offer insights into the effectiveness of this approach in improving digital literacy among undergraduate students. Moreover, understanding students' perceptions and acceptance of gamification can provide crucial feedback for educators and instructional designers to refine gamified learning experiences in digital literacy courses. The study's outcomes may also inform policy decisions regarding the integration of gamification and technology in higher education curricula, paving the way for more engaging and effective teaching practices in the digital age.

Research Questions

This study seeks to contribute to the existing body of knowledge by examining the following research questions:

1. What is the impact of gamification on the learning outcomes of students in a digital literacy course at the undergraduate level?
2. What factors affect the acceptance of learners toward gamified learning experiences in a digital literacy course?
3. How does learners' prior experience with gamified learning environments affect their acceptance of gamification in a digital literacy course at the undergraduate level?

Based on the above research questions, the following hypotheses are proposed:

- H1: Gamification has a significant positive effect on learning outcomes of students in a digital literacy course at the undergraduate level.
- H2: Age and the major of students' degree programs are significant factors influencing learners' acceptance of gamification in a digital literacy class.
- H3: Learners' prior experience with gamified learning environments has a significant effect on their acceptance of gamification in a digital literacy course.

Methods

Participants

Using convenience sampling, this study involved a diverse group of 168 undergraduate students enrolled in a digital literacy course at a Saudi university. While convenience sampling may not provide a representative sample of the population, it is the most practical in small-scale studies or when studying hard-to-reach populations [28]. Likewise, this study justifies this approach by emphasizing the need to assess the initial impact of a specific pedagogical approach in a real-world setting before considering broader applications or generalizations. All of these participants were native Arabic speakers. The students' age ranged from 18 to 22 years with an average of 20.6 years.

Intervention students participated in the 16-week digital literacy course. The course is offered to the students enrolled in various undergraduate programs. The participants were distributed into 2 groups for a true experimental design [29,30]. There were 84 students in the experimental group and 84 in the control group. Participants in this study were randomly assigned to either the control group or the experimental group using a computerized randomization procedure.

The process of randomization of the participants into the 2 groups started with the creation of a list of all participating students, including their names and class roll numbers. A random number generator was used to create a sequence of random numbers corresponding to the total number of participants (N=168). This sequence was generated using a computer software program designed for randomization to eliminate any potential bias. The first 84 numbers in the sequence were assigned to the experimental group, and the remaining 84 numbers were assigned to the control group. The allocation process was conducted by a researcher who was not involved in the instructional process to ensure blinding. This helped minimize any potential bias in group assignment. This approach ensured that each participant had an equal chance of being placed in either group, thereby minimizing potential biases and confounding variables. By implementing random assignment, the study aimed to create comparable groups, ensuring that any observed differences in digital literacy skills and engagement could be attributed to the gamified intervention rather than to pre-existing differences between the participants.

Experiment Process

The teaching process took place partially in a lecture room and a computer lab. The duration of each lesson was 1 hour. The whole experiment lasted for 16 weeks. In order to ensure that participants in the experimental (game-based) and control groups (nongame based or conventional) had the same knowledge, participants in both groups were asked to appear in a pretest. All the participants in both groups received teaching through conventional teaching methods until week 10 of the semester. However, the experimental and control groups were taught using 2 diverse teaching methods over the next 6 weeks (weeks 11 - 16). During these 6 weeks, the participants in the experimental group experienced gamification to facilitate their learning of the topics taught during this period (see Figure 1 and Figure 2 to view the snapshots of one of the games used to teach cybersecurity), while the participants in the control group continued their learning of the same topics through conventional teaching methods.

After the 16-week teaching-learning process, both groups took an achievement test for the full term with some questions dedicated to assessing the topics covered in weeks 11-16 (the intervention time when the 2 groups of participants were taught using different teaching strategies, ie, game based and conventional). At the same time, a survey questionnaire was administered in the experimental group to assess their acceptability of gamification to support their learning in a digital literacy course (see Figure 3 for the graphical flow of the experimental process of this study). The study was authorized by the ethics committee of the corresponding author's university. All the participants provided written informed consent before their participation in the study, and they did not receive any monetary or other sort of compensation.

Figure 1. Snapshot 1 of “cybersecurity lab,” a cybersecurity game.

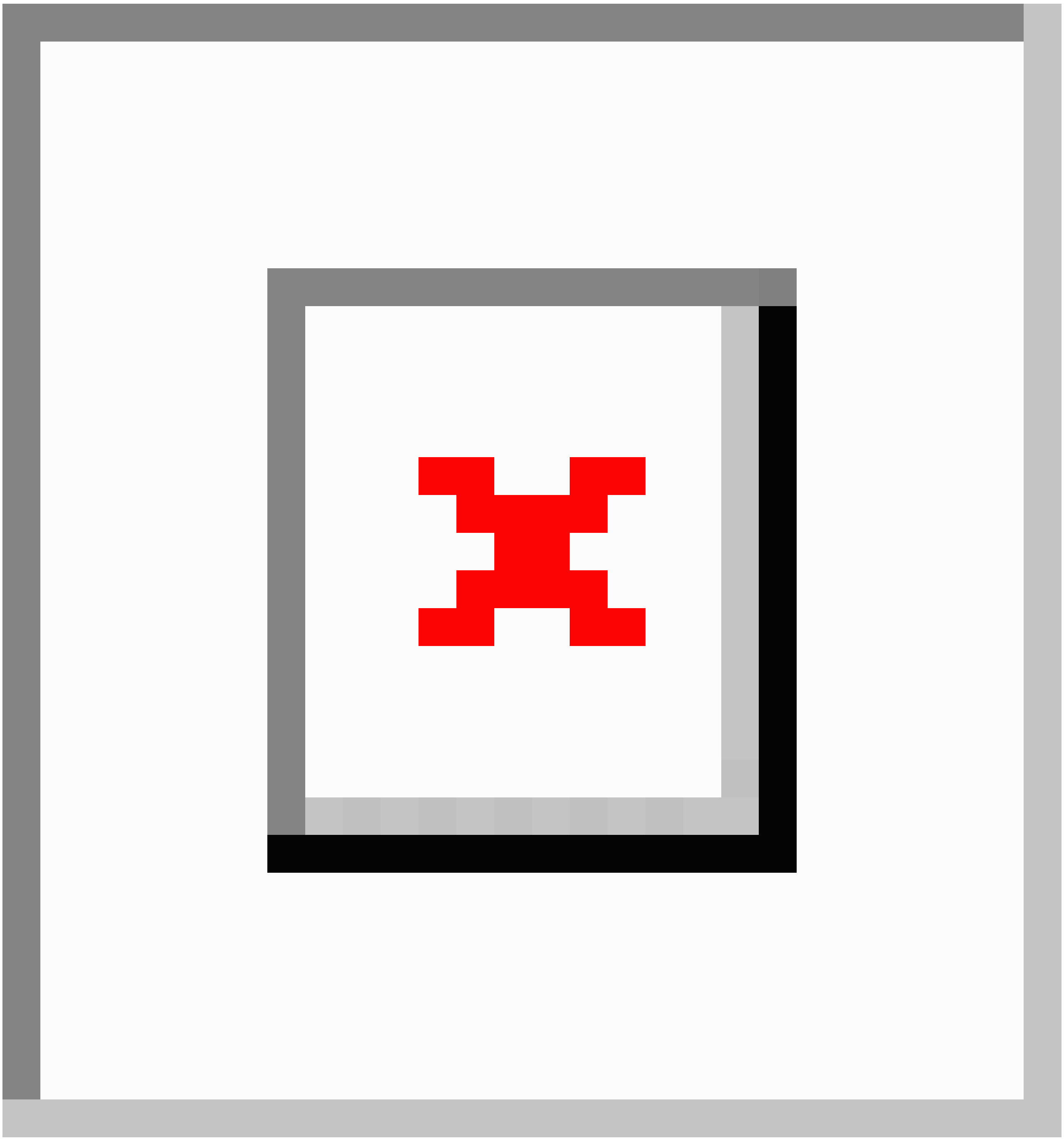


Figure 2. Snapshot 2 of "cybersecurity lab," a cybersecurity game.

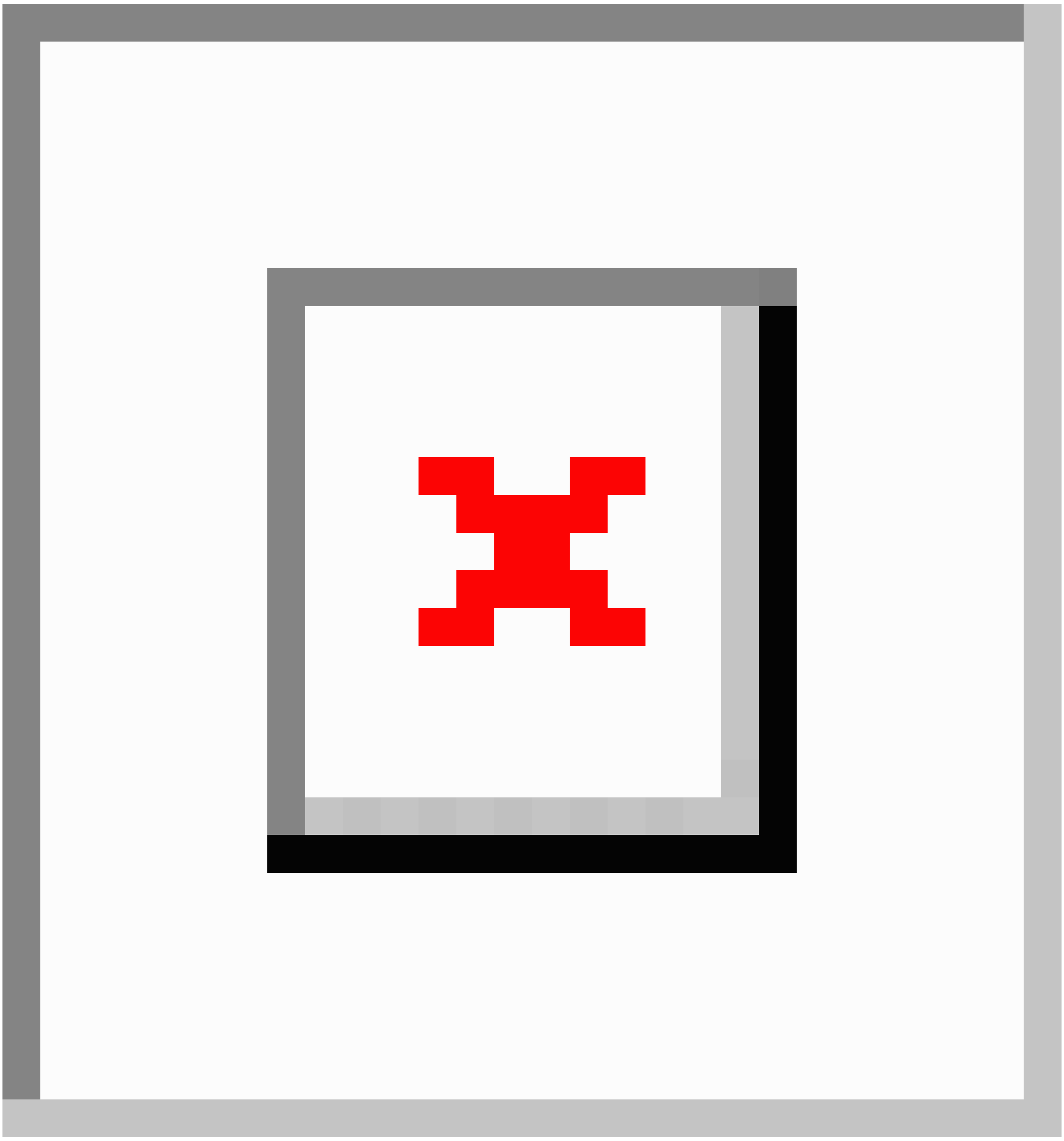
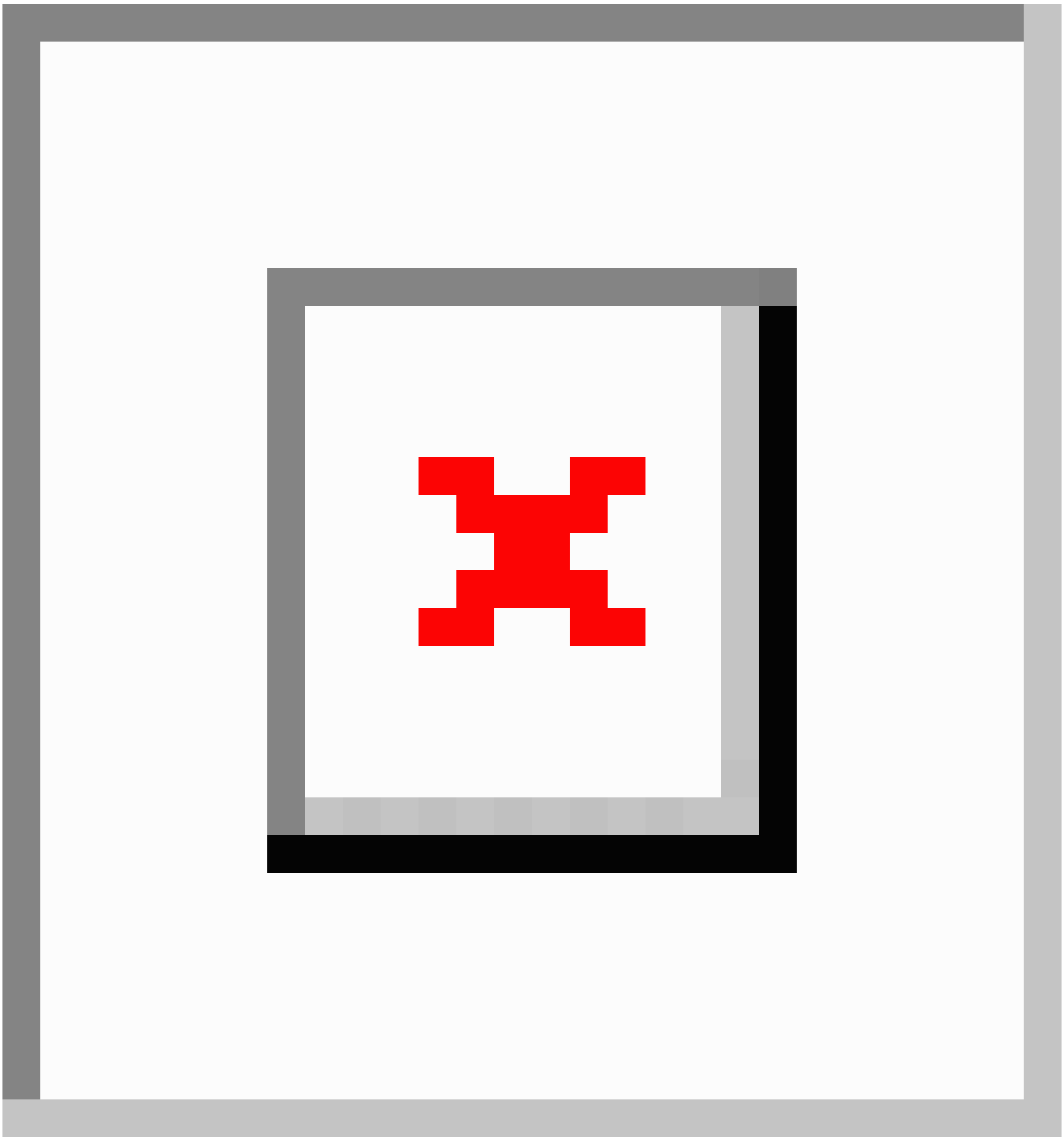


Figure 3. Experiment process.



Game Application and Elements

The experimental group was introduced to NOVA Labs, a digital platform that engages learners in games and interactives that foster authentic scientific exploration. This platform incorporated several gamification elements to enhance students' engagement and motivation:

1. Points system: Each task or activity within the course was assigned a point value. Students accumulated points for every completed activity, which contributed to their overall score and ranking on the leaderboard.
2. Badges: Students could earn badges for completing various tasks and reaching milestones. For instance, badges were awarded for completing weekly assignments on time, achieving high scores on quizzes, and participating in group discussions and collaborative projects.
3. Leaderboards: A dynamic leaderboard was displayed within the application, showcasing the top performers in the class. This element fostered a sense of competition and encouraged students to improve their performance to climb the ranks.
4. Challenges and quests: The course content was structured into thematic challenges and quests. Each week, students embarked on a new quest, which consisted of a series of tasks and activities related to the week's learning objectives. Completing a quest unlocked new content and additional rewards.

The application supporting these gamification elements was user-friendly and accessible through both desktop and mobile devices. It featured an intuitive interface that guided students through their learning journey. Key functionalities included:

1. Dashboard: A personalized dashboard where students could track their progress, view earned badges, and see their current standing on the leaderboard.
2. Interactive lessons: Multimedia-rich lessons incorporating videos, interactive simulations, and practice exercises.
3. Real-time feedback: Immediate feedback on quizzes and assignments to help students identify areas for improvement.
4. Collaboration tools: Features enabling group work and peer-to-peer interactions, such as discussion forums and collaborative project spaces.

Data Collection Tools

The study mainly used quantitative data collection methods. The data were collected through three tools: (1) pretest, (2) posttest, and (3) achievement. The aim of the pretest of academic achievement was to check if participants in the experimental and control groups fulfilled the same minimum criteria of prior knowledge and skills in digital literacy. While the purpose of the posttest was to assess participants' assessment of digital literacy taught throughout the semester and to investigate whether there were differences between the 2 groups (experimental and control) using different teaching approaches, that is, game based and conventional, during weeks 11 - 16 of the semester. These tests were developed by the instructor of the course and were validated by 2 senior professors of digital literacy with more than 10 years' experience of teaching courses in the area of digital literacy.

The survey questionnaire used in this study was administered to the participants in the experimental group only. It was aimed to measure participants' perceptions of gamification regarding their perceived usefulness, ease of use, and their intension to use gamification in future. This was modified from a scale developed by Ghani et al [30]. It consisted of 20 items such as "The educational digital game will improve my learning performance," "I find the educational digital game is easy to use," and "Studying using the educational digital game is a good idea." (see Multimedia Appendix 1). The responses for the items were scored on a 5-point scale, with 1, 2, 3, 4, and 5 representing "strongly disagree," "disagree," "neutral," "agree," and "strongly agree," respectively. The internal consistency of the scale was found to be acceptable with a Cronbach α value of 0.81 [31].

Data Analysis

All data preprocessing and analyses were conducted using SPSS (version 21; IBM Corp). Statistical assumptions for parametric tests were checked and confirmed before running the main analyses [31]. Data analysis included descriptive statistics and inferential analysis including independent sample *t* test, 1-way between-subject ANOVA, and linear regression. Prior to conducting inferential statistical analyses, we tested for the assumptions underlying each statistical test. For the independent sample *t* test and ANOVA, we assessed the normality of data distribution using the Shapiro-Wilk test and ensured homogeneity of variances through Levene test. Additionally, the assumptions of linearity, independence of residuals, and homoscedasticity were confirmed for regression analysis. Moreover, Cronbach α was assessed to check the reliability of the instruments used.

Ethical Considerations

This randomized controlled trial study was approved by the standing committee for Scientific Research Ethics of King Saud University (approval KSU-HE-22 - 871). However, the trial was not registered by a clinical trial registration organization as it did not involve an explicit medical treatment.

Results

This study primarily aimed to investigate the impact of gamification on students' learning outcomes in a digital literacy course and to understand their perceptions and acceptance of gamification for their learning. The main results of this study are presented question-wise in the following sections.

Impact of Gamification on Learning Outcomes

Our first question was focused on investigating whether there was any significant impact of the gamified learning experience on students' learning outcomes. In specific terms, we wanted to check if there were significant differences in the academic achievement between the students who were taught using the gamified learning approach and those who were taught using conventional teaching methods in a digital literacy course at the undergraduate level. This question was analyzed using two main variables: (1) students' scores in the achievement test conducted at the end of the term as dependent variable in the analysis and (2) students' group (control and experimental) as an independent variable in the analysis. The dependent variable was measured on a scale with quantitative values ranging from 0 to 25, while the independent variable was recorded as a nominal variable with only 2 possible values: 1 representing the control group and 2 representing the experimental group.

The results of an independent-sample *t* test indicated that there was significant differences in students' academic performance between those who were taught using conventional teaching methods (median 15.87, SD 2.15) and those whose learning was supported with gamification (mean 21.00, SD 1.88) in a digital literacy class (t_{166}=−16.435; *P*=.001; Table 1). In simpler words, the experimental group performed better than the control group in the academic achievement test that was conducted as a summative assessment in the class. Overall, this finding suggests that the use of the gamification approach in the teaching of digital literacy proves to be an effective instructional approach at the undergraduate level.

Table . Independent *t* test results to test the impact of gamification.

		Participants, n (%)	Mean (SD)	*t* test (*df*)	*P* value
Group				–16.435 (166)	.001
	Control	84 (50)	15.87 (2.15)		
	Experimental	84 (50)	21.00 (1.88)		

Acceptance of Gamification by the Students

The second research question proposed in this study seeks to find the factors that influence the acceptance of students toward gamified learning experiences in a digital literacy course. Students' acceptance of a gamified learning experience was measured in terms of their attitude toward gamification for learning digital literacy. This variable was measured through Likert scale (ordinal type variables) items and was computed by taking the average of all items within the scale. Overall, the results indicated that the participants reported a high level of satisfaction (mean 3.857, SD 0.61) with their gamified learning experience in a digital literacy course. Regarding the factors influencing the acceptance of gamification by the students, we tested factors including age and the major of their degrees. Depending on the nature of variables, that is, scale variables and nominal variables with more than 2 levels, we used different statistical tests for the analysis of this research question.

The results of simple linear regression and 1-way between-subject ANOVA revealed that the major of the degree program was not a significant factor for participants' acceptance of gamification in a digital literacy class (P=.06). Similarly, a simple linear regression analysis did not find age as a significant predictor of acceptance of gamification (P=.06). All in all, these results suggest that neither age nor major of degree are factors for participants' high level of acceptance of gamified learning. The participants in this study showed a positive attitude toward the acceptance of gamification in a digital literacy course regardless of their age or the academic discipline of their degrees.

Effect of Learners' Prior Experience With Gamification

In the third question of the study, we were interested to confirm if participants' prior experience with gamification affects their reported acceptance of gamified learning for a digital literacy course. For the analysis of this question, participants' acceptance score served as the dependent variable, which was measured as a scale type variable, while their prior experience served as the independent variable, which was measured as a nominal variable with three levels: "no experience," "little experience," and "extended experience." About one-third of the participants (n=26, 31%) had no prior experience of gamification. A total of 33 (39.3%) the participants reported that they had little experience of using games for learning. Likewise, 29.8% (n=25) of the participants had extended prior experience of gamification.

Since the levels of the independent variable were more than two, a 1-way between-subject ANOVA was chosen for the analysis of this question. The ANOVA results showed that there was no statistically significant difference in participants' acceptance of gamified learning for the digital literacy course in respect to their prior experience of gamification ($F_{2,81}$=1.319; P=.27; Table 2 for details). Since the main result of the ANOVA analysis was found to be nonsignificant, further post hoc analysis was not needed. These results suggested that regardless of the participants' differences in their prior experience of gamification (some were not exposed to this teaching strategy, while some were experienced with it), their current level of acceptance for the gamified learning experience of digital literacy was almost similar.

Table . ANOVA results showing the influence of prior experience of gamification.

		Sum of squares (*df*)	Mean square	*F* test (*df*)	*P* value
Analyses				1.319 (2,81)	.27
	Between groups	0.974 (2)	0.487		
	Within groups	29.914 (81)	0.369		
Total		30.888 (83)	—[a]	—	—

[a]Not applicable.

Discussion

Principal Results

The ultimate goal of the study was to investigate the impact and acceptance of gamification in a digital literacy course at the undergraduate level. To achieve this goal, we addressed 3 research questions. The findings of these questions are discussed below.

For RQ1, we found that there was a significant difference in academic performance between students who were taught using conventional teaching methods and those who experienced gamified learning. Specifically, the experimental group, which was exposed to gamification, achieved a higher mean score (21.00) compared to the control group (mean 15.87), which received traditional teaching methods. This finding supports the notion that gamification can be an effective educational tool to enhance learning outcomes [32]. This finding has practical implications for educators and instructional designers. By implementing the game-based elements into the learning environment, instructors can potentially construct a more engaging and satisfying learning environment for learners [33]. Gamification techniques, such as scoring, badges, leaderboards, and interactive challenges, may foster a sense of competition, achievement, and enjoyment, helping to achieve better learning outcomes.

The data for this study show a high level of satisfaction (mean 3.857) with the gamified learning experience in a digital literacy course. The mean value being close to the upper end of the Likert scale (5) suggests that students generally had a positive attitude toward gamification as a learning approach. Our results are not different to the results of other studies that provide evidence for learners' higher level of satisfaction with gamified learning [34-38]. This positive attitude is a promising indicator of the potential effectiveness of gamification in enhancing student engagement and motivation.

The study findings demonstrate that age is not a significant factor influencing students' acceptance of gamified learning experiences. In this context, it means that students across different age groups showed similar levels of acceptance toward gamification. Although this finding suggests that gamified learning experiences can be suitable for a diverse range of age groups within the undergraduate level, it is important to acknowledge that the limited age range (18-22 years) of the participants in this study restricts our ability to draw strong conclusions about the impact of age on gamification acceptance. In addition, the finding of this study about age is different from what is suggested by Denden et al [38]. They found that "age" moderates the relationship between their experience of participating in a gamification program and perceived self-efficacy, such that it exerts a greater influence on older people.

Further, we found that the major of degree programs does not significantly impact students' acceptance of gamification in a digital literacy course. In other words, regardless of their academic disciplines, students were similarly receptive to the gamified learning approach. This result is encouraging, as it suggests that gamification can be applied in various subject areas without compromising its acceptance among students.

The nonsignificant effects of age and the major of degree on students' acceptance of gamified learning experiences have practical implications. It indicates that gamification has broad applicability and can be effectively integrated into digital literacy courses, irrespective of students' demographic characteristics or academic backgrounds. Educators can use gamification as a versatile tool to enhance student engagement and motivation across diverse student populations.

Lastly, we found that the participants had varying levels of prior experience with gamification. However, there was no statistically significant difference in participants' acceptance of gamified learning based on their prior experience with gamification. The 1-way between-subject ANOVA did not show any significant effect of prior experience (with 3 levels: no experience, little experience, and extended experience) on participants' acceptance scores. In other words, whether participants had no exposure to gamification or considerable experience with it, their acceptance of gamified learning in a digital literacy course remained similar. It suggests that even students who have never been exposed to gamified learning strategies can still embrace and appreciate the approach. Additionally, students with previous experience with gamification did not necessarily have a more positive attitude toward it compared to their peers who had no experience. This implies that the effectiveness and acceptance of gamified learning can extend to a wide range of students with different levels of exposure to gamification.

Limitations

While this study contributes valuable insights into the impact and acceptance of gamification in a digital literacy course at the undergraduate level, it is essential to acknowledge certain limitations. First, the research was conducted in a specific educational context, which may restrict the generalizability of the findings to other disciplines or educational levels. The use of convenience sampling, while practical for this study, further constrains the diversity of the participant pool. This sampling technique can lead to selection bias, as the sample may not be representative of the larger population. Future studies should aim to replicate this research with more diverse populations to explore the potential differences in gamification acceptance across various demographic and cultural groups.

The participant group, consisting of 168 native Arab speakers aged between 18 and 22 years from a single university, represents a fairly homogenous demographic. This homogeneity, while facilitating a focused analysis within this specific context, limits the broader applicability of our results. Another critical limitation arising from this homogeneity is the reduced variability in cultural background and academic exposure. Given the relatively homogeneous nature of our sample, caution should be exercised in extrapolating our conclusions to populations with greater age diversity or differing backgrounds. Future research endeavors should strive to recruit participants from a wider age spectrum to better understand the influence of age on acceptance of gamification in educational contexts. Additionally, exploring cultural and contextual factors beyond age could provide further insights into the acceptance and effectiveness of gamified learning experiences across diverse learner populations.

In addition, the study focused on a single course, and the participants were from a particular institution, which may limit the diversity of the sample. Additionally, the self-reported nature of some data, such as prior experience and acceptance scores, may have introduced response bias. Future research could address these limitations by conducting similar investigations in diverse educational settings with larger and more representative samples. Future studies could also explore the long-term effects of gamified learning experiences on students' retention of knowledge, skill development, and continued motivation in subsequent courses or academic years. Longitudinal research can offer deeper insights into the sustained benefits of gamification in educational settings.

Conclusions

In conclusion, this study sheds light on the impact and acceptance of gamification in a digital literacy course at the undergraduate level. The findings of our study have both theoretical and practical implications. The research reveals that gamified learning experiences positively influence students' academic performance, leading to higher achievement compared to conventional teaching methods. This study also demonstrates that the students' level of acceptability for gamified learning is not affected by prior experience of gamification. Educators and instructional designers can leverage the insights gained from this study to create more engaging and effective learning environments that foster student motivation and satisfaction. As the educational landscape continues to evolve, the integration of gamification into pedagogical practices stands as a promising approach to enriching the learning experiences of students in digital literacy courses and beyond.

Acknowledgments

The author acknowledges the Researchers Supporting Project (RSP2024R314) King Saud University, Riyadh, Saudi Arabia for funding this study.

Data Availability

The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey questionnaire.
[DOCX File, 15 KB - games_v12i1e52017_app1.docx]

References

1. Purwanto A, Fahmi K, Cahyono Y. The benefits of using social media in the learning process of students in the digital literacy era and the education 4.0 era. J Manag Inf Syst 2023 Apr;2(2):1-7. [doi: 10.4444/jisma.v2i2.296]
2. Reddy P, Sharma B, Chaudhary K. Digital literacy: a review of literature. Int J Technoethics 2020 Jan;11(2):65-94. [doi: 10.4018/IJT.20200701.oa1]
3. Gutiérrez-Ángel N, Sánchez-García JN, Mercader-Rubio I, García-Martín J, Brito-Costa S. Digital literacy in the university setting: a literature review of empirical studies between 2010 and 2021. Front Psychol 2022 Sep;13:896800. [doi: 10.3389/fpsyg.2022.896800] [Medline: 36176798]
4. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev 2020 Mar;32(1):77-112. [doi: 10.1007/s10648-019-09498-w]
5. Saleem AN, Noori NM, Ozdamli F. Gamification applications in e-learning: a literature review. Tech Know Learn 2022 Mar;27(1):139-159. [doi: 10.1007/s10758-020-09487-x]
6. Krath J, Schürmann L, von Korflesch HFO. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comput Human Behav 2021 Dec;125:106963. [doi: 10.1016/j.chb.2021.106963]
7. Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. Presented at: 47th Hawaii International Cconference on System Sciences; Jan 6-9, 2014; Waikoloa, HI. [doi: 10.1109/HICSS.2014.377]
8. Trinidad M, Ruiz M, Calderon A. A bibliometric analysis of gamification research. IEEE Access 2021 Mar;9:46505-46544. [doi: 10.1109/ACCESS.2021.3063986]
9. Huang R, Ritzhaupt AD, Sommer M, et al. The impact of gamification in educational settings on student learning outcomes: a meta-analysis. Education Tech Research Dev 2020 Jul;68(4):1875-1901. [doi: 10.1007/s11423-020-09807-z]
10. Molina-Torres G, Rodriguez-Arrastia M, Alarcón R, et al. Game-based learning outcomes among physiotherapy students: comparative study. JMIR Serious Games 2021 Mar 24;9(1):e26007. [doi: 10.2196/26007] [Medline: 33759800]
11. Aibar-Almazán A, Castellote-Caballero Y, Carcelén-Fraile MDC, Rivas-Campo Y, González-Martín AM. Gamification in the classroom: Kahoot! as a tool for university teaching innovation. Front Psychol 2024 Mar;15:1370084. [doi: 10.3389/fpsyg.2024.1370084] [Medline: 38646120]
12. Garza MC, Olivan S, Monleón E, et al. Performance in Kahoot! activities as predictive of exam performance. BMC Med Educ 2023 Jun 6;23(1):413. [doi: 10.1186/s12909-023-04379-x] [Medline: 37280600]
13. Schöbel SM, Janson A, Söllner M. Capturing the complexity of gamification elements: a holistic approach for analysing existing and deriving novel gamification designs. Eur J Inf Syst 2020 Nov 1;29(6):641-668. [doi: 10.1080/0960085X.2020.1796531]

14. Toda AM, Klock ACT, Oliveira W, et al. Analysing gamification elements in educational environments using an existing gamification taxonomy. Smart Learn Environ 2019 Dec;6(1):1-14. [doi: 10.1186/s40561-019-0106-1]
15. Manzano-León A, Camacho-Lazarraga P, Guerrero MA, et al. Between level up and game over: a systematic literature review of gamification in education. Sustainability 2021 Feb;13(4):2247. [doi: 10.3390/su13042247]
16. Ishaq K, Mat Zin NA, Rosdi F, Jehanghir M, Ishaq S, Abid A. Mobile-assisted and gamification-based language learning: a systematic literature review. PeerJ Comput Sci 2021 May;7:e496. [doi: 10.7717/peerj-cs.496] [Medline: 34084920]
17. Shortt M, Tilak S, Kuznetcova I, Martens B, Akinkuolie B. Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. Comput Assist Lang Learn 2021 Jul;36(3):517-554. [doi: 10.1080/09588221.2021.1933540]
18. Atin S, Syakuran RA, Afrianto I. Implementation of gamification in mathematics m-learning application to creating student engagement. Int J Adv Comput Sci Appl 2022 Aug;13(7):542-556. [doi: 10.14569/IJACSA.2022.0130765]
19. Dondio P, Gusev V, Rocha M. Do games reduce maths anxiety? A meta-analysis. Comput Educ 2023 Mar;194. [doi: 10.1016/j.compedu.2022.104650]
20. Malvasi V, Gil-Quintana J, Bocciolesi E. The projection of gamification and serious games in the learning of mathematics multi-case study of secondary schools in Italy. Mathematics 2022 Jan;10(3):336. [doi: 10.3390/math10030336]
21. Ninaus M, Cortez R, Hazin I, et al. The added value of game elements: better training performance but comparable learning gains. Education Tech Research Dev 2023 Oct;71(5):1917-1939. [doi: 10.1007/s11423-023-10263-8]
22. Deci EL, Ryan RM. Self-determination theory. In: Handbook of Theories of Social Psychology: Sage Publications Ltd; 2012, Vol. 1:416-436. [doi: 10.4135/9781446249215]
23. Silva P. Davis' technology acceptance model (TAM) (1989). In: Al-Suqri MN, Al-Aufi AS, editors. Information Seeking Behavior and Technology Adoption: IGI Global; 2015:205-219. [doi: 10.4018/978-1-4666-8156-9.ch013]
24. Bertrand M, Bouchard S. Applying the technology acceptance model to VR with people who are favorable to its use. J Cyber Ther Rehabil 2008 Jan;1(2):200-210 [FREE Full text]
25. Dehghanzadeh H, Farrokhnia M, Dehghanzadeh H, Taghipour K, Noroozi O. Using gamification to support learning in K ‐ 12 education: a systematic literature review. Brit J Educational Tech 2024 Jan;55(1):34-70. [doi: 10.1111/bjet.13335]
26. Nadi-Ravandi S, Batooli Z. Gamification in education: a scientometric, content and co-occurrence analysis of systematic review and meta-analysis articles. Educ Inf Technol (Dordr) 2022 Apr;27(7):10207-10238. [doi: 10.1007/s10639-022-11048-x] [Medline: 35431599]
27. Tan CW, Yu PD, Lin L. Teaching computational thinking using mathematics gamification in computer science game tournaments. In: Computational Thinking Education: Springer; 2019:167-181. [doi: 10.1007/978-981-13-6528-7]
28. Lundstedt T, Seifert E, Abramo L, et al. Experimental design and optimization. Chemometr Intell Lab Syst 1998 Aug;42(1-2):3-40. [doi: 10.1016/S0169-7439(98)00065-3]
29. Kirk RE. Experimental design. In: Sage Handbook of Quantitative Methods in Psychology: Sage; 2009:23-45. [doi: 10.4135/9780857020994]
30. Ghani MTA, Hamzah M, Ramli S, et al. A questionnaire-based approach on technology acceptance model for mobile digital game-based learning. Journal of Global Business and Social Entrepreneurship 2019 Mar;5(14):11-21.
31. Field A. Discovering Statistics Using IBM SPSS Statistics: Sage; 2013.
32. Alsawaier RS. The effect of gamification on motivation and engagement. Int J Inf Learn Technol 2018 Jan 2;35(1):56-79. [doi: 10.1108/IJILT-02-2017-0009]
33. Landers RN, Armstrong MB, Collmus AB. How to use game elements to enhance learning: applications of the theory of gamified learning. In: Serious Games and Edutainment Applications: Springer; 2017, Vol. 2:457-483. [doi: 10.1007/978-3-319-51645-5]
34. Legaki NZ, Xi N, Hamari J, Karpouzis K, Assimakopoulos V. The effect of challenge-based gamification on learning: an experiment in the context of statistics education. Int J Hum Comput Stud 2020 Dec;144:102496. [doi: 10.1016/j.ijhcs.2020.102496] [Medline: 32565668]
35. Amory A, Naicker K, Vincent J, Adams C. The use of computer games as an educational tool: identification of appropriate game types and game elements. Brit J Educational Tech 1999 Oct;30(4):311-321. [doi: 10.1111/1467-8535.00121]
36. Aldemir T, Celik B, Kaplan G. A qualitative investigation of student perceptions of game elements in a gamified course. Comput Human Behav 2018 Jan;78:235-254. [doi: 10.1016/j.chb.2017.10.001]
37. Ade-Ibijola A, Young K, Sivparsad N, et al. Teaching students about plagiarism using a serious game (Plagi-Warfare): design and evaluation study. JMIR Serious Games 2022 Feb 16;10(1):e33459. [doi: 10.2196/33459] [Medline: 35171103]
38. Denden M, Tlili A, Essalmi F, Jemni M, Chen NS, Burgos D. Effects of gender and personality differences on students' perception of game design elements in educational gamification. J Hum Comput Stud 2021 May;154:102674. [doi: 10.1108/IJSMS-02-2020-0020]

Abbreviations

SDT: self-determination theory
TAM: telephone acceptance model

Edited by R Baranyi, T Leung; submitted 20.08.23; peer-reviewed by K Soomro, N Torstensson, S Otero; revised version received 13.06.24; accepted 14.06.24; published 23.08.24.

Please cite as:
Alnuaim A
The Impact and Acceptance of Gamification by Learners in a Digital Literacy Course at the Undergraduate Level: Randomized Controlled Trial
JMIR Serious Games 2024;12:e52017
URL: https://games.jmir.org/2024/1/e52017
doi:10.2196/52017

Interprofessional Team Training With Virtual Reality: Acceptance, Learning Outcome, and Feasibility Evaluation Study

Andrea N Neher[1,2,*], MScN; Rafael Wespi[1,2,*], MSc; Benjamin D Rapphold[3], MScN, MME; Thomas C Sauter[1], MME, Prof Dr Med; Juliane E Kämmer[1,*], Prof Dr, PhD; Tanja Birrenbach[1,*], MME, Dr med, PD

1

2

3

[*]these authors contributed equally

Corresponding Author:
Tanja Birrenbach, MME, Dr med, PD

Abstract

Background: Effective interprofessional teamwork is vital for ensuring high-quality patient care, especially in emergency medicine. However, interprofessional education often fails to facilitate meaningful interaction among health care disciplines. It is therefore imperative to afford early opportunities for cultivating interprofessional teamwork skills. While in-person simulation-based training has been shown to improve performance, this is resource-intensive, especially if it involves multiple professions. Virtual reality (VR)–based training is an innovative instructional approach that demands fewer resources and offers the flexibility of location-independent learning.

Objective: This study aimed to develop and evaluate the acceptance, learning outcome, and feasibility of an interprofessional team (INTEAM) training course that included a VR simulation of a neurological emergency case.

Methods: This 1-group study used a pre- and posttest design to evaluate the 2-hour INTEAM training course for nursing and medical students. The course included an e-learning part, VR simulation, and debriefing. The main learning objectives were derived from the entrustable professional activity 6, namely to handle a common problem in emergency medicine (headache due to subarachnoid hemorrhage and epileptic seizure) that requires interprofessional collaboration, including a structured handover. We used validated and self-constructed questionnaires, pre- and posttests, and open questions to assess the acceptance, learning outcome, and feasibility of the course.

Results: The data of 42 students (21 nursing and 21 medical students) were analyzed and showed good usability in the System Usability Scale (median 72.5, IQR 65 - 80). The perception of usefulness (median 6, IQR 5.8 - 6.9) and ease of use (median 5.9, IQR 5.1 - 6.3) was good among all students. There was a significant increase in the handover performance from pre- (median 8, IQR 6 - 9) to posttraining (median 8, IQR 7 - 9; z=−2.01; P=.045; r=0.33) and of the confidence in caring for patients with seizures (median 3, IQR 2 - 3 and median 3.5, IQR 3 - 4, respectively; z=−3.8; P<.001; r=0.60). In 67% (14/21) of the simulations, technical issues occurred, but all simulations could be carried out completely.

Conclusions: The new INTEAM training course was well received by nursing and medical students. The handover skills and confidence in caring for patients with seizures were improved after the course. Despite technical challenges with the VR simulations, none required termination, and this demonstrates that our approach is feasible. These promising results encourage the use of VR simulations for team training in the education of nursing and medical students.

(JMIR Serious Games 2024;12:e57117) doi:10.2196/57117

KEYWORDS

medical education; simulation; virtual reality; VR; emergency medicine; interprofessional team training; nursing students; medical students; evaluation study; assessment; effectiveness; patient care; simulation-based training; hemorrhage; epileptic seizure; headache

Introduction

In all branches of medicine, seamless collaboration between health care professionals, such as physicians and nurses, is essential in achieving optimal patient care and a high level of patient safety [1-3]. This need is particularly pronounced in emergency medicine, where rapid and coordinated teamwork is critical to managing life-threatening situations [2]. Although teamwork is so important for patient outcomes, there is hardly any time to practice this during work [4], especially in the emergency department. One way to counteract this is to integrate interprofessional education (IPE) into the curriculum of health professionals [2,5-7]. IPE involves students from at least 2 different health care professions, who must learn to understand their respective roles and responsibilities and practice their communication skills and thereby enhance their ability to work seamlessly together in real emergency situations [4,7-10].

The importance of IPE during undergraduate studies is widely recognized, and efforts are underway worldwide to integrate this into health care education [4,11-13]. In Switzerland, for example, the Federal Law on Health Professions requires "familiarity with the interaction between different health professions" as a necessary expertise for graduates [14]. Consequently, Swiss nursing education has made progress in promoting interprofessional collaboration through joint educational activities involving diverse professions, such as midwives, nutritionists, and physiotherapists [15]. In Swiss medical education, the importance of interprofessional work has also been recognized, as is seen in its inclusion in the medical licensing examination. However, it has not been consistently integrated into the curriculum [16]. Previous interprofessional courses have faced challenges and some have been discontinued, in part due to the COVID-19 pandemic [17,18]. As a result, there remains a lack of interaction between nursing and medical students. Addressing this issue is critical given the lifelong collaboration between these professions in the medical field.

To foster interprofessional collaboration and enhance student teamwork skills, team training is an effective approach, as evidenced by numerous reviews [5-7,19]. Team training is commonly conducted through real-life simulations, "a technique...to replace or amplify real experiences with guided experiences, often immersive in nature, that evoke or replicate substantial aspects of the real world in a fully interactive fashion" [20]. Such simulations are frequently carried out in simulation centers, although these require substantial resources, and this limits their accessibility [6,21]. Virtual reality (VR)–based training is an innovative and resource-efficient alternative, which offers immersive experiences that simulate real-world scenarios [22,23]. Current research suggests that VR simulations are at least as effective as real-life simulations and offer advantages such as reduced costs, accessibility, and the ability to practice challenging scenarios [24-28].

Our objective was therefore to develop and evaluate an interprofessional team (INTEAM) training course for students that uses a VR simulation to provide them with an authentic training experience of handling an emergency case, coupled to interprofessional communication. We now report the development, content, and evaluation of the INTEAM training course "Patient handover and headache." We specifically aimed to assess the acceptance (usability, VR-induced sickness, sense of presence, workload, user satisfaction, and technology acceptance) of the VR simulation among nursing and medical students, with comparisons between study programs and genders; its learning outcome among nursing and medical students, including comparisons between study programs; and its feasibility.

Methods

Study Design and Setting

In this evaluation study, a 1-group pre- and posttest design was used. Data collection took place in May 2023 digitally and in a training facility of the University Hospital of Bern, Switzerland.

Beyond the reported data, additional data on the validation of the TEAM (Team Emergency Assessment Measure) instrument [29] collected during the project was reported elsewhere [30]. The data presented in the 2 papers do not overlap, with the exception of demographic information.

Ethical Considerations

The local ethics committee (Kantonale Ethikkommission Bern) deemed our study to be exempt from full ethical approval, as it is not covered by the Human Research Act (BASEC-Nr: Req-2023 - 00208).

Participants and Eligibility Criteria

The study population was a convenience sample of adult (≥18 years of age) final-year medical students of the University of Bern and adult (≥18 years of age) final-year bachelor nursing students of Bern University of Applied Sciences. The project was presented to them by their teachers. Students could volunteer for the course at their university (nursing) or enroll themselves through the course portal (medicine) as one of several optional courses in the spring semester. No compensation was provided. Written informed consent for study participation and publication of study results was obtained from each student.

Exclusion criteria included unwillingness to participate or give informed consent. Students experiencing epilepsy or other sensitivity to flashing light were also excluded.

Training Course

Overview

The development of the INTEAM training course in emergency medicine involved experts from various disciplines, including emergency physicians, nurses, medical educators, and psychologists, who determined the learning objectives based on the entrustable professional activity (EPA) and designed the course content and material and the VR case scenario. Given the course's focus on EPA 6—"Recognize a patient requiring urgent/emergency care, initiate evaluation and management" [16]—the team specifically selected a common emergency medical scenario that required seamless interprofessional collaboration, including 2 structured handovers.

To reduce the time needed for the on-site training, we decided to include an e-learning part that had to be accomplished at home in preparation of the on-site part. The e-learning part aimed at refreshing the major contents of the simulation (headache with red flags, epileptic seizure, and the handover tool "Introduction, Situation, Background, Assessment, and Recommendation" [ISBAR] [31-33]), which had already been covered by the respective curricula.

The on-site part was scheduled for 3 consecutive days, with each day consisting of 3 separate 3-hour time slots. Per slot, 6 students were invited (3 nursing and 3 medical students), as we had 3 rooms and 3 moderators available. Upon arrival, students were given a prebriefing and an orientation tour to introduce them to the upcoming VR simulation. In the VR simulation, the nursing and medical students had the opportunity to practice their future role as nurses or physicians, including their teamwork, as interprofessional pairs in a specific emergency situation. After the simulation, students participated in a debriefing to reflect on their experience and learning points.

e-Learning

Students received a link to the e-learning part approximately 1 week before the on-site session. The e-learning part consisted of 2 videos of 10 minutes each. In the first video, an emergency physician and an advanced practice nurse—with the support of a facilitator—gave an overview of headaches and epileptic seizures (ie, clinical manifestation, diagnostics, and therapy) and shared their best practices. In the second video, the same team presented the dos and don'ts of handovers and emphasized the ISBAR handover tool together with a typical demonstration. Students were also pointed to a written summary of the presentation, further examples, and background material.

VR Simulation

Development

The interdisciplinary author team developed the simulation case, with input from a company for immersive technologies in the health care sector (StellDirVor GmbH). The implementation in VR was carried out by the VR medical simulation company SimX.

The case displayed a male, 56-year-old patient initially presenting to the emergency department with a thunderclap headache and subtle neurological findings and then progressing to a generalized tonic-clonic epileptic seizure due to subarachnoid hemorrhage. The developers carefully considered the main learning objectives (following EPA 6): emergency management of a patient with thunderclap headache and epileptic seizure (including history taking); physical (including neurological) examination; use of the ABCDE (Airway, Breathing, Circulation, Disability, Exposure) approach [34] in the emergency setting; drug therapy; and conducting a structured handover using ISBAR.

Technical Details and Simulation Setup

The VR simulation setup consisted of the application of a fully immersive semiautomated supervised-learning VR training scenario using a Meta Quest 2 VR headset with Elite Straps with built-in battery, touch controllers (Meta Platforms, Inc), and noise-canceling headphones (JBL Tune 760NC). The VR simulation was controlled by a moderator using an OMEN Gaming Laptop by HP (HP Development Company), giving the appropriate prerecorded verbal responses and initiating the appropriate physiological patient response. There were 3 moderators (2 medical students and 1 PhD student with a background in nursing) and a coordinator (PhD student with a background in psychology) who received training from the study team for approximately 5 hours in the emergency scenario and the technical setup. We deliberately selected students as moderators, in concordance with the principles of peer-tutor teaching. The coordinator also served as a substitute moderator and was familiar with the VR system to help the moderators troubleshoot technical issues.

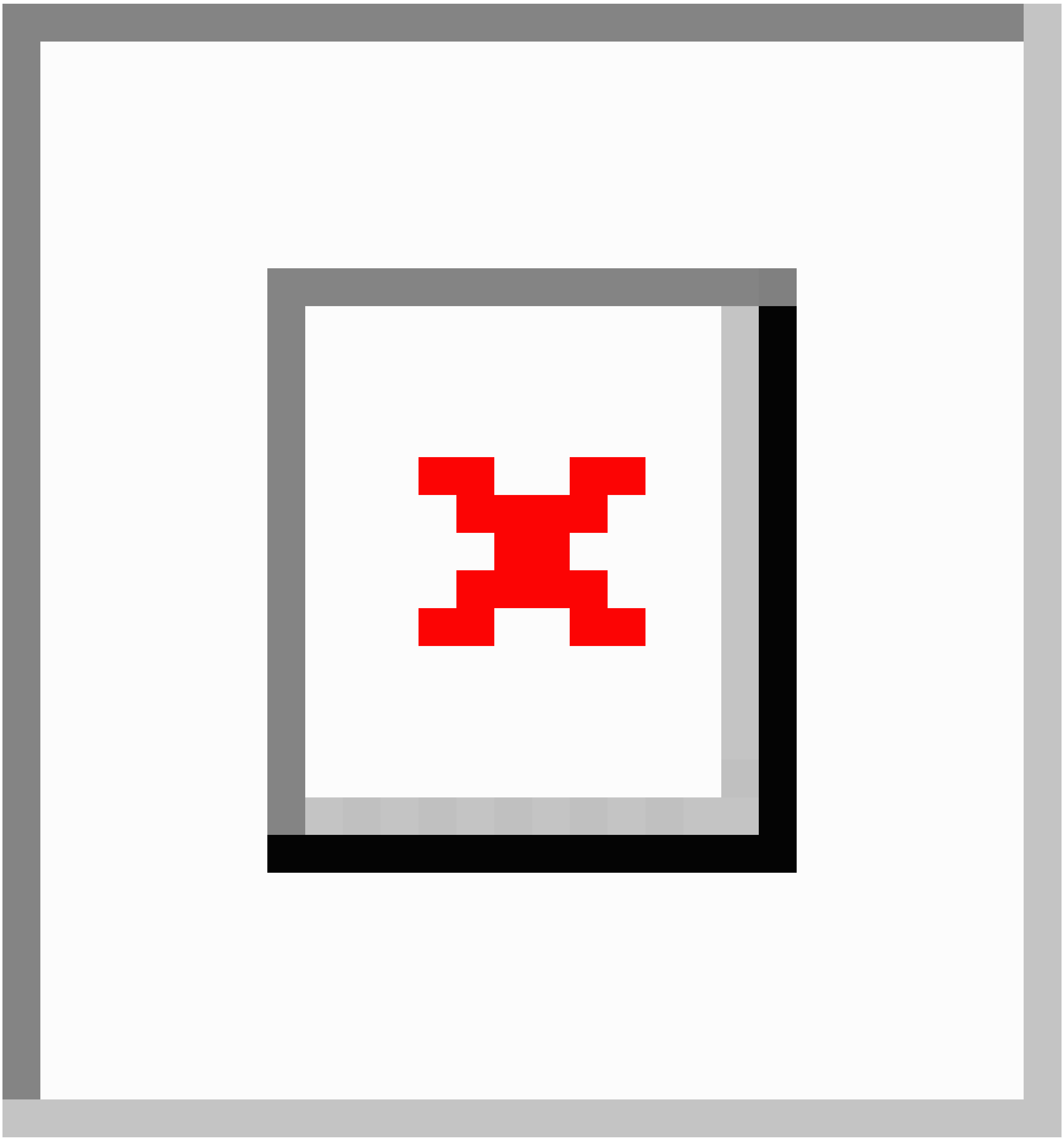
Prebriefing

At the beginning of the on-site part, we conducted a prebriefing with all 6 students to communicate the key aspects, including safety guidelines for the VR experience. We emphasized the importance of creating a safe learning environment and assured students that their performance would only be analyzed for study purposes and in pseudonymized form. We also made clear that the main aim of the course was not to assess their clinical skills but rather to explore the use of VR and an interprofessional course as an effective learning tool.

Orientation Tour

After the prebriefing, students were randomly assigned to pairs of a nursing and a medical student. Each pair was accompanied to a separate room by one of the moderators. Prior to the VR simulation, the student pairs were guided through the VR environment by the moderator in a standardized procedure. During this orientation tour, the team could familiarize themselves with their virtual surroundings and practice how to interact with the virtual environment. They were informed that their teammates' avatars would just be displayed as heads and hands in VR (Figure 1) due to the technical setup of the simulation software. The students were also instructed to speak in standard German, without using dialects, but the patient would respond in English.

Figure 1. Participants in the VR simulation (**A**) in the real training room and (**B**) in the VR patient room. The circles illustrate the corresponding avatars. VR: virtual reality.



Content of VR Simulation

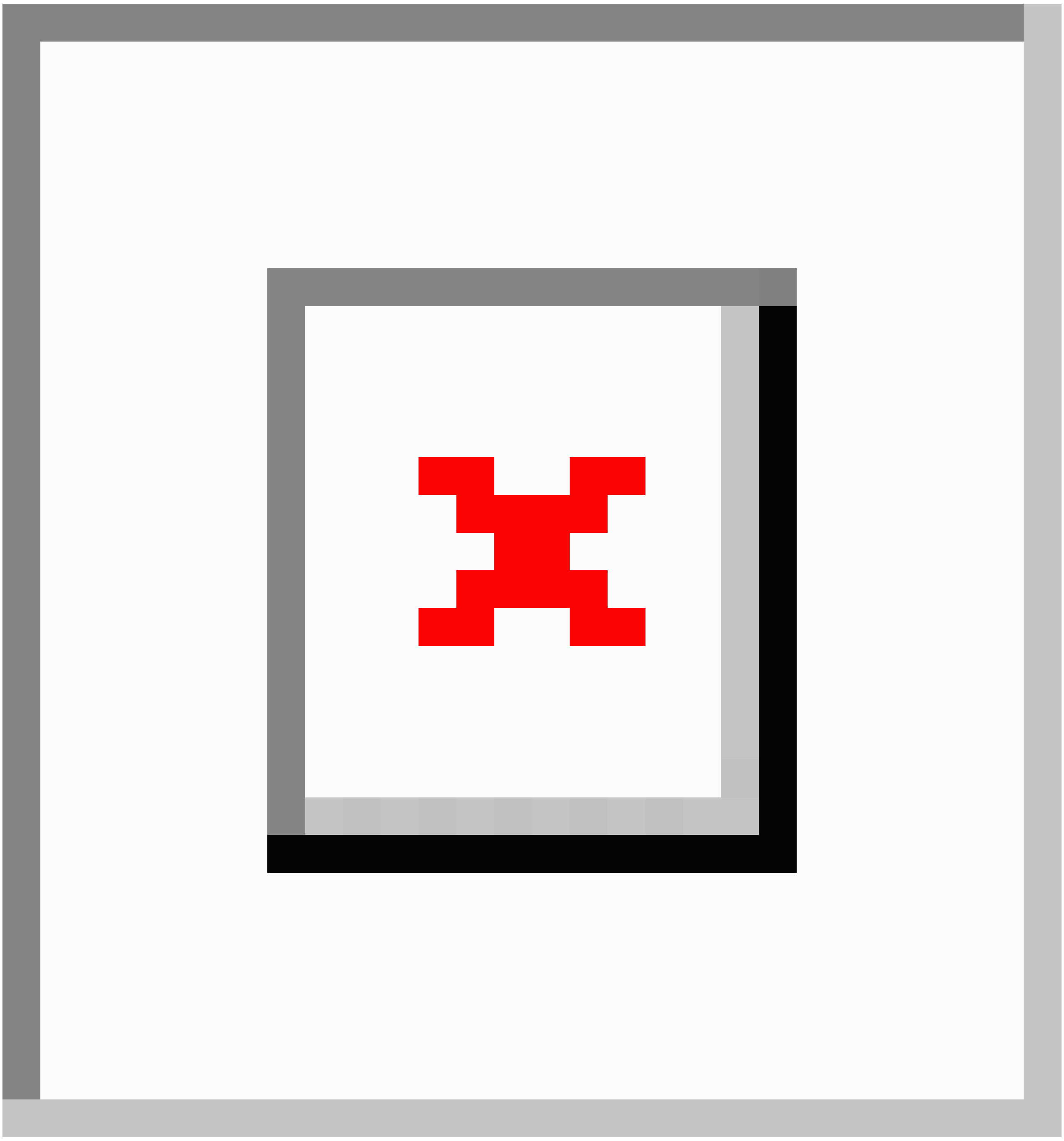
Before entering the virtual patient room, the nursing student received a brief written instruction from the moderator:

> *You are a nurse working in a local hospital. You are called to a room in the emergency department to see a new patient. The patient has walked in on his own and has not yet been seen by a physician. Please perform the initial assessment. The physician will soon come to support you. The physician on duty will knock on the door and then support you.*

The student nurse then started to take the history and examination of the patient who presented with a severe headache and who was accompanied by his wife (nurse assessment, Figure 2). In the meantime, the medical student stayed in the same (physical) room, but was unable to watch or listen to the nurse-patient interaction, and also received written instructions:

> *You are a physician working in a regional hospital. You are called to see a new walk-in patient in the emergency department. The nurse is already there and asks for your assistance. If you hear the knock [of the moderator], you can enter and introduce yourself to the nurse. If you don't get a handover from the nurse, ask for one.*

Figure 2. Virtual reality simulation scenario.



The medical student entered the VR simulation after 5 minutes, and the first handover from the nursing student to the medical student took place. The team was then able to continue taking the history and physical (including neurological) examination (team assessment). After 9 minutes (in total), the patient had an epileptic seizure, which required emergency care such as placing him in a stable side position and administering oxygen and medications. The seizure lasted until it was treated with benzodiazepine or anticonvulsants or ceased spontaneously without treatment at minute 17. After the seizure, the patient was unresponsive for 3 minutes (postictal), after which his cognitive state gradually improved. During this phase (team treatment), the students were able to order further diagnostics and call the attending physician. The VR simulation ended with either a self-initiated handover to the attending physician or a moderator-triggered handover at minute 23 via telephone.

Debriefing

The 30-minute lasting debriefing was conducted after the VR simulation by the moderators according to a debriefing guideline. The guideline was developed from an expert in simulation training (TCS) following the 3D Model of Debriefing [35] and included 4 subjects: defusing (How did it go? How did it feel?); discovering, including medical aspects (What was the medical problem? What did you do?); handover (Did you get all the information? What were the problems?); and deepening (What have you learned for the interprofessional teamwork?

What will you take with you to the next emergency handover?). The debriefing session was neither recorded nor analyzed.

Data Collection

Overview

Before and after the VR simulation, we administered questionnaires as outlined in Table 1. The self-constructed questionnaires can be found in Multimedia Appendix 1.

Table . Overview of all questionnaires.

Category and variable measured		Questionnaire	Point in time
Baseline			
	Age, gender, study program, prior medical education (eg, health care assistant), previous experience with VR[a] simulation and gaming, previous communication training	Self-constructed questionnaire (digital)	Baseline (approximately 2 weeks before the on-site part)
	Visual aid	Self-constructed questionnaire	Presimulation
	Acquaintance	Self-constructed questionnaire	Postsimulation
Acceptance			
	Usability	System Usability Scale [36]	Postsimulation
	VR-induced sickness	Simulator Sickness Questionnaire [37]	Postsimulation
	Sense of presence	Slater, Usoh, and Steed [38]	Postsimulation
	Subjective workload	NASA-Task Load Index [39]	Postsimulation
	User satisfaction	User Satisfaction Evaluation Questionnaire [40]	Postsimulation
	Technology acceptance	Fast Form-Technology Acceptance Model [41]	Postsimulation
Learning outcomes			
	Handover skills	Clinical case vignette 1 and 2 (digital) and rated with Handover Assessment Tool [42]	Baseline and postdebriefing
	Confidence	Self-constructed questionnaire	Baseline and postdebriefing
	Perceived effectiveness	Training Evaluation Inventory [43]	Postsimulation
Feasibility			
	Duration, technical problems, attendance of on-site part	Self-constructed questionnaire (moderator)	On-site part
	Evaluation of the e-learning, VR simulation, and debriefing regarding the achievement of the learning objectives, grade, and suggestions for improvement	Self-constructed questionnaire	Postdebriefing
	Overall comments (regarding feasibility)	Free text	Postdebriefing

[a]VR: virtual reality.

Baseline Data

All students had to fill in a web-based questionnaire via SosciSurvey before starting the INTEAM training course. The link to the questionnaire was sent to students approximately 2 weeks before the on-site session and 1 week before the e-learning. It included questions about sociodemographic factors (age, gender, study program, and prior medical education), previous experience with VR simulation and gaming, and previous communication training (in hours). The use of visual aids was assessed in a survey right before the VR simulation. The students were also asked whether they knew their assigned team partner.

Acceptance

The following questionnaires were completed by the students immediately after the VR simulation.

Usability was assessed using the System Usability Scale (SUS) [36], which consisted of 10 questions to be rated on a 5-point

Likert scale, ranging from 1=strongly disagree to 5=strongly agree. Ratings were then converted so that the resulting SUS score ranges from 0 to 100, with scores lower than 50 being regarded as concerning [36,44].

VR-induced sickness was assessed with the Simulator Sickness Questionnaire (SSQ) [37], where the students had to rate 16 symptoms (eg, nausea and headache) on a 4-point Likert scale from 0=none to 3=severe. Ratings were then converted into 3 subscores and a total score. A total SSQ score above 20 is an indicator of a poor simulator according to Stanney et al [45].

Sense of presence in the virtual world was assessed using the 6-item questionnaire developed by Slater, Usoh, and Steed [38,46], using a semantic differential scale (eg, 1=being elsewhere and 7=sense of being in the virtual environment), with a mean score of 7 representing the strongest sense of presence.

Perceived subjective workload was assessed using the NASA (National Aeronautics and Space Administration)–Task Load Index (NASA-TLX) [39]. The NASA-TLX was calculated by weighting 6 dimensions (mental demand, physical demand, temporal demand, performance, effort, and frustration) assigned by the respondent, with each dimension's rating, and summing these weighted values. The calculated total score can range from 0 to 100, with higher scores indicating more perceived subjective workload.

User satisfaction was assessed using the User Satisfaction Evaluation Questionnaire [40], comprising 6 questions with a 5-point Likert scale from 1=not at all to 5=very much. The total score can range from 6=poor satisfaction to 30=excellent satisfaction.

Technology acceptance was measured using the Fast Form-Technology Acceptance Model (FF-TAM) [41,47], comprising 12 items scored on a 7-point semantic differential scale (eg, 1=ineffective and 7=effective). The FF-TAM mean score ranges from 1 to 7. Higher scores reflect an increased likelihood of technology acceptance based on the subscales usefulness (items 1 - 6) and ease of use (items 7 - 12).

Learning Outcome

As one of the learning objectives was to practice a structured handover using ISBAR, handover skills were assessed with a pre- and posttest to assess the training's learning outcome. The pretest took place as part of the web-based questionnaire mentioned earlier prior to the commencement of the e-learning. For this, students were given 4 minutes to read a clinical case vignette (case 1) and take notes. The screen then automatically showed the next page, where they were instructed to record a verbal, structured handover of the case within 1 minute. The posttest took place after the debriefing, following the same procedure as in the pretest, now with a different clinical case vignette (case 2; Multimedia Appendix 2). Both handovers were transcribed and assessed by 1 trained rater (ANN) using the Handover Assessment Tool (HAT) [42], which was adapted for our study purposes (Multimedia Appendix 3). The tool comprises 12 items and follows the ISBAR framework, resulting in a total score of 0 - 12 points, with higher values indicating better adherence to the ISBAR framework. To capture the perceived learning outcome, confidence was assessed at baseline in the web-based survey and after the INTEAM training course using a self-constructed 4-item questionnaire ("rate your confidence when (1) caring for a patient with a seizure, (2) making a structured handover of an emergency patient, (3) recognizing when to call for help in an emergency situation, and (4) working with a person from another profession") on a 5-point Likert scale (1=very low to 5=very high). The confidence items were based on items from Kolbe et al [12] and have already been tested in a study by Birrenbach et al [48] as well as in the pilot.

In addition, the perceived effectiveness of the VR simulation was measured immediately after the simulation using the Training Evaluation Inventory [43], which consists of 17 statements regarding 5 subscores: subjective enjoyment, perceived usefulness, perceived difficulty, subjective knowledge gain, and attitude toward the training, scored on a 5-point Likert scale ranging from 1=strongly disagree to 5=strongly agree.

Feasibility

To assess the feasibility of the training, notes were taken by the moderator on the duration of the orientation and the VR simulation as well as on any technical problems or other comments regarding the on-site part. The notes of all training were summarized and analyzed descriptively.

The students also evaluated the e-learning, the VR simulation, and the debriefing separately after the debriefing session by answering 3 questions: "How well did this part contribute to achieving the learning objectives?" (1=not at all to 6=very well), "What grade would you give this part?" (1=worst to 6=best), and "Do you have any suggestions for improvement?" (free text). Students were also asked to write down their most important learning experience during the course (free text). There was also space for overall comments regarding feasibility (free text). The free-text responses were coded by one of the authors (ANN), then summarized, and analyzed descriptively.

Statistical Analysis

The data were analyzed using SPSS (version 28.0; IBM Corp) and stored in pseudonymized form. Only the data of complete teams were analyzed. Descriptive statistics, the Mann-Whitney *U* test, and the Fisher exact test were used to compare the baseline characteristics between groups.

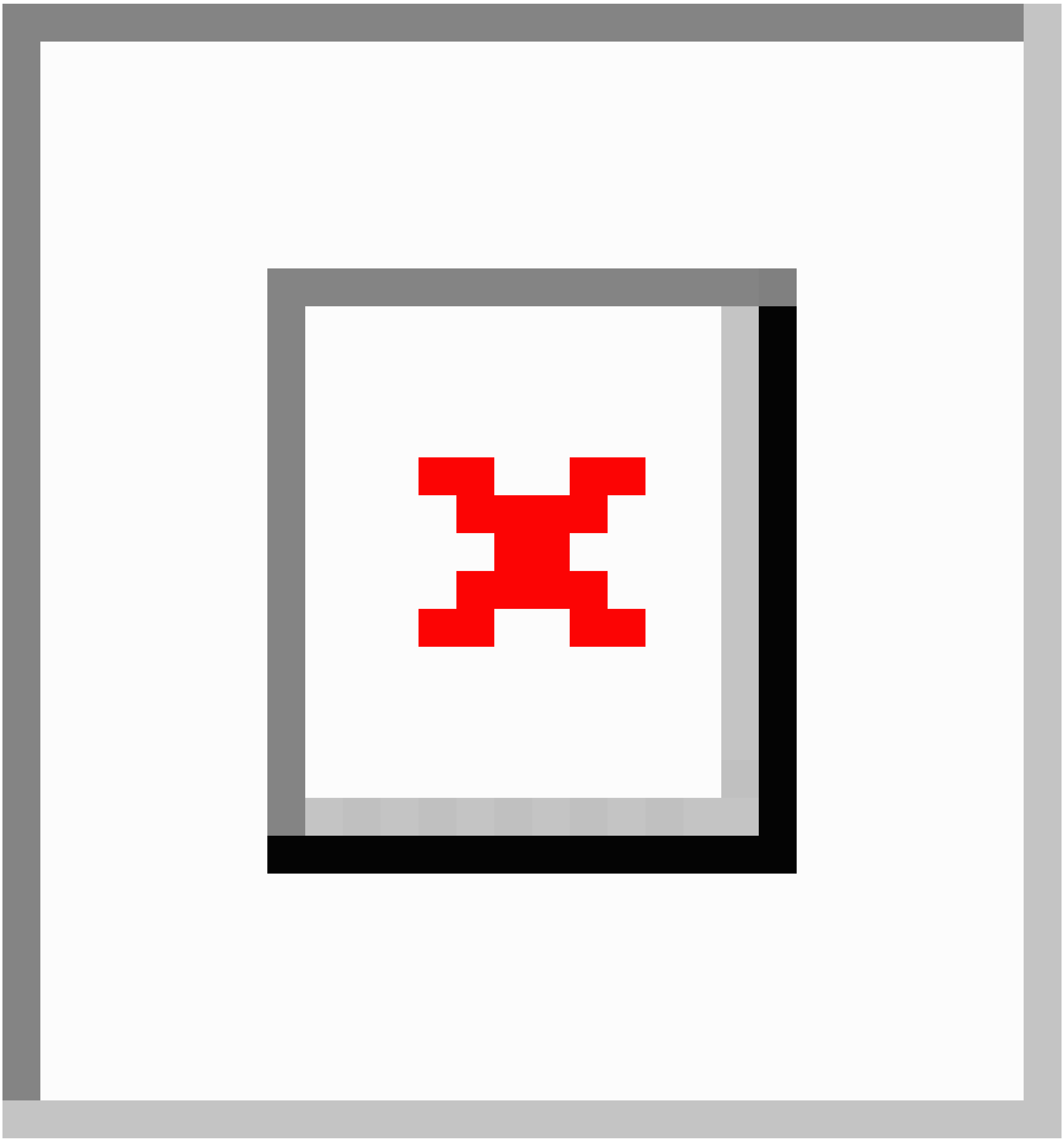
Wilcoxon signed rank tests were performed for pre- and postsimulation comparisons (ie, handover skills and confidence). The Mann-Whitney *U* test was used to compare the questionnaire results of medical with those of nursing students and of women with those of men. A P value of <.05 was considered statistically significant.

Results

Overview

As shown in Figure 3, 54 students enrolled in the course. Among them, 46 completed the e-learning part. Due to no-shows, 4 students had no team partner for the VR simulation. In these cases, the nonpaired students were still invited to conduct the simulation in a modified manner (shorter, moderator as a partner) but excluded from the evaluation. The final sample thus comprised 42 students, including 21 nursing and 21 medical students.

Figure 3. Flowchart. VR: virtual reality.



Sample Characteristics

Table 2 shows the demographic characteristics of our final sample. One nursing student did not complete the baseline questionnaire, but gender and course of study were used, as this was known. Only 1 team was familiar with each other; they had known each other for 6 years. Nursing students differed from medical students in terms of age, prior medical education, and hours of communication training, which may be explained by the different lengths of their curricula (nursing 3 years and medicine 6 years).

Table . Baseline characteristics.

Characteristics	Nursing students (n=20)	Medical students (n=21)	Total (n=41)	*P* value
Age (years), median (IQR)	23 (22 - 24)	26 (25 - 28.5)	25 (23 - 26)	*<.001*[a, b]
Female, n (%)	16 (76)[c]	10 (48)	26 (62)[d]	.06[a]
Visual aids, n (%)				.48[a]
Glasses	5 (24)[c]	6 (30)[e]	11 (27)	
Lenses	3 (14)[c]	4 (20)[e]	7 (17)	
Communication training (hours), median (IQR)	8.5 (7 - 18.8)	15 (12 - 25)	14 (8.5 - 20)	*.006* [a]
Prior medical education, n (%)	6 (30)	0 (0)	6 (15)	*.007* [a]
Playing computer games, n (%)				.13[f]
Several times a week	2 (10)	2 (10)	4 (10)	
About weekly	1 (5)	4 (19)	5 (12)	
1 - 2 times per month	4 (20)	0 (0)	4 (10)	
1 - 2 times per year	1 (5)	5 (24)	6 (15)	
Less than 1 - 2 times per year	3 (15)	2 (10)	5 (12)	
Never	9 (45)	8 (38)	17 (42)	
Virtual reality simulations, n (%)				.44[f]
Several times a week	0 (0)	0 (0)	0 (0)	
About weekly	0 (0)	0 (0)	0 (0)	
1 - 2 times per month	1 (5)	0 (0)	1 (2)	
1 - 2 times per year	0 (0)	1 (5)	1 (2)	
Less than 1 - 2 times per year	4 (20)	7 (33)	11 (27)	
Never	15 (75)	13 (62)	28 (68)	

[a]Mann-Whitney *U* test.
[b]Significant values are present in italics format.
[c]n=21.
[d]n=42.
[e]n=20.
[f]Fisher exact test.

Acceptance

To quantify the acceptance of the VR simulation, we report the median scores of the usability, VR-induced sickness, sense of presence, subjective workload, user satisfaction, and technology acceptance scales (Table 3).

Table . Acceptance variables.

	Nursing students (n=21), median (IQR)	Medical students (n=21), median (IQR)	Total (n=42), median (IQR)	*P* value[a]
Usability: SUS[b] (score ranges 0 to 100, higher=better)				
Total score	70 (65 - 81.3)	77.5 (67.5 - 80)	72.5 (65-80)	.34
Virtual reality–induced sickness: SSQ[c] (total score ranges 0 to 236, higher=more symptoms)				
Total score	30 (11 - 43)	26 (13 - 37)	26 (11 - 41)	.34
Nausea	19.1 (9.5 - 28.6)	19.1 (0 - 28.6)	19.1 (9.5 - 28.6)	.11
Oculomotor	45.5 (15.2 - 60.6)	37.9 (15.2 - 60.6)	37.9 (15.2 - 60.6)	.99
Disorientation	41.8 (20.9 - 83.5)	41.8 (27.8 - 62.6)	41.8 (27.8 - 73.1)	.66
Sense of presence: Slater, Usoh, and Steed (score ranges 0 to 7, 7=strongest sense of presence)				
Total score	5 (4.3 - 5.4)	4.5 (3.9 - 5.5)	4.8 (4.1 - 5.5)	.86
Subjective workload: NASA-TLX[d] (score ranges 0 to 100, higher=more workload)				
Total score	64.3 (60 - 69.2)	64.7 (55.2 - 71.3)	64.5 (59.6 - 70.7)	.89
User satisfaction: USEQ[e] (score ranges from 6=poor satisfaction to 30=excellent satisfaction)				
Total score	24 (21.5 - 28)	25 (23 - 27)	25 (22.8 - 27)	.93
Technology acceptance: FF-TAM[f] (score ranges from 1 to 7, higher=better)				
Subscore				
Usefulness	6.3 (5.8 - 6.8)	6 (5.5 - 6.5)	6 (5.8 - 6.9)	.78
Ease of use	6 (4.7 - 6.3)	5.8 (5.3 - 6.3)	5.9 (5.1 - 6.3)	.49

[a]Mann-Whitney *U* test.

[b]SUS: System Usability Scale.

[c]SSQ: Simulator Sickness Questionnaire.

[d]NASA-TLX: NASA-Task Load Index.

[e]USEQ: User Satisfaction Evaluation Questionnaire.

[f]FF-TAM: Fast Form-Technology Acceptance Model.

Usability was rated as good (SUS: median 72.5, IQR 65-80 of a maximum of 100). SSQ analysis showed moderate VR-induced sickness, primarily due to eye-related problems (eg, difficulty focusing and blurred vision). Perceived participant presence (Slater, Usoh, and Steed) was good, with a median rating of 4.8 (IQR 4.1-5.5) of a possible 7. Students also expressed positive satisfaction (User Satisfaction Evaluation Questionnaire) with the system, giving a median score of 25 (IQR 22.8-27) of a maximum 30. Moreover, technology acceptance (FF-TAM) was deemed satisfactory, with a usefulness median of 6 (IQR 5.8-6.9), and an ease of use median of 5.9 (IQR 5.1-6.3) of a possible 7. However, the workload was rated as notably high, as evidenced by a NASA-TLX median score of 64.5 (IQR 59.6-70.7) of 100.

As can be seen in the last column of Table 3, there were no significant differences between nursing and medical students (all $P≥.11$). However, there were significant differences between female (median 71.3, IQR 62.5 - 80) and male students (median 78.8, IQR 70.6 - 86.3) in the total score of the SUS, with male students rating usability higher ($z=−2.380$; $P=.02$; $r=0.37$). Additionally, female students (median 35.5, IQR 20.3 - 49.8) had significantly higher scores than male students (median 20.5, IQR 8 - 26) in the total score of the SSQ ($z=−2.586$; $P=.01$; $r=0.39$) and its subscore disorientation (median 62.6, IQR 27.8 - 90.5 and median 34.8, IQR 13.9 - 52.2, respectively; $z=−2.481$; $P=.01$; $r=0.38$).

Learning Outcome

For 36 students, we obtained complete recordings of the pre- and posttests on handover skills, which were evaluated with the HAT [42]. The recordings of 5 nursing students and 1 medical student could not be evaluated due to recording errors (blank recording, recording canceled, and recording not executed). Results indicated that nursing students had a lower pretest score than medical students (median 7, IQR 5 - 8 and median 8, IQR 7 - 9, respectively; $z=−1.93$; $P=.045$; $r=0.33$). The HAT score (all students) showed a significant increase from the pretest (median 8, IQR 6 - 9) to posttest (median 8, IQR 7 - 9; $z=−2.01$; $P=.045$; $r=0.33$).

For 41 students, we obtained complete pre- and postdebriefing confidence ratings (1 student did not fill out the pretest). At pretest, students rated their confidence as medium to high (Table 4). For 2 items (caring for patient and call for help), ratings improved at posttest but not for the other 2 ($z=−3.8$; $P<.001$; $r=0.60$ and $z=−3.0$; $P=.003$; $r=0.47$, respectively). Interestingly, medical students indicated lower confidence than nursing students at pretest for the item “working interprofessionally” ($z=−2.3$; $P=.02$; $r=0.36$).

Table . Confidence.[a]

Confidence: items	Nursing students (n=20)			Medical students (n=21)			Total (n=41)		
	Pre, median (IQR)	Post, median (IQR)	*P* value[b]	Pre, median (IQR)	Post, median (IQR)	*P* value[b]	Pre, median (IQR)	Post, median (IQR)	*P* value[b]
Caring for patient with seizure	3 (2 - 3.8)	3 (3-4)	.06	3 (2-3)	4 (3-4)	*<.001*[c]	3 (2-3)	3.5 (3-4)	*<.001*
Making structured handover	3 (3-4)	3 (3-4)	.26	3 (3-4)	3 (3-4)	>.99	3 (3-4)	3 (3-4)	.47
Recognizing when to call for help	4 (3-4)	4 (4-4)	*.03*	4 (3-4)	4 (4-4)	*.03*	4 (3-4)	4 (4-4)	*.003*
Working interprofessionally	4 (4 - 4.8)	4 (4-4)	.25	4 (3-4)	4 (3-4)	.36	4 (4-4)	4 (4-4)	.87

[a]Likert scale from 1=very low to 5=very high.

[b]Wilcoxon signed rank test.

[c]Significant values are present in italics format.

The results of the Training Evaluation Inventory indicate a high perceived effectiveness. The subscores subjective enjoyment, perceived usefulness, and attitudes toward training were rated high (median 5, IQR 4 - 5). The subscores perceived difficulty and subjective knowledge gain received medium ratings (median 4, IQR 4 - 5 and median 4, IQR 3 - 5, respectively). No significant differences were found between nursing and medical students (all *P*≥.25).

Feasibility

The mean length of the orientation tour was 26 (SD 6) minutes, and that of the VR simulation part was 20 (SD 4) minutes. In nearly 67% (n=14) of the 21 VR simulations, technical difficulties arose, but all simulations could be completed. The most serious problems were hardware and software related and involved tracking, resulting in incorrect spawn or disconnection from the VR simulation. There was also 1 case of continuous controller vibration. Software-related problems included the patient overstretching his head and not speaking when the hand (or the controller) came too close to his neck. In addition, some VR simulation materials that had been placed on the patient, such as the blood pressure cuff, were no longer attached to the patient after switching to another phase of the simulation (eg, postictal). There were also situations where either the patient's responses were no longer heard by both students or 1 student could no longer hear the other student and the patient. Other issues were reported, but they were not related to technical problems, like blurred vision or double hearing. Students heard the team member through noise-canceling headphones due to the microphone setup (for recording), but also heard others in real time, causing a confusing mix of delayed and overlapping sounds.

Analyses of the questionnaires revealed that students rated both the e-learning and VR simulation parts as contributing well to achieving the learning objectives (median 5, IQR 4 - 5). They also gave the same high grade to both parts (median 5, IQR 4 - 5). The debriefing received similarly high ratings (median 5, IQR 4.8 - 6) and grades (median 5, IQR 5 - 6).

The coding system and some examples of the free-text responses can be found in Multimedia Appendix 4. Overall, there were only a few suggestions for improvement. Suggestions were made on improving the e-learning content, some of which were inconsistent, such as indicating that the e-learning was overloaded versus that it should be expanded. Feedback on the VR simulation pointed to technical limitations and the need for a longer duration. Suggestions for debriefing related to content and structure. Key learning experiences included interprofessional working, teamwork, handover, and the effectiveness of VR as a learning tool. The overall comments covered a range of issues, with some mentioning technical distractions and a preference for more human-like avatars. Some students found the VR experience new and stressful and suggested repetition for better adaptation. Students emphasized the importance of the orientation tour and the need for clear instructions. Again, some students mentioned the duration of the simulation (too short) and that the VR was a valuable learning tool. Some specifically highlighted its effectiveness for team training and the acquisition of standards.

Discussion

Principal Findings

This evaluation study examined the acceptance, learning outcomes, and feasibility of the newly developed INTEAM training course, in which 21 pairs of nursing and medical students trained together their medical and handover skills in a VR simulation. Results of questionnaire analyses and pre- and posttest comparisons indicate that the course was perceived as highly acceptable, feasible, and may be effective in improving handover skills. In the following, we summarize and discuss the principal findings.

Acceptance

The acceptance of the VR simulation as part of the INTEAM training course was measured by usability, VR-induced sickness, sense of presence, subjective workload, user satisfaction, and technology acceptance. The INTEAM training course demonstrated good usability [44], despite a relatively high rating for VR-induced sickness using the SSQ [37]. Previous research suggests that high levels of simulator or VR-induced sickness can lead to reduced usability [49,50]. However, it is possible that the SSQ is not the most appropriate tool for capturing usability in a VR setting, as cybersickness can differ from simulator sickness [45]. Instead, the use of a specific tool such as the Virtual Reality Sickness Questionnaire may provide more accurate results [51]. In both measures, VR-induced sickness and usability, we observed a gender difference, namely that men rated usability higher and reported less VR-induced sickness, and this is consistent with findings from other studies [52,53].

Compared to other studies [54], the workload was judged as fairly high, potentially due to most students being in a VR simulation for the first time. Although some research indicates a lower workload in VR [28], this pertained to single-player skill training without time constraints. Conversely, other studies have noted increased workload under stress [54,55]. Students mentioned that they had felt additional stress due to the unfamiliarity of the VR tool, particularly during the already tense emergency simulation. Nonetheless, despite these challenges, students expressed a high level of satisfaction.

Learning Outcomes

The learning outcomes of the INTEAM training course were assessed through students' performance in structured handover, self-reported confidence level, and their evaluation of the effectiveness of the training. When comparing handover quality before and after the training, we observed a slight improvement in ISBAR-related learning, which is consistent with studies that showed enhanced outcomes when ISBAR training is coupled with simulations rather than solely relying on theoretical instruction [42,56].

At baseline, medical students showed less confidence in interprofessional work than nursing students possibly because the latter had attended more interdisciplinary lectures on various topics. Students' confidence in handling seizure cases increased significantly after the training, probably due to the integration of theory and practice in an authentic yet controlled environment [26]. The positive comments about VR as a learning tool support the assumption that VR is particularly effective for Generation Z students [57].

Feasibility

Feasibility was captured by recording technical aspects such as the duration of the orientation and VR simulation session, any technical problems, and comments from the moderator. In terms of content, students evaluated e-learning, VR simulation, and debriefing using predefined questions and shared their main learning experiences.

We experienced technical difficulties, but most of them were easily resolved with minimal effort. None of the VR simulations had to be terminated due to technical malfunctions. This may be attributed to the presence of a coordinator who was always available and had high technical expertise. Some disturbances, like blurred vision, were attributed to slight movements of the head-mounted displays, although the Elite Strap helped secure them; nevertheless, text in VR was acknowledged as being potentially less clear than in reality. The additional battery in the Elite Strap proved beneficial and ensured that there were no battery-related interruptions during the training sessions.

One issue with the feasibility was the absence of any replacements who could take over if a student failed to appear. This is also a crucial consideration when incorporating such training into a curriculum [58]. There will typically be a larger number of students, and they will need to be divided into smaller groups. One potential approach is to involve several students per session, with some observing while others participate, as they did in other studies [59]. This facilitates a peer-tutor dynamic while ensuring that at least 1 nursing and 1 medical student are present, even if attendance is not consistent.

As indicated by the grading of the participants, the course content seems feasible. Even though there were a few improvement suggestions, some of these were inconsistent. The positive ratings of the learning objectives by both groups lead us to believe that we have successfully developed valuable learning objectives that are not only interprofessional but also individually relevant for each group.

Students expressed positivity toward the interprofessional approach and would appreciate the inclusion of interprofessional training in their curricula. Other studies showed similar findings [60,61].

Strengths and Limitations

The study's key strengths include the development and evaluation of the INTEAM training by an interdisciplinary team comprising experts in medical education, emergency medicine, psychology, and nursing. This ensured a comprehensive and well-rounded approach to the project. Furthermore, the study used validated and widely used tools for assessing VR-specific outcomes, enhancing the reliability and relevance of the findings. The study also adhered to methodological rigor, contributing to the robustness of the results. However, this study has some limitations worth noting. First, due to the study design, such as the absence of a control group and the relatively small sample size, the statistical robustness and generalizability of the results may be affected. While the findings offer valuable initial insights, it is important to note that no causal relationships can be inferred. Second, voluntary student participation might have introduced self-selection bias. Third, the assessment of handover quality as an indicator of learning outcome relied on a single evaluator and thus may have introduced observer bias despite the evaluator's extensive practical experience. Fourth, not varying the order of vignettes between pre- and posttest of handover quality might have limited the assessment's sensitivity. However, the second vignette was recognized by experts as the more challenging one, which further supports the results that suggest an increase in handover performance.

Future Directions

We believe that there should be careful consideration and further research on the implementation of interprofessional training courses that include a VR simulation. It would be beneficial to investigate whether the technical aspects become less prominent when the VR simulation is repeatedly practiced. Once improved software and hardware become available, it will, of course, be worth exploring whether this leads to overall improvements.

Furthermore, different INTEAM training courses should be developed to explore all the learning objectives that can be achieved through such training. In this case, it is essential to select learning objectives that are significant for all involved students such as, in our case, medical and nursing students. Moreover, we need further research that investigates long-term effects and whether students can strengthen their teamwork skills through these courses and derive practical benefits for their future professional lives, as is the case with traditional team training [10].

Conclusions

The INTEAM training course, including a VR simulation in emergency medicine, was well received by the nursing and medical students, and their handover skills and confidence in managing patients with seizure were improved after the course. Although some technical problems occurred during the VR simulations, none resulted in dropout, thus confirming the feasibility of the approach. Technical enhancements and organizational considerations are advisable for further improvement. These promising results encourage the use of VR simulations for team training in the education of nursing and medical students.

Acknowledgments

The authors would like to thank Fabienne Gschwind, Lukas Schwendimann, and Britta Haupt for their invaluable support with preparing and conducting the study. The authors also would like to thank BeLEARN for the funding and for providing the facilities to conduct the pilot study. ChatGPT3.5 was used to edit selected parts of the manuscript to improve the language, and the changes were subsequently reviewed by a native speaker. This project received financial support from BeLEARN, an association established by 5 Bern universities to advance digitalization in education. The content of this scientific paper reflects the authors' views alone and does not necessarily align with those of BeLEARN. The authors are solely responsible for the paper's content, with no implied influence or endorsement by the funder.

Data Availability

The datasets generated during and analyzed during this study are available in the Open Science Framework repository [62].

Authors' Contributions

ANN was involved in acquisition, analysis, interpretation, drafting, and revision. RW was involved in conception, design, acquisition, drafting, and revision. BDR was involved in conception and revision. TCS and JEK were involved in conception, design, interpretation, and revision. TB was involved in conception, design, interpretation, drafting, and revision. All authors have approved the submitted version and have agreed both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated and resolved, and the resolution documented in the literature.

Conflicts of Interest

TCS holds the endowed professorship of emergency telemedicine at the University of Bern sponsored by the Touring Club Switzerland. The sponsor has no influence on the research or decision to publish. All other authors have nothing to disclose.

Multimedia Appendix 1
Self-constructed questionnaires.
[DOCX File, 35 KB - games_v12i1e57117_app1.docx]

Multimedia Appendix 2
Clinical case vignettes.
[DOCX File, 18 KB - games_v12i1e57117_app2.docx]

Multimedia Appendix 3
Adapted Handover Assessment Tool.
[DOCX File, 19 KB - games_v12i1e57117_app3.docx]

Multimedia Appendix 4
Code system—gives further information about the code system from the free text questions and comments.

[DOCX File, 25 KB - games_v12i1e57117_app4.docx]

References

1. Institute of Medicine (US) Committee on Quality of Health Care in America. In: Kohn LT, Corrigan JM, Donaldson MS, editors. To Err Is Human: Building a Safer Health System: National Academies Press (US); 2000.
2. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. Acta Anaesthesiol Scand 2009 Feb;53(2):143-151. [doi: 10.1111/j.1399-6576.2008.01717.x] [Medline: 19032571]
3. Reime MH, Johnsgaard T, Kvam FI, et al. Simulated settings; powerful arenas for learning patient safety practices and facilitating transference to clinical practice. A mixed method study. Nurse Educ Pract 2016 Nov;21:75-82. [doi: 10.1016/j.nepr.2016.10.003] [Medline: 27769018]
4. Chakraborti C, Boonyasai RT, Wright SM, Kern DE. A systematic review of teamwork training interventions in medical student and resident education. J Gen Intern Med 2008 Jun;23(6):846-853. [doi: 10.1007/s11606-008-0600-6] [Medline: 18386100]
5. Buljac-Samardzic M, Doekhie KD, van Wijngaarden JDH. Interventions to improve team effectiveness within health care: a systematic review of the past decade. Hum Resour Health 2020 Jan 8;18(1):2. [doi: 10.1186/s12960-019-0411-3] [Medline: 31915007]
6. Hughes AM, Gregory ME, Joseph DL, et al. Saving lives: a meta-analysis of team training in healthcare. J Appl Psychol 2016 Sep;101(9):1266-1304. [doi: 10.1037/apl0000120] [Medline: 27599089]
7. Sezgin MG, Bektas H. Effectiveness of interprofessional simulation-based education programs to improve teamwork and communication for students in the healthcare profession: a systematic review and meta-analysis of randomized controlled trials. Nurse Educ Today 2023 Jan;120:105619. [doi: 10.1016/j.nedt.2022.105619] [Medline: 36343420]
8. Kämmer JE, Ewers M. Stereotypes of experienced health professionals in an interprofessional context: results from a cross-sectional survey in Germany. J Interprof Care 2022;36(3):350-361. [doi: 10.1080/13561820.2021.1903405] [Medline: 34014142]
9. Kleib M, Jackman D, Duarte-Wisnesky U. Interprofessional simulation to promote teamwork and communication between nursing and respiratory therapy students: a mixed-method research study. Nurse Educ Today 2021 Apr;99:104816. [doi: 10.1016/j.nedt.2021.104816] [Medline: 33662866]
10. Salas E, DiazGranados D, Klein C, et al. Does team training improve team performance? A meta-analysis. Hum Factors 2008 Dec;50(6):903-933. [doi: 10.1518/001872008X375009] [Medline: 19292013]
11. Chen AS, Yau B, Revere L, Swails J. Implementation, evaluation, and outcome of TeamSTEPPS in interprofessional education: a scoping review. J Interprof Care 2019;33(6):795-804. [doi: 10.1080/13561820.2019.1594729] [Medline: 31009273]
12. Kolbe M, Goldhahn J, Useini M, Grande B. “Asking for help is a strength”—how to promote undergraduate medical students’ teamwork through simulation training and interprofessional faculty. Front Psychol 2023;14:1214091. [doi: 10.3389/fpsyg.2023.1214091] [Medline: 37701867]
13. Guraya SY, Barr H. The effectiveness of interprofessional education in healthcare: a systematic review and meta-analysis. Kaohsiung J Med Sci 2018 Mar;34(3):160-165. [doi: 10.1016/j.kjms.2017.12.009] [Medline: 29475463]
14. Bundesgesetz über die Gesundheitsberufe. Fedlex. URL: https://www.fedlex.admin.ch/eli/cc/2020/16/de [accessed 2024-10-08]
15. Interprofessionelle Lehre. Berner Fachhochschule. URL: https://www.bfh.ch/gesundheit/de/forschung/forschungsbereiche/kompetenzzentrum-interprofessionalitaet/lehre/ [accessed 2024-10-08]
16. Michaud P, Jucker-Kupper P, Members of the PROFILES Working Group. PROFILES; principal relevant objectives and framework for integrated learning and education in Switzerland. : Joint Commission of the Swiss Medical Schools; 2017 URL: https://www.profilesmed.ch/doc/Profiles_2017.pdf [accessed 2024-08-10]
17. Herrmann G, Woermann U, Schlegel C. Interprofessional education in anatomy: learning together in medical and nursing training. Anat Sci Ed 2015 Jul;8(4):324-330. [doi: 10.1002/ase.1506]
18. Woermann U. Interprofessionelles Lernen—50 Jahre. : IML Institut für Medizinische Lehre; 1971 URL: https://www.iml.unibe.ch/attachment/356/download/IML50_final_web_28082023.pdf [accessed 2024-10-14]
19. O’Brien B, Bevan K, Brockington C, Murphy J, Gilbert R. Effects of simulation-based cardiopulmonary and respiratory case training experiences on interprofessional teamwork: a systematic review. Can J Respir Ther 2023;59:85-94. [doi: 10.29390/cjrt-2022-060] [Medline: 36999003]
20. Gaba DM. The future vision of simulation in health care. Qual Saf Health Care 2004 Oct;13(Suppl 1):i2-10. [doi: 10.1136/qhc.13.suppl_1.i2] [Medline: 15465951]
21. Colonna AL, Robbins R, Stefanucci J, et al. Trauma bay virtual reality—a game changer for ATLS instruction and assessment. J Trauma Acute Care Surg 2022 Sep 1;93(3):353-359. [doi: 10.1097/TA.0000000000003569] [Medline: 35170584]
22. Plotzky C, Lindwedel U, Sorber M, et al. Virtual reality simulations in nurse education: a systematic mapping review. Nurse Educ Today 2021 Jun;101:104868. [doi: 10.1016/j.nedt.2021.104868] [Medline: 33798987]
23. Kilmon CA, Brown L, Ghosh S, Mikitiuk A. Immersive virtual reality simulations in nursing education. Nurs Educ Perspect 2010;31(5):314-317. [Medline: 21086871]

24. Berg H, Steinsbekk A. Is individual practice in an immersive and interactive virtual reality application non-inferior to practicing with traditional equipment in learning systematic clinical observation? A randomized controlled trial. BMC Med Educ 2020 Apr 22;20(1):123. [doi: 10.1186/s12909-020-02030-7] [Medline: 32326948]
25. Birrenbach T, Zbinden J, Papagiannakis G, et al. Effectiveness and utility of virtual reality simulation as an educational tool for safe performance of COVID-19 diagnostics: prospective, randomized pilot trial. JMIR Serious Games 2021;9(4):e29586. [doi: 10.2196/29586]
26. Dhar E, Upadhyay U, Huang Y, et al. A scoping review to assess the effects of virtual reality in medical education and clinical care. Dig Health 2023;9:20552076231158022. [doi: 10.1177/20552076231158022] [Medline: 36865772]
27. Kyaw BM, Saxena N, Posadzki P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959. [doi: 10.2196/12959] [Medline: 30668519]
28. Rickenbacher-Frey S, Adam S, Exadaktylos AK, Müller M, Sauter TC, Birrenbach T. Development and evaluation of a virtual reality training for emergency treatment of shortness of breath based on frameworks for serious games. GMS J Med Educ 2023;40(2):Doc16. [doi: 10.3205/zma001598] [Medline: 37361243]
29. Cooper S, Cant R, Porter J, et al. Rating medical emergency teamwork performance: development of the Team Emergency Assessment Measure (TEAM). Resuscitation 2010 Apr;81(4):446-452. [doi: 10.1016/j.resuscitation.2009.11.027] [Medline: 20117874]
30. Wespi R, Schwendimann L, Neher A, et al. TEAMs go VR—validating the TEAM in a virtual reality (VR) medical team training. Adv Simul (Lond) 2024 Sep 11;9(1):38. [doi: 10.1186/s41077-024-00309-z] [Medline: 39261889]
31. Burgess A, van Diggele C, Roberts C, Mellis C. Teaching clinical handover with ISBAR. BMC Med Educ 2020 Dec 3;20(Suppl 2):459. [doi: 10.1186/s12909-020-02285-0] [Medline: 33272274]
32. Müller M, Jürgens J, Redaèlli M, Klingberg K, Hautz WE, Stock S. Impact of the communication and patient hand-off tool SBAR on patient safety: a systematic review. BMJ Open 2018 Aug 23;8(8):e022202. [doi: 10.1136/bmjopen-2018-022202] [Medline: 30139905]
33. Andreasen EM, Høigaard R, Berg H, Steinsbekk A, Haraldstad K. Usability evaluation of the preoperative ISBAR (Identification, Situation, Background, Assessment, and Recommendation) desktop virtual reality application: qualitative observational study. JMIR Hum Factors 2022 Dec 29;9(4):e40400. [doi: 10.2196/40400] [Medline: 36580357]
34. Soltan M, Kim M. The ABCDE approach explained. BMJ 2016 Oct 24;355:i4512. [doi: 10.1136/sbmj.i4512] [Medline: 31055337]
35. Zigmont JJ, Kappus LJ, Sudikoff SN. The 3D model of debriefing: defusing, discovering, and deepening. Semin Perinatol 2011 Apr;35(2):52-58. [doi: 10.1053/j.semperi.2011.01.003] [Medline: 21440811]
36. Brooke J. SUS—a quick and dirty usability scale. In: Usability Evaluation in Industry: Taylor and Francis; 1996:189-194.
37. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
38. Slater M, Steed A, McCarthy J, Maringelli F. The influence of body movement on subjective presence in virtual environments. Hum Factors 1998 Sep;40(3):469-477. [doi: 10.1518/001872098779591368] [Medline: 9849105]
39. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv Psychol 1988;52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
40. Gil-Gómez JA, Manzano-Hernández P, Albiol-Pérez S, Aula-Valero C, Gil-Gómez H, Lozano-Quilis JA. USEQ: a short questionnaire for satisfaction evaluation of virtual rehabilitation systems. Sensors (Basel) 2017 Jul 7;17(7):1589. [doi: 10.3390/s17071589] [Medline: 28686174]
41. Chin WW, Johnson N, Schwarz A. A fast form approach to measuring technology acceptance and other constructs. MIS Q 2008;32(4):687. [doi: 10.2307/25148867]
42. Holt N, Crowe K, Lynagh D, Hutcheson Z. Is there a need for formal undergraduate patient handover training and could an educational workshop effectively provide this? A proof-of-concept study in a Scottish medical school. BMJ Open 2020 Feb 12;10(2):e034468. [doi: 10.1136/bmjopen-2019-034468] [Medline: 32051318]
43. Ritzmann S, Hagemann V, Kluge A. TEI. Training Evaluations Inventar. ZPID (Leibniz Institute for Psychology). 2020. URL: https://www.psycharchives.org/handle/20.500.12034/3069 [accessed 2024-10-08]
44. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
45. Stanney KM, Kennedy RS, Drexler JM. Cybersickness is not simulator sickness. Proc Hum Factors Ergon Soc Annu Meet 1997 Oct;41(2):1138-1142. [doi: 10.1177/107118139704100292]
46. Usoh M, Catena E, Arman S, Slater M. Using presence questionnaires in reality. Pres Teleoperators Virtual Environ 2000 Oct;9(5):497-503. [doi: 10.1162/105474600566989]
47. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
48. Birrenbach T, Stuber R, Müller CE, et al. Virtual reality simulation to enhance advanced trauma life support trainings—a randomized controlled trial. BMC Med Educ 2024 Jun 17;24(1):666. [doi: 10.1186/s12909-024-05645-2] [Medline: 38886688]

49. Schultheis MT, Rebimbas J, Mourant R, Millis SR. Examining the usability of a virtual reality driving simulator. Assist Technol 2007;19(1):1-8. [doi: 10.1080/10400435.2007.10131860] [Medline: 17461285]
50. Voinescu A, Petrini K, Stanton Fraser D. Presence and simulator sickness predict the usability of a virtual reality attention task. Virtual Real 2023 Mar 24;27(3):1-17. [doi: 10.1007/s10055-023-00782-3] [Medline: 37360806]
51. Kim HK, Park J, Choi Y, Choe M. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl Ergon 2018 May;69:66-73. [doi: 10.1016/j.apergo.2017.12.016] [Medline: 29477332]
52. Munafo J, Diedrick M, Stoffregen TA. The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. Exp Brain Res 2017 Mar;235(3):889-901. [doi: 10.1007/s00221-016-4846-7] [Medline: 27915367]
53. Mahling M, Wunderlich R, Steiner D, Gorgati E, Festl-Wietek T, Herrmann-Werner A. Virtual reality for emergency medicine training in medical school: prospective, large-cohort implementation study. J Med Internet Res 2023 Mar 3;25:e43649. [doi: 10.2196/43649] [Medline: 36867440]
54. Grier RA. How high is high? A meta-analysis of NASA-TLX global workload scores. Proc Hum Factors Ergon Soc Annu Meet 2015 Sep;59(1):1727-1731. [doi: 10.1177/1541931215591373]
55. Favre-Félix J, Dziadzko M, Bauer C, et al. High-fidelity simulation to assess Task Load Index and performance: a prospective observational study. Turk J Anaesthesiol Reanim 2022 Aug;50(4):282-287. [doi: 10.5152/TJAR.2022.21234] [Medline: 35979975]
56. Darbyshire D, Gordon M, Baker P. Teaching handover of care to medical students. Clin Teach 2013 Feb;10(1):32-37. [doi: 10.1111/j.1743-498X.2012.00610.x] [Medline: 23294741]
57. Chicca J, Shellenbarger T. Connecting with Generation Z: approaches in nursing education. Teach Learn Nurs 2018 Jul;13(3):180-184. [doi: 10.1016/j.teln.2018.03.008]
58. Liaw SY, Ooi SW, Rusli KDB, Lau TC, Tam WWS, Chua WL. Nurse-physician communication team training in virtual reality versus live simulations: randomized controlled trial on team communication and teamwork attitudes. J Med Internet Res 2020 Apr 8;22(4):e17279. [doi: 10.2196/17279] [Medline: 32267235]
59. Kiessling A, Amiri C, Arhammar J, et al. Interprofessional simulation-based team-training and self-efficacy in emergency medicine situations. J Interprof Care 2022;36(6):873-881. [doi: 10.1080/13561820.2022.2038103] [Medline: 35341425]
60. Krielen P, Meeuwsen M, Tan E, Schieving JH, Ruijs A, Scherpbier ND. Interprofessional simulation of acute care for nursing and medical students: interprofessional competencies and transfer to the workplace. BMC Med Educ 2023 Feb 11;23(1):105. [doi: 10.1186/s12909-023-04053-2] [Medline: 36774481]
61. Mahler C, Schwarzbeck V, Mink J, Goetz K. Students' perception of interprofessional education in the bachelor programme “Interprofessional Health Care” in Heidelberg, Germany: an exploratory case study. BMC Med Educ 2018 Jan 25;18(1):19. [doi: 10.1186/s12909-018-1124-3] [Medline: 29370784]
62. Neher AN. INter-professional TEAM (INTEAM) training with virtual reality: evaluation study assessing its acceptance, learning outcome, and feasibility. Open Science Framework. 2023. URL: https://osf.io/yha76/ [accessed 2024-10-09]

Abbreviations

ABCDE: Airway, Breathing, Circulation, Disability, Exposure
EPA: entrustable professional activity
FF-TAM: Fast Form-Technology Acceptance Model
HAT: Handover Assessment Tool
INTEAM: interprofessional team
IPE: interprofessional education
ISBAR: Introduction, Situation, Background, Assessment, and Recommendation
NASA: National Aeronautics and Space Administration
NASA-TLX: NASA-Task Load Index
SSQ: Simulator Sickness Questionnaire
SUS: System Usability Scale
TEAM: Team Emergency Assessment Measure
VR: virtual reality

Edited by A Coristine, T Leung; submitted 05.02.24; peer-reviewed by E Andreasen, M Aksoy; revised version received 30.07.24; accepted 04.09.24; published 04.11.24.

Please cite as:
Neher AN, Wespi R, Rapphold BD, Sauter TC, Kämmer JE, Birrenbach T
Interprofessional Team Training With Virtual Reality: Acceptance, Learning Outcome, and Feasibility Evaluation Study
JMIR Serious Games 2024;12:e57117
URL: https://games.jmir.org/2024/1/e57117
doi:10.2196/57117

Impact of Video-Based Error Correction Learning for Cardiopulmonary Resuscitation Training: Quasi-Experimental Study

Yun Wang[1,*], MD; Junzuo Fu[1,*], MD; Shaoping Wang[1], RN; Huijuan Wang[1], MD; Wei Gao[2], RN; Lina Huang[1], MD, PhD

1

2

*these authors contributed equally

Corresponding Author:
Lina Huang, MD, PhD

Abstract

Background: Video-based error correction (VBEC) in medical education could offer immediate feedback, promote enhanced learning retention, and foster reflective practice. However, its application in cardiopulmonary resuscitation (CPR) training has not been investigated.

Objective: The objective of this study is to assess whether the VBEC procedure could improve the training performance of CPR among anesthesiology residents.

Methods: A quasi-experimental study was conducted among anesthesiology residents between December 2022 and April 2023. Primary outcomes included a posttraining knowledge test and practical assessment scores. Secondary outcomes included the number of residents who correctly conducted CPR at each step, the rate of common mistakes during the CPR process, and the self-assessment results. A total of 80 anesthesiology residents were divided into a VBEC group (n=40) and a control group (n=40). The VBEC group underwent a 15-minute VBEC CPR training, whereas the control group underwent a 15-minute video-prompting CPR training.

Results: The posttraining knowledge test score of the VBEC group was significantly higher than that of the control group (73, SD 10.5 vs 65.1, SD 11.4; *P*=.002). The residents in the VBEC group had lower error rates in “failure to anticipate the next move” (n=3, 7.5% vs n=13, 32.5%; *P*=.01) and “failure to debrief or problem solve after the code” (n=2, 5% vs n=11, 27.5%; *P*=.01), as well as better performance in the “secure own safety” step (n=34, 85% vs n=18, 45%; *P*<.001) than those in the control group. The VBEC group showed significantly higher confidence in CPR than the control group (n=?, 62.5% vs n=?, 35%; *P*=.03).

Conclusions: VBEC may be a promising strategy compared to video prompting for CPR training among anesthesiology residents.

(JMIR Serious Games 2024;12:e53577) doi:10.2196/53577

KEYWORDS

video-based error correction; video-prompting; cardiopulmonary resuscitation training; anesthesiology resident; quasi-experimental study; anesthesiology; cardiopulmonary; cardiopulmonary resuscitation; training; video; learning; residents; CPR training; CPR; video prompting

Introduction

Background

Cardiac arrest remains a significant health care burden despite substantial improvements in survival rates in the past several decades [1]. High-quality cardiopulmonary resuscitation (CPR) is the key link in the chain of survival, and successful CPR may increase the survival rate of patients experiencing cardiac arrest [2,3]. For anesthesiology residents, the likelihood of administering CPR is high, and thus CPR training is of utmost importance [4]. Unfortunately, due to limited training quality and efficiency, the overall quality of CPR remains low, which may further lead to undesired outcomes in clinical practice [5,6]. Thus, CPR instruction modalities are needed to improve the quality of CPR training of anesthesiology residents to maximize their performance.

Video prompting has been shown to be an effective instructional method in the acquisition of a variety of skills, such as vocational. By providing step-by-step guidance and immediate feedback, video prompting can facilitate the development of the necessary skills and confidence in learners required to perform effectively a task [7-9]. Furthermore, emerging evidence suggests that the additional procedure of video-based error correction (VBEC), during which participants who fail to complete a step correctly are interrupted and shown the video prompt again within a short time, can promote skill acquisition more efficiently [10]. VBEC in medical education could offer immediate feedback, promote enhanced learning retention, and foster reflective practice. It allows learners to quickly recognize and correct mistakes, aiding in the memorization of correct procedures through visual demonstration. This method also encourages learners to observe their actions critically, supporting self-improvement and deeper understanding. By providing accessible and consistent instructional content, it ensures equitable learning opportunities. Therefore, introducing the VEBC procedure into CPR training performed with video prompting may enhance the residents' learning efficiency and performance.

Objective

In this study, we hypothesized that VBEC would improve the CPR skills of anesthesiology residents.

Methods

Study Design and Participants

This quasi-experimental study was conducted in the Simulation Teaching Center of Hospital among anesthesiology residents between December 2022 and April 2023. Participants who met all of the following criteria were included: (1) ≥18 years old; (2) anesthesiology residents; (3) attended CPR training sessions; and (4) primary skill levels for CPR. Residents in elective rotations were excluded.

Ethical Considerations

Informed consent was obtained from all residents. Ethical approval was received from the Ethics Committee of Shanghai General Hospital (approval #2022KY115). Written informed consent was obtained from all participants.

Intervention

All residents received a web-based pretest consisting of knowledge and practical assessment based on the Resusci Anne trainer (Laerdal Medical Corporation), with the practical assessment involving the execution of the complete steps of CPR on a manikin. Subsequently, all residents attended a didactic lecture using PowerPoint (Microsoft Corporation) on the American Heart Association (AHA) Guidelines for CPR and Emergency Cardiovascular Care (ECC) [11] and watched a standard CPR video. Then, the residents were divided into VBEC and control groups in the 1:1 ratio.

The training process for the VBEC group was a 15-minute video error correction segment. There were ten CPR video segments with errors in the steps, after which trainees discussed and corrected the errors and practiced the correct steps on a manikin (Multimedia Appendix 1). At the end of each video clip, the trainer announced, "Okay, let's discuss this step." The residents were then given 30 seconds to point out the mistake in the video and perform that step on the Resusci Anne trainer in the correct way. This procedure was repeated for all 10 video clips. Further, all residents received 30-minute hands-on standard CPR training, during which they were asked to perform a complete CPR procedure on the Resusci Anne trainer.

The training process for the control group involved a 15-minute video reinforcement segment. There were ten correct CPR step-by-step video segments, after which trainees were asked to practice the correct steps on a manikin following each segment. Then they were asked to perform that step on the Resusci Anne trainer, one at a time. A 30-minute hands-on standard CPR training, the same as that in the VBEC group, was conducted after the video learning in the control group.

At the end of lessons, all residents were asked to independently complete the web-based knowledge test and practical assessment, with the practical test involving performing the entire CPR process on a manikin. The final practical test of each participant was videotaped for evaluation. Finally, all residents completed a web-based questionnaire for self-assessment regarding their performance and satisfaction in this training.

Outcomes and Measurement

Primary outcomes included the posttraining knowledge test and practical assessment scores. Secondary outcomes were the number of residents who correctly conducted CPR at each step, the rate of common mistakes during the CPR process, and the self-assessment results of the residents after the training.

The practical assessment score was the on-target chest compressions (CCs) indicated by the Resusci Anne trainer, that is, the number of CCs meeting the minimum performance metrics for both CC depth and CC rate. The performance and mistakes of the residents were evaluated by both the Resusci Anne trainer and the human trainers, primarily considering the correct steps of high-quality CPR and the number of common mistakes.

Statistical Analysis

Data were analyzed using SPSS (version 15; SPSS Inc). Percentages with numbers in parentheses are used to present categorical data. Continuous data are expressed as mean (SD). Continuous variables were compared using the Student 2-tailed *t* test. Categorical variables were compared with the χ^2 test. A 2-sided *P* value <.05 was considered to indicate a statistically significant difference.

Results

Baseline Characteristics

Of 83 eligible residents, 80 were subjected to analysis as 3 residents missed part of the training session (Figure 1). The baseline characteristics of the residents are presented in Table 1.

Figure 1. Study flowchart. CPR: cardiopulmonary resuscitation.

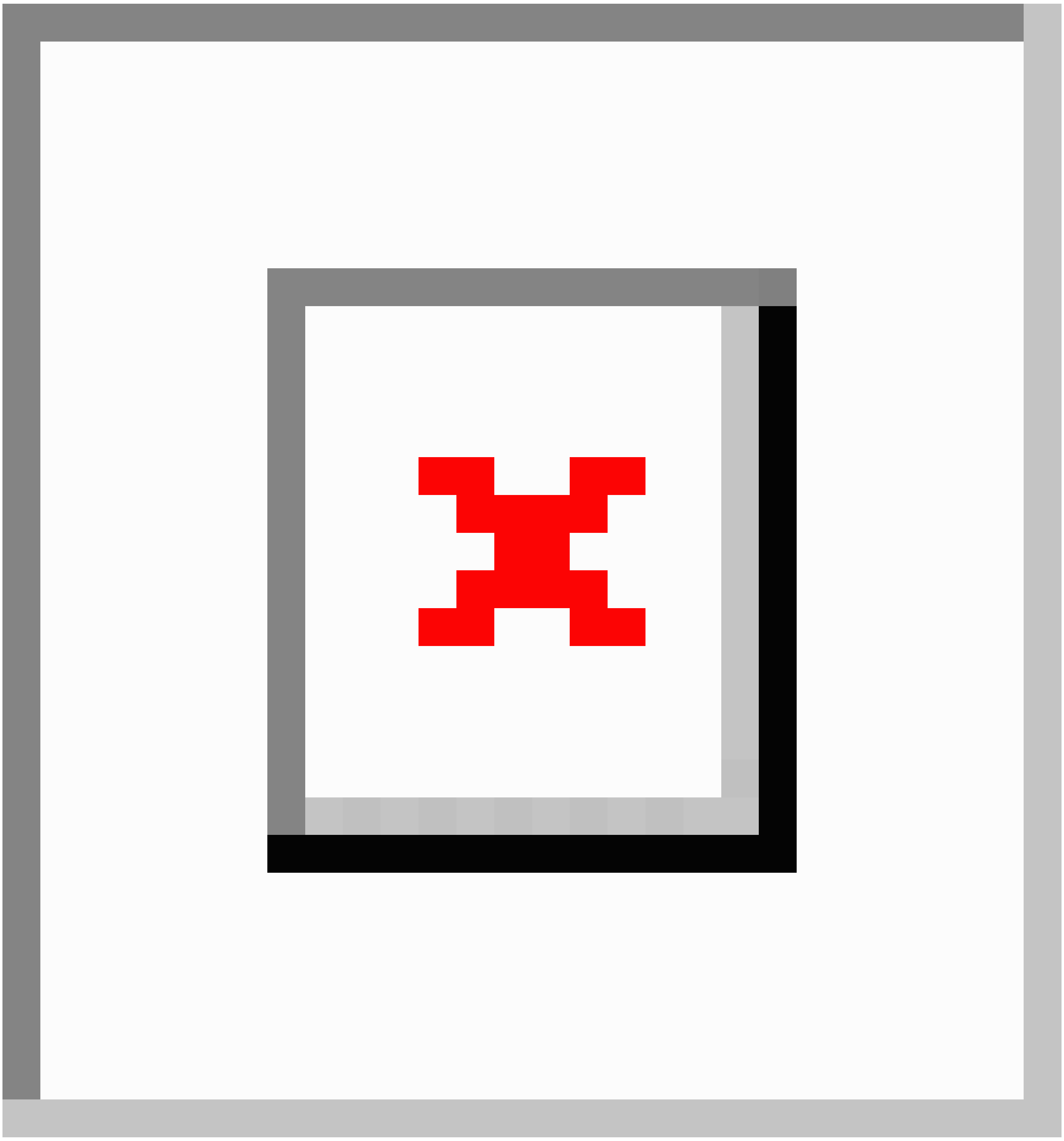


Table . Baseline characteristics (n=80).

Characteristics	VBEC[a] (n=40)	Control (n=40)	*P* value
Age (year), mean (SD)	28 (5)	27 (5)	.45
Male, n (%)	20 (50)	22 (55)	.83
Doctors with postgraduate degrees, n (%)	22 (55)	24 (60)	.66

[a]VBEC: video-based error correction.

Primary Outcome Results

The pretraining theoretical score (63.9, SD 10.7 vs 63.4, SD 11.6; *P*=.85) and practical scores were comparable between the two groups (81.5 , SD 8.5 vs 82, SD 11.7; *P*=.82). The mean posttraining theoretical score of the VBEC group was significantly higher than that of the control group (73, SD 10.5 vs 65.1, SD 11.4; *P*=.002). No significant difference in the posttraining practical scores was observed between the two groups (86.9, SD 8.8 vs 86.3, SD 9.9; *P*=.77) (Table 2).

Table . Knowledge test score and practical assessment score before and after the training between the VBEC[a] and control groups.

	VBEC (n=40)	Control (n=40)	*P* value	*t* (*df*)	95% CI	Effect size
Knowledge test score, mean (SD)						
Before training	63.9 (10.7)	63.4 (11.6)	.85	0.190 (78)	−4.490 to 5.440	.043
After training	73 (10.5)	65.1 (11.4)	.002	3.228 (78)	3.028 to 12.772	.722
Practical assessment score, mean (SD)						
Before training	81.5 (8.5)	82 (11.7)	.82	−0.229 (78)	−5.082 to 4.032	.051
After training	86.9 (8.8)	86.3 (9.9)	.77	0.299 (78)	−3.537 to 4.787	.067

[a]VBEC: video-based error correction

Secondary Outcome Results

In the stepwise comparison of CPR performance, performance for the "Secure own safety" step was significantly higher in the VBEC group (85% vs 45%; *P*<.001). Among the common mistakes, residents in the VBEC group showed a lower error rate of "failure to anticipate the next move" (7.5% vs 32.5%; *P*=.01) and "failure to debrief or problem solve after the code" (5% vs 27.5%; *P*=.01) than the control group (Table 3).

Table . Comparison of correct performance and common mistakes of cardiopulmonary resuscitation between the VBEC[a] and control groups.

	VBEC, n (%)	Control, n (%)	*P* value
Alert emergency services	34 (85)	36 (90)	.74
Recognize illness/injury	40 (100)	40 (100)	>.99
Secure own safety	34 (85)	18 (45)	<.001
Examine patient	34 (85)	30 (75)	.40
Recovery position	36 (90)	38 (95)	.68
Decide if CPR[b] should be started	38 (95)	32 (80)	.09
Effective chest compressions	38 (95)	35 (87.5)	.43
Open the airway and check breathing	37 (92.5)	36 (90)	>.99
Give rescue breaths	40 (100)	38 (95)	.49
Check rhythms and resume CPR	39 (97.5)	35 (87.5)	.20
Failure to recognize arrest	2 (5)	3 (7.5)	>.99
Failure to act (or act rationally) and failure to call for help appropriately	2 (5)	8 (20)	.09
Failure to provide effective compressions	2 (5)	5 (12.5)	.43
Failure to provide effective ventilations for the patient	0 (0)	1 (2.5)	>.99
Failure to anticipate the next move	3 (7.5)	13 (32.5)	.01
Failure to debrief or problem solve after the code	2 (5)	11 (27.5)	.01

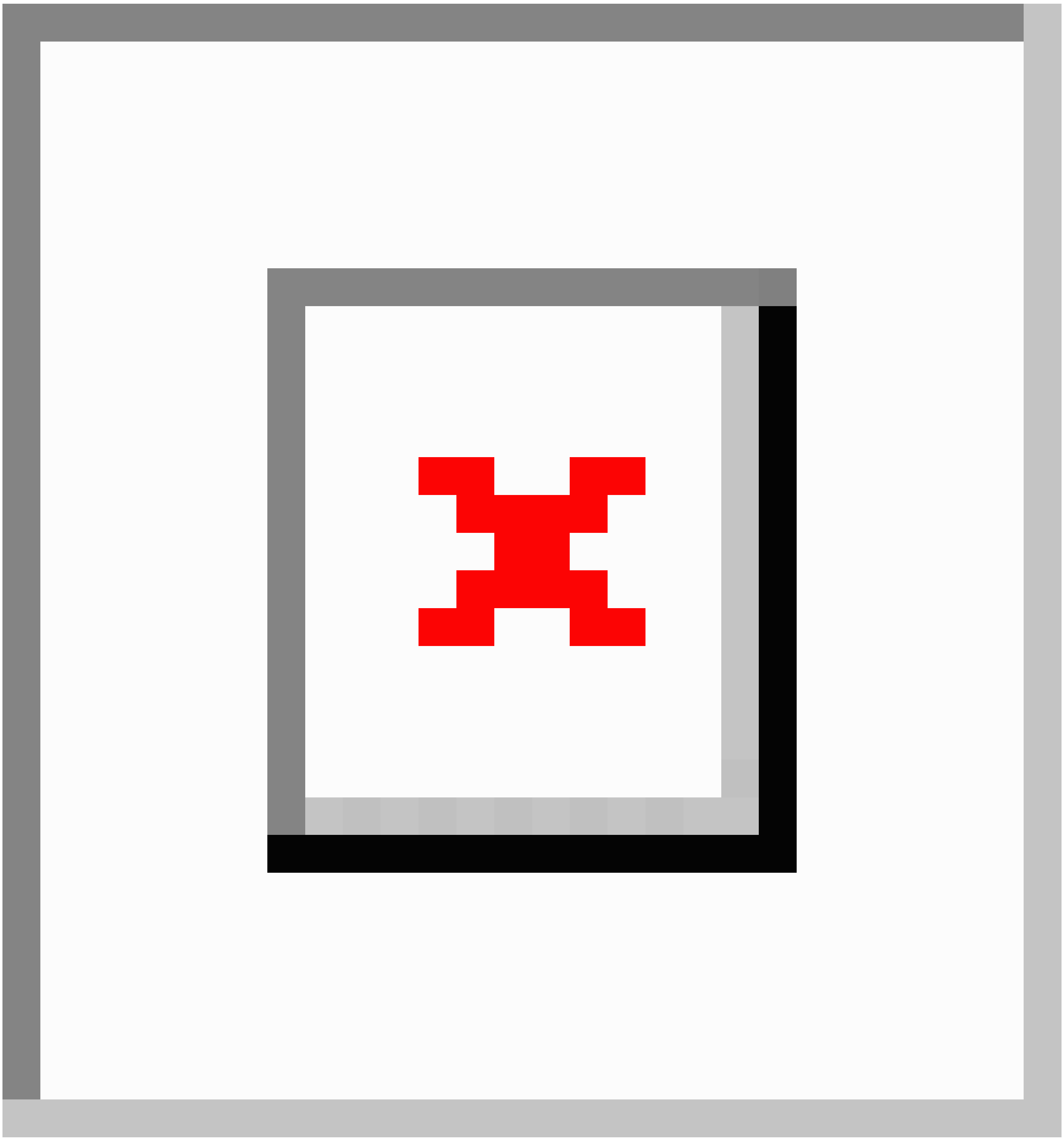
[a]VBEC: video-based error correction.

[b]CPR: cardiopulmonary resuscitation.

Performance After Training

The residents' perceptions after the training regarding self-assessment of their performance and related experiences are presented in Figure 2. The residents in the VBEC group had significantly higher confidence in CPR than the control group (n=?, 62.5% vs n=?, 35%; *P*=.03).

Figure 2. Participants' positive reactions to activities (%). VBEC: video-based error correction. Asterisk indicates statistical significance.



Discussion

Principal Findings

In this study, it was revealed that VBEC improved the quality of CPR training of anesthesiology residents. Specifically, the VBEC-trained residents had improved posttraining theoretical scores, a lower error rate for common mistakes, and better performance at certain steps of the CPR procedure compared with the control group. The residents who received the VBEC training also showed higher confidence in performing basic CPR than the control group. These findings suggest that VBEC may be applicable in anesthesiology residents' training.

Comparison to Prior Work

High-quality CPR can dramatically increase the chances of survival in cardiac arrest. The AHA CPR Guidelines emphasize that to bridge the gap between knowledge and practice, educators should develop a method for improving CPR education [12,13]. Errorless, or near-errorless, learning procedures involve attempting to prevent errors during all teaching sessions [14]. A video prompting technique is to begin teaching with the most assistive prompt to minimize the likelihood of an error [15]. Error correction, on the other hand, involves procedures that are employed following an incorrect response that would increase the probability of a correct response on subsequent trials [16]. Previous research has evidenced the effectiveness of the error correction procedures and has confirmed their benefits [15,17-20]. In this study, we used "error correction learning" as the primary method to improve CPR education and optimize the studying experience. VBEC may provide the learners with the opportunity to think and respond independently to the situation before external instructions are given, which may deepen their understanding of the CPR procedure. This approach may explain the improvement in the posttraining knowledge test among residents in the VBEC group. While some studies argue against the use of error correction since this may cause trainees to pick up incorrect information [21], this study does not support such a conclusion since none of the outcomes of the VBEC group were inferior to those of the control group. On the contrary, the VBEC group had a lower chance of making mistakes that are common in CPR procedures and paid more attention to the rescuer's safety. Therefore, the VBEC may highlight the error-prone points in standard CPR training. Moreover, Goodson et al demonstrated that VBEC is effective for learners who do not fully benefit from video prompting, as all participants achieved 100% accuracy in task analysis. This outcome was noteworthy, especially since the study included individuals with developmental disabilities [10]. Considering the group-based nature of medical education, this suggests that VBEC could be a valuable tool in the training of medical residents.

Strengths

No difference was found between the VBEC and the control groups in practical performance after the training. This result may be attributed to the fact that on-target CCs were used to evaluate the practical performance; CC is a muscle memory-based skill that requires more physical practice. However, the 15-minute VBEC or video prompt training plus the 30-minute hands-on training is relatively short for substantially improving the practical performance of the residents. Nevertheless, CC is a central but not the only key part of the whole CPR procedure. Successful CPR in real situations also relies on non-technical skills, such as fluency in the CPR code and physicians' self-confidence [22,23]. Our results showed that the VBEC group had a lower error rate and higher self-confidence in conducting CPR. More specifically, the fewer "failure to anticipate the next move" in the VBEC group may indicate an enhanced fluency throughout the CPR code. These improvements are essential for the improvement of the overall CPR quality, even with the same CC.

Limitations

This study is not without limitations. First, the small sample size limits the generalizability of the results, indicating a need for future studies with larger populations to confirm these findings. Second, the brief duration of the testing period suggests the necessity for extended follow-up to fully understand the long-term implications. Third, further investigation is required to explore how VBEC may be applied to improve practical skills such as CC. Moreover, the study's design constraints, including the absence of power estimation, rigorous randomization, and blinding, hinder our ability to draw definitive causal conclusions, warranting a cautious interpretation of the results.

Conclusions

In summary, VBEC may be an efficient teaching technique used in CPR training, not only for improving residents' cognitive performance and self-confidence but also for increasing the rate of completion in providing a fluent CPR sequence with fewer mistakes. However, further research with longer-term follow-up periods is needed to determine whether error correction learning improves long-term outcomes.

Data Availability

The dataset supporting the results of this work is included in the article.

Authors' Contributions

YW and JF conceived and coordinated the study; designed, performed, and analyzed the study; and wrote the paper. SW, HW, WG, and LH carried out the data collection, data analysis, and revised the paper. All authors reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scenario script of error-correction learning.
[DOCX File, 15 KB - games_v12i1e53577_app1.docx]

References

1. Merchant R, Yang L, Becker L, et al. American Heart Association get with the guidelines-resuscitation investigators: incidence of treated cardiac arrest in hospitalized patients in the United States. Crit Care Med 2011;39(11):2401-2406. [doi: 10.1097/CCM.0b013e3182257459]
2. Rao P, Kern KB. Improving community survival rates from out-of-hospital cardiac arrest. Curr Cardiol Rev 2018;14(2):79-84. [doi: 10.1097/CCM.0b013e3182257459]

3. Hubail D, Mondal A, Al Jabir A, Patel B. Comparison of a virtual reality compression-only cardiopulmonary resuscitation (CPR) course to the traditional course with content validation of the VR course – a randomized control pilot study. Ann Med Surg 2022;73:103241. [doi: 10.1016/j.amsu.2022.103241]
4. Potts J, Lynch B. The American Heart Association CPR Anytime Program: the potential impact of highly accessible training in cardiopulmonary resuscitation. J Cardiopulm Rehabil 2006;26(6):346-354. [doi: 10.1097/00008483-200611000-00002] [Medline: 17135853]
5. Wik L, Kramer-Johansen J, Myklebust H, Sreb H. Quality of cardiopulmonary resuscitation during out-of-hospital cardiac arrest. ACC Curr J Rev 2005 Apr;14(4):45-46. [doi: 10.1016/j.accreview.2005.03.041]
6. Abella BS, Alvarado JP, Myklebust H, et al. Quality of cardiopulmonary resuscitation during in-hospital cardiac arrest. J Am Med Assoc 2005 Jan 19;293(3):305-310. [doi: 10.1001/jama.293.3.305] [Medline: 15657323]
7. Cannella-Malone HI, Fleming C, Chung YC, Wheeler GM, Basbagill AR, Singh AH. Teaching daily living skills to seven individuals with severe intellectual disabilities: a comparison of video prompting to video modeling. J Posit Behav Interv 2011 Jul;13(3):144-153. [doi: 10.1177/1098300710366593]
8. Heider AE, Cannella-Malone HI, Andzik NR. Effects of self-directed video prompting on vocational task acquisition. Career Dev Transit Except Individ 2019 May;42(2):87-98. [doi: 10.1177/2165143417752901]
9. Cannella-Malone HI, Jimenez ED, Schaefer JM, Miller M, Byrum H. Examination of the effects of video prompting across different types of tasks. Career Dev Transit Except Individ 2018 Nov;41(4):200-211. [doi: 10.1177/2165143417739609]
10. Goodson J, Sigafoos J, O'Reilly M, Cannella H, Lancioni GE. Evaluation of a video-based error correction procedure for teaching a domestic skill to individuals with developmental disabilities. Res Dev Disabil 2007 Oct;28(5):458-467. [doi: 10.1016/j.ridd.2006.06.002]
11. Neumar RW, Otto CW, Link MS, et al. Part 8: adult advanced cardiovascular life support: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2010 Nov 2;122(18 Suppl 3):S729-S767. [doi: 10.1161/CIRCULATIONAHA.110.970988] [Medline: 20956224]
12. Peberdy MA, Callaway CW, Neumar RW, Geocadin RG, Kronick S. Part 9: post-cardiac arrest care: 2010 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. . 2010(18 Suppl 3) p. S768-S786. [doi: 10.1161/CIRCULATIONAHA.110.971002]
13. Chien CY, Tsai SL, Huang CH, et al. Effectiveness of blended versus traditional refresher training for cardiopulmonary resuscitation: prospective observational study. JMIR Med Educ 2024 Apr 29;10:e52230. [doi: 10.2196/52230] [Medline: 38683663]
14. Mueller MM, Palkovic CM, Maynard CS. Errorless learning: review and practical application for teaching children with pervasive developmental disorders. Psychol Sch 2007 Sep;44(7):691-700. [doi: 10.1002/pits.20258]
15. Leaf JB, Cihon JH, Ferguson JL, Milne CM, Leaf R, McEachin J. Comparing error correction to errorless learning: a randomized clinical trial. Anal Verbal Behav 2020 Jun;36(1):1-20. [doi: 10.1007/s40616-019-00124-y] [Medline: 32699736]
16. Cariveau T, La Cruz Montilla A, Gonzalez E, Ball S. A review of error correction procedures during instruction for children with developmental disabilities. J Appl Behav Anal 2019 May;52(2):574-579. [doi: 10.1002/jaba.524] [Medline: 30468249]
17. Kessels RPC, van Loon E, Wester AJ. Route learning in amnesia: a comparison of trial-and-error and errorless learning in patients with the Korsakoff syndrome. Clin Rehabil 2007 Oct;21(10):905-911. [doi: 10.1177/0269215507077309] [Medline: 17981849]
18. Evans JJ, Wilson BA, Schuri U, et al. A comparison of "errorless" and "trial-and-error" learning methods for teaching individuals with acquired memory deficits. Neuropsychol Rehabil 2000 Jan;10(1):67-101. [doi: 10.1080/096020100389309]
19. Carroll RA, Joachim BT, St Peter CC, Robinson N. A comparison of error-correction procedures on skill acquisition during discrete-trial instruction. J Appl Behav Anal 2015;48(2):257-273. [doi: 10.1002/jaba.205] [Medline: 25913105]
20. Plaisance L, Lerman DC, Laudont C, Wu WL. Inserting mastered targets during error correction when teaching skills to children with autism. J Appl Behav Anal 2016 Jun;49(2):251-264. [doi: 10.1002/jaba.292] [Medline: 26790395]
21. Gast DB. A rejoinder to leaf: what constitutes efficient, applied, and trial and error. Evid Based Commun Assess Interv 2011 Dec;5(4):234-238. [doi: 10.1080/17489539.2012.689601]
22. Kern KB, Hilwig RW, Berg RA, Ewy GA. Efficacy of chest compression-only BLS CPR in the presence of an occluded airway. Resuscitation 1998 Dec;39(3):179-188. [doi: 10.1016/s0300-9572(98)00141-5] [Medline: 10078808]
23. Verplancke T, De Paepe P, Calle PA, De Regge M, Van Maele G, Monsieurs KG. Determinants of the quality of basic life support by hospital nurses. Resuscitation 2008 Apr;77(1):75-80. [doi: 10.1016/j.resuscitation.2007.10.006] [Medline: 18083286]

Abbreviations

CC: chest compression
CPR: cardiopulmonary resuscitation
VBEC: video-based error correction

Edited by C Prahm, TDA Cardoso; submitted 11.10.23; peer-reviewed by R Khanmohammadi, W Langewitz; revised version received 01.07.24; accepted 02.08.24; published 03.10.24.

Please cite as:
Wang Y, Fu J, Wang S, Wang H, Gao W, Huang L
Impact of Video-Based Error Correction Learning for Cardiopulmonary Resuscitation Training: Quasi-Experimental Study
JMIR Serious Games 2024;12:e53577
URL: https://games.jmir.org/2024/1/e53577
doi:10.2196/53577

An Augmented Reality Serious Game for Children's Optical Science Education: Randomized Controlled Trial

Bo Liu, MA; Xinyue Wan, MA; Xiaofang Li, MA; Dian Zhu, PhD; Zhao Liu, PhD

School of Design, Shanghai Jiao Tong University, Shanghai, China

Corresponding Author:
Zhao Liu, PhD

Abstract

Background: Knowledge construction in the context of children's science education is an important part of fostering the development of early scientific literacy. Nevertheless, children sometimes struggle to comprehend scientific knowledge due to the presence of abstract notions.

Objective: This study aimed to evaluate the efficacy of augmented reality (AR) games as a teaching tool for enhancing children's understanding of optical science education.

Methods: A total of 36 healthy Chinese children aged 6-8 years were included in this study. The children were randomly divided into an intervention group (n=18, 50%) and a control group (n=18, 50%). The intervention group received 20 minutes of AR science education using 3 game-based learning modules, whereas the control group was asked to learn the same knowledge for 20 minutes with a non-AR science learning app. Predict observe explain tests for 3 topics (animal vision, light transmission, and color-light mixing) were conducted for all participants before and after the experiment. Additionally, the Intrinsic Motivation Inventory, which measures levels of interest-enjoyment, perceived competence, effort-importance, and tension-pressure, was conducted for children after the experiment.

Results: There was a statistically significant difference in light transmission (z=−2.696; P=.008), color-light mixing (z=−2.508; P=.01), and total predict observe explain test scores (z=2.458; P=.01) between the 2 groups. There were also variations between the groups in terms of levels of interest-enjoyment (z=−2.440; P=.02) and perceived competence (z=−2.170; P=.03) as measured by the Intrinsic Motivation Inventory.

Conclusions: The randomized controlled trial confirmed that the AR-based science education game we designed can correct children's misconceptions about science and enhance the effectiveness of science education.

Trial Registration: ClinicalTrials.gov NCT06184022; https://classic.clinicaltrials.gov/ct2/show/NCT06184022

(JMIR Serious Games 2024;12:e47807) doi:10.2196/47807

KEYWORDS

augmented reality; serious game; science education; childhood education; cognition; children; scientific cognition; cognitive process; effectiveness

Introduction

Children's level of scientific concept generation is representative of their inquiry, comprehension, and application of natural events and phenomena and reveals their cognitive capacities and developmental stages [1]. Knowledge construction in children's science education contributes to early scientific literacy development, which improves children's cognitive level by enhancing thinking skills, and is being emphasized by scholars and parents [2]. Traditionally, children build domain knowledge in science through films, literature, and lectures in science education [3]. Although some forms of educational learning are accessible, they often uses a monotonous instructional format, and confusing content hinders the transmission of scientific knowledge [4].

Serious games provide a more engaging interactive environment and an accessible cognitive framework to facilitate effective learning [5]. Studies have shown that serious games have more effective learning outcomes than traditional methods of science education (eg, face-to-face lectures and book-based knowledge transfer) [6,7]. It is suitable for children's investigation of natural phenomena because the game's visual design simulates paranormal phenomena that cannot be produced in real life. Lester et al [8] constructed virtual environments that generate natural phenomena, allowing children to assume roles in open-world environments and to freely rely on their knowledge of the geography and biology of natural environments. Laine et al [9] permitted children to interact with hosts in virtual narrative game scenarios and to investigate the geometry of the virtual environment with the protagonist. The concept of light, a prevalent natural phenomenon, was selected as the subject of this research to explore its design for enhancing children's cognitive abilities. Optical science education programs are still presented in a 2D format, which has been demonstrated to be ineffective [10].

Due to the spatial complexity and abstract nature of optics, it is challenging to accurately convey knowledge through flat visual representations [11]. Therefore, it is necessary to blur the boundaries between the 3D real world and the 2D digital world to reduce the distance between children's learning of science concepts and their learning environments [12]. 3D representations and interactions in augmented reality (AR) games have the potential to enhance spatial cognition, thereby facilitating children's comprehension of spatially abstract scientific concepts [13], such as simulating the movement of the sun in a classroom environment [14]. Sahin and Yilmaz [15] demonstrated that students who used AR technology to improve their science literacy performed better on tests than those who learned using traditional methods. This is as a result of AR technology's ability to enhance the dynamic potential of human consciousness to comprehend the science learning process [16]. In addition, motivational improvement was mentioned as one of the frequently observed AR outcomes [17]. Using AR apps increased student motivation relative to other instructional aids [18]. Our study investigated whether designing optical science education with more comprehensible 3D interactions for children can enhance science education and promote children's motivation.

The study designed the "AR Serious Game for Optical Science" and conducted a randomized controlled trial to determine the efficacy of this AR game product in enhancing children's science education. The primary objective of this study was to validate the efficacy of AR science education games for children; the secondary objective was to investigate the intrinsic motivation of children toward them.

Methods

Study Design

Guardians of children with independent mobility provided informed written consent for their participation in the study. Participants were randomly assigned to the intervention and control groups using a randomization list, which was maintained by members of the study group uninvolved in any other aspect of the study. Participants' guardians received and opened opaque, sealed envelopes containing group assignments following the initial evaluation. The evaluator in charge of assessing the results of the AR science education course had no access to participant information or group assignment.

Sample size calculations were performed using PASS software (NCSS LLC) based on the predict observe explain (POE) test scores from the preintervention questions. Group sample sizes of 18 and 18 achieve 90.118% power to reject the null hypothesis of equal means, when the population mean difference is $\mu1 - \mu2 = 3.2 - 1.0 = 2.2$, with SDs of 2.0 and 1.9 for the 2 groups and with a significance level (α) of .05 using a bilateral, 2-sample, equal-variance, 2-tailed *t* test.

Participants

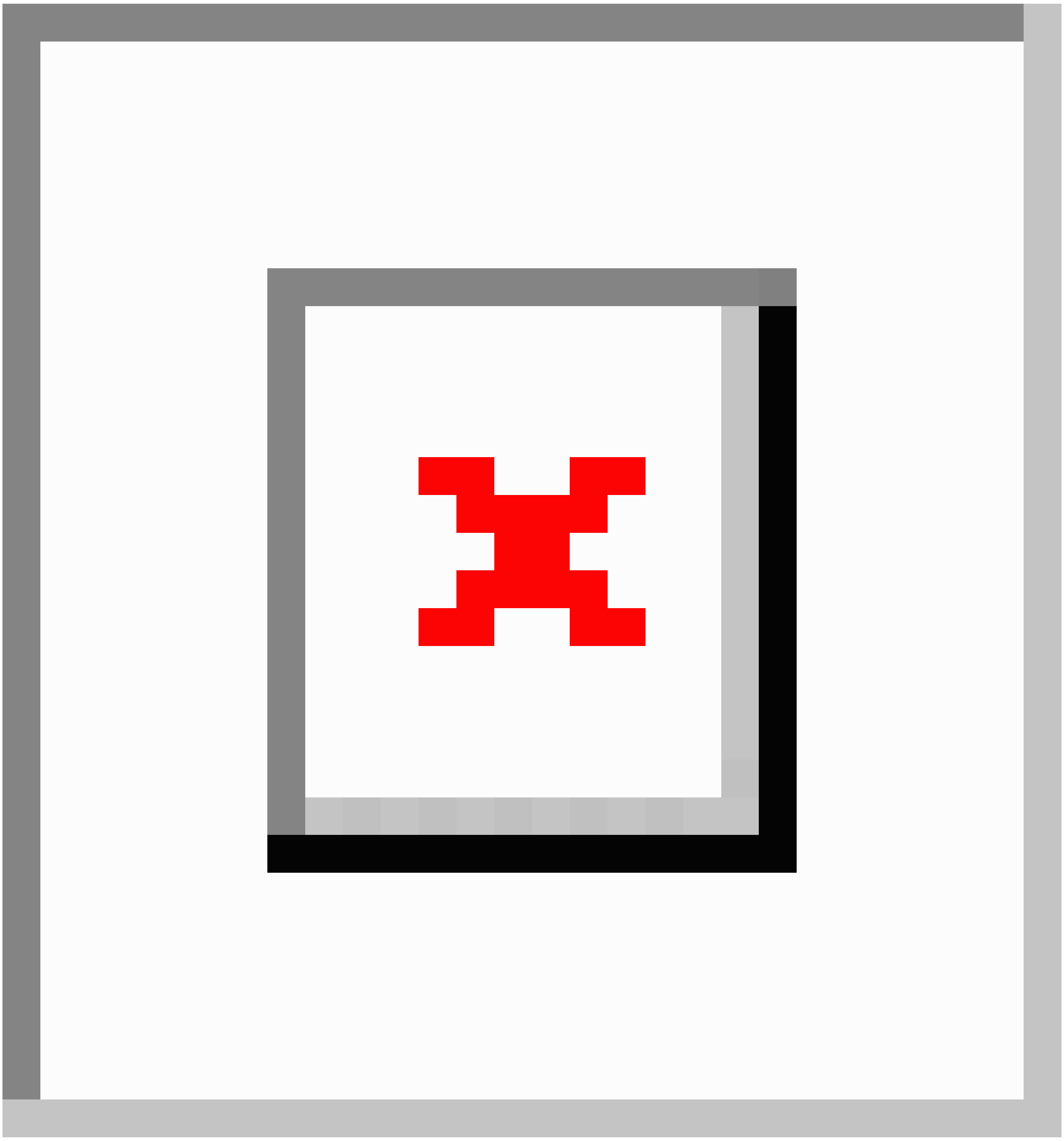
A total of 36 Chinese children (aged 6-8 y) were recruited from Jiangyin Children's Education Center and Jiangyin Wuxi Community in Jiangsu Province and divided into the intervention (n=18, 50%) and control (n=18, 50%) groups.

AR Science Education Game Design

During the learning phase, children are required to engage in physical activities, such as walking around with a handheld device, to interact with the AR scene's content to discover what is unique about the light phenomenon. When children touch the interactive points, the content is explained by animation and voice-overs. This study developed several interactive approaches for children within AR games, such as through in-game visual representations, speech, and interactive methods, which permit children to connect game content to unfamiliar information as they explore. The advantages are as follows: (1) children can use more familiar physical activities with light concepts to establish metaphorical mappings related to orientation, not just gestural touch; (2) rendering light with 3D attributes in the real world reduces the cognitive load generated by children's linkage of abstract knowledge and the real phenomena; and (3) adding various kinds of digital augmentation effects in the AR scene helps children understand the concepts. The project created 3 games based on the characteristics of scientific understanding (Multimedia Appendix 1 and Figure S1 in Multimedia Appendix 2 [9,14,19-28]).

Game 1 introduces children to the fundamentals of animal vision (Figure 1). Animal vision concepts are investigated through AR scenes. By clicking on the icons in the lower-left corner, the game transforms to an animal simulation. In each scenario, a voice-over narration instructs children to identify the visual differences between the animal and the human. When the handheld device is trained on a specific target, a voice-over narration and feedback animation will play.

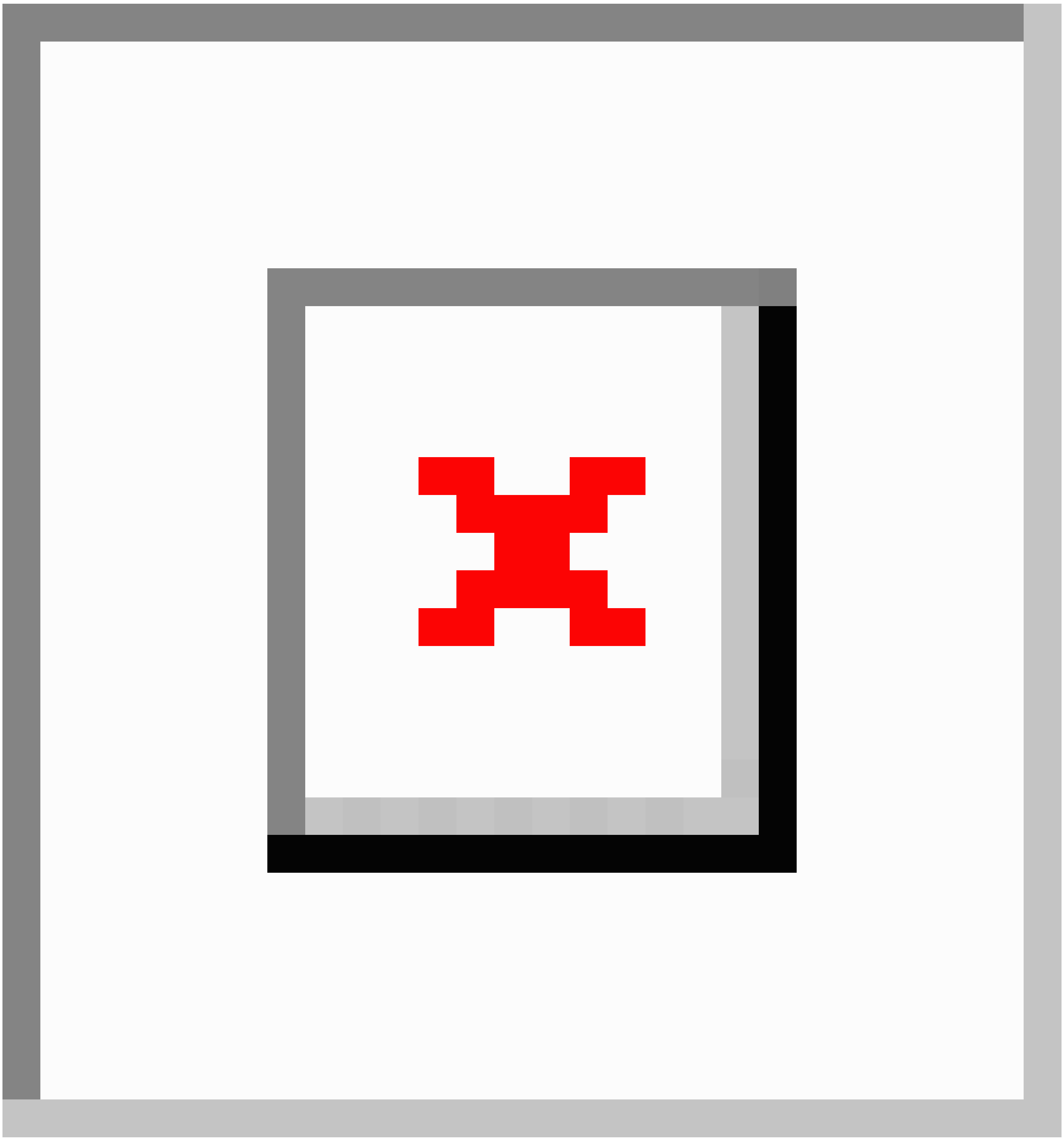
Figure 1. Animal vision augmented reality (AR) game introduction.



In the design of the interaction mode, 3 display modes were established for the game's interactive elements: far, medium, and near (hybrid camera mode). The concept of invisible light is introduced to children in greater detail based on the ray distance between the device's camera and the target element. The far view provides children with an intuitive impression of the invisible light's overall effect; the medium view uses transition animation to illustrate the invisible light's characteristics; and the near view uses special effect particles to illustrate the invisible light's trajectory.

Game 2 introduces children to light transmission–related concepts (Figure 2). In the AR scenario, children navigate the environment with a handheld device and activate energy panels by interacting with flat mirrors and optics. By targeting AR-enhanced prop objects and manipulating the angle of light emission to investigate how light propagates, voice-over explanations and feedback animations are activated.

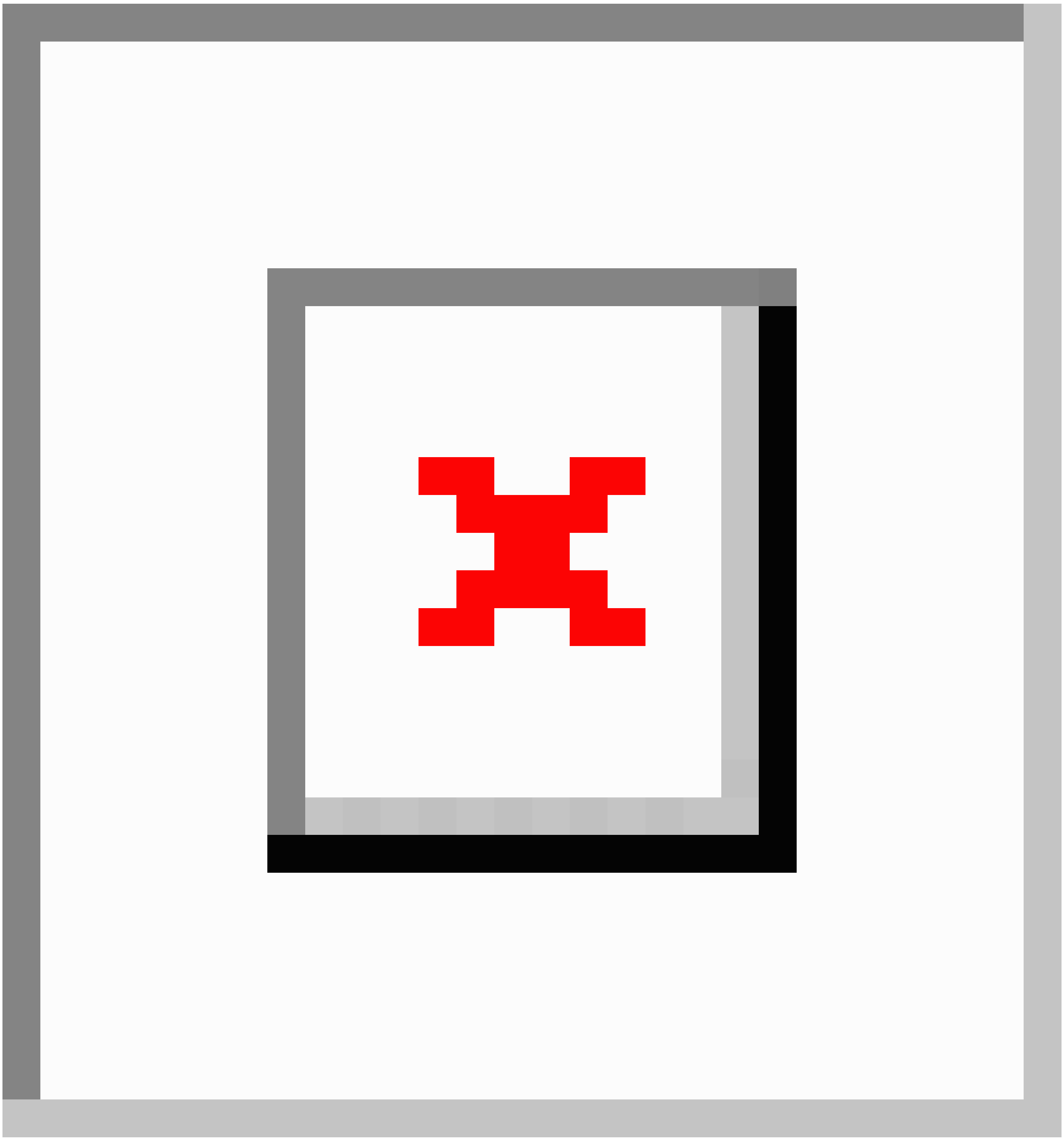
Figure 2. Light transmission augmented reality (AR) game introduction.



According to the voice-over prompts, children can hold the device and manipulate the flashlight from a first-person perspective (spatial exploration mode) as part of the interactive design. They then complete 3 steps: locating the interactive elements (mirrors, ice crystals, etc), adjusting the flashlight's tilt angle, and using the flashlight to complete the light-up task. The progression encourages children to investigate the principles of light transmission through the game.

Game 3 introduces children to color-light mixing concepts (Figure 3). Children were instructed to walk around with the device in hand and explore the color changes of props such as AR-enhanced birds, which are illuminated with various colors of lights. Collecting the target color's shadow initiates a voice-over explanation and feedback animation.

Figure 3. Color-light mixing augmented reality (AR) game introduction.



Regarding the interactive design, children need to hold the device to illuminate the creatures and cast shadows on the present wall, and then they need to press the button to turn the light on and off (projection irradiation mode). The objective of the game required children to perform single-color illumination, 2-color mixing, and 3-color mixing to achieve the desired hue. In another vibrant nursery game, children were instructed to move plants to receive various colors of light and to observe the plants' root elongation and leaf dispersal.

This game design used Unity 3D (Unity Technologies) as the development engine, and the app was installed on an Apple iPad (2018) with a screen resolution of 2048×1536 (264 pixels per inch). The AR component made use of the Vuforia AR SDK (Parametric Technology Corporation) to accomplish the fundamental duties of plane identification and virtual object generation. The interaction section used lens focus to determine the interactions; when the device camera's output rays collide with the target virtual object and the distance is close, it is deemed to have located the target effectively. To imitate the illusion of invisible light, Unity's *Post Processing* module was applied to the camera filter. The principle entailed presenting the camera screen into the buffer of Unity and applying filters and effects prior to displaying it; it can be applied to both the camera screen and the virtual item.

Procedure

This experiment was a randomized controlled trial, and the participants were randomly separated into the intervention group and the control group. The random numbers were generated by applying the SAS software analysis system (SAS Institute) on a computer simulation, and no experimental group was allowed to be selected at random. Every child was tested in the company of a guardian and 2 researchers.

The independent variable was the type of game (an optical science education app called "Light and Color" or the AR game we designed; see Figure S2 in Multimedia Appendix 2 for a comparison of the differences between the 2 games). The dependent variables for both intervention and control group participants were the differences between the pre- and posttest results of the POE tests and the children's motivation to play the game. To create control variables for the experiment, both games included the topics of animal vision, light transmission, and color-light mixing, and neither game involved a human teacher. In addition, there were no significant sex (P=.49) or age (P=.67) differences between the 2 groups.

Intervention Group

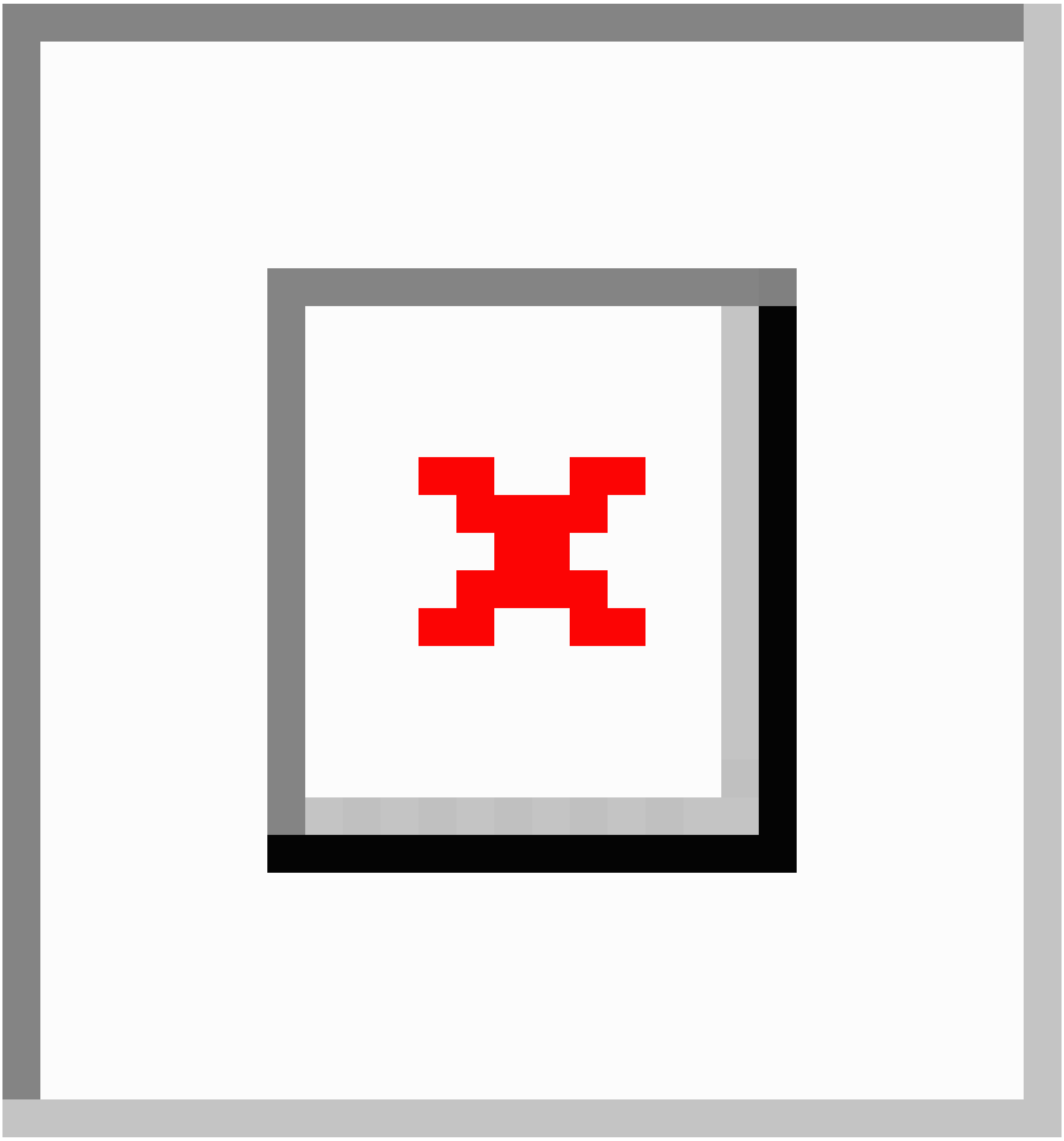
Before the test started, the researcher provided the basic information of the experiment to the participants, including the test topic, test technique, test time, and other information. The participants were asked to complete a cognitive exam on the notion of light and perform a POE test for each topic to find out how well they comprehend the content, without being told whether their answers were correct.

After completing the pretest, intervention group participants were instructed to complete the 3 game-based learning modules of the AR science education app on the iPad regarding animal vision, light transmission, and color-light mixing. On their initial encounter with the game, respondents were given around 10 minutes to comprehend its mechanics. The intervention group's total learning time was limited to 20 minutes, the testing process was completed under the supervision of the instructor and the experimenter, and the children's behavioral characteristics were recorded. During the experiment, the participants were not disturbed in any way; researchers only intervened when they faced difficulties or requested assistance. The participants were given a 15-minute respite at the conclusion of the trial to take another POE test. Before and after the experiment, each participant's performance on the game was recorded. The researcher then read aloud and described the items on the intrinsic motivation and cognitive load scales to the participants, who scored the scale items using a 5-point "smiley face" scale.

Control Group

The control group was also introduced to the experiment and given a preintervention POE test to assess their prior knowledge of the learning material. The control group completed the same 3 game courses for a maximum of 20 minutes using the non-AR app "Light and Color" after completing the pretest. The participants took a 15-minute break at the conclusion of the trial to complete another POE test and the Intrinsic Motivation Inventory (IMI) scale (Figure 4).

Figure 4. Photos of the experimental process: (A) the process of using the augmented reality (AR) game for participants in the intervention group; (B) the process of using the "Light and Color" app game for participants in the control group, and (C) the process of filling out the questionnaire by the participants.



Evaluation Metrics

The study was validated based on several experiments.

The POE test is commonly used in science classes and tries to expose students' expectations about certain events and the rationale for these predictions [29]. It is used to demonstrate scientific experiments to pupils and is advantageous for fostering children's critical thinking and assessing students' grasp of scientific topics. The investigator then displays the relevant physical events to the students using basic prop materials after requiring the students to independently determine the correct answers to the questionnaire along with their justifications. Finally, students are instructed to alter or supplement their explanations in light of the observations. Since children may appear to be able to answer the question properly but not comprehend the reasoning behind it, for each topic, it is possible that they do not comprehend the underlying concept. In this study, individuals' accurate answers and explanations were recorded, and different situations were rated differently based on a 2-tier test [30] (Table 1). This scoring method is frequently used to evaluate students' conceptual understanding [31]. The outcomes were categorized as correct answer+correct explanation, correct answer+incorrect explanation, incorrect answer+correct explanation, and incorrect answer+incorrect explanation. Each topic's overall score was included in the subsequent analysis. To avoid disruptions caused by children's memorization of answers, the experimental posttest questionnaire in this study was different from the pretest questionnaire but was founded on the same scientific concepts. The examination topics are provided in Table S1 in Multimedia Appendix 2.

Table . Two-tier test assessment criteria.

Level of conceptual understanding	Score
Correct answer+correct explanation	2
Correct answer+incorrect explanation	1
Incorrect answer+correct explanation	1
Incorrect answer+incorrect explanation	0

Due to the young age of the study participants, the simplified version of the IMI adapted by Vos et al [32] was selected for this research. It was developed under game conditions with 3 subscales: interest-enjoyment, perceived competence, and effort-importance, to assess the perceived levels of motivation, enjoyment, and perceived difficulty of the participants. To investigate the negative emotions of children using the AR game, the study inserted questions from the original scale's tension-stress section [33] (Table S2 in Multimedia Appendix 2). Participants were asked to rate the extent to which they concurred with the statement using a 5-point Likert scale depicting 5 smiling faces. A score of 5 indicated that the child participant strongly agreed with the statement. To minimize the effect of differences in reading ability, the researcher read the questionnaire audibly to the child participants, who then completed the questionnaire independently.

Ethical Considerations

The study was approved by the Human Research Ethics Committee of Shanghai Jiao Tong University (H2022041I) in China. Informed consent was signed by guardians and the data were deidentified. A toy with a value of CNY ¥50 (US $7.01) was provided as compensation.

Results

A total of 36 healthy Chinese children aged 6-8 years were recruited in May 2022, including 22 male and 14 female children, all of whom participated in the experiment with the consent of their guardians and of their own volition. The 36 participants were randomly assigned to the intervention group (n=18, 50%) and the control group (n=18, 50%), with the mean age of the intervention group being 7.16 (SD 0.76) years and that of the control group being 7.06 (SD 0.78) years. Baseline demographic data and POE test scores for the intervention and control groups are shown in Table 2. The statistical analysis revealed that there was no statistically significant distinction observed between the 2 groups across all variables (all *P*>.05). This suggests that the intervention and control groups exhibited a similar overall comprehension level prior to the commencement of the trial. The experimental procedure is provided in Figure 5.

Table . Baseline data for the intervention and control groups.

Variable	Intervention group[a] (n=18)	Control group[b] (n=18)	*z* score or chi-square (*df*)	*P* value[c]
Male sex, n (%)	10 (56)	12 (67)	0.467 (1)[d]	.49
Age (y), mean (SD)	7.17 (0.76)	7.06 (0.78)	−0.421[e]	.67
POE[f] test score for animal vision, mean (SD)	1.83 (1.10)	1.94 (1.21)	−0.296[e]	.77
POE test score for light transmission, mean (SD)	2.83 (1.72)	2.67 (2.20)	−0.437[e]	.66
POE test score for color-light mixing, mean (SD)	1.50 (1.09)	1.94 (1.16)	−1.031[e]	.30
Total POE test score, mean (SD)	6.17 (2.28)	6.56 (2.12)	−0.273[e]	.78

[a]Augmented reality game.

[b]Non–augmented reality game.

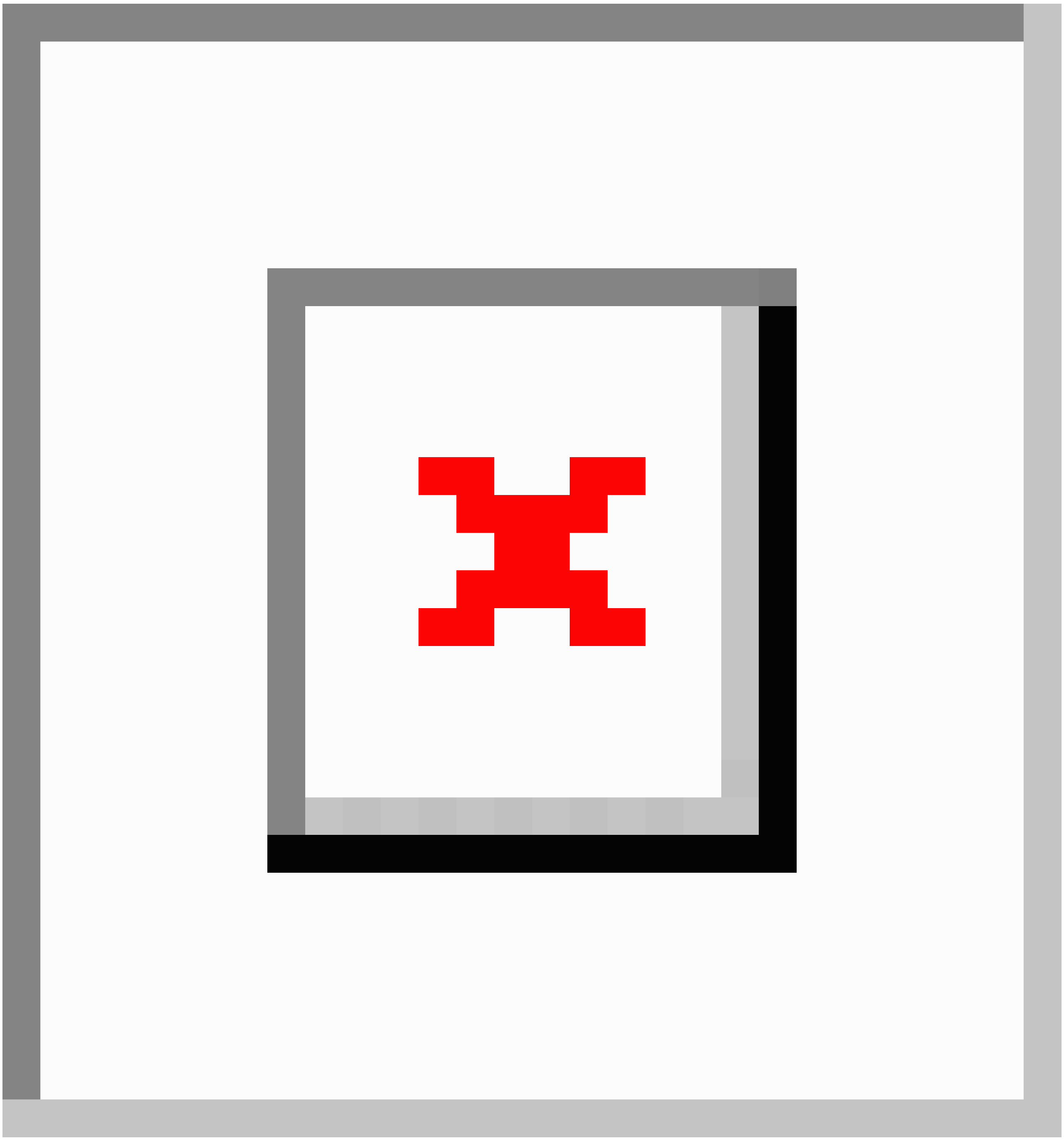
[c]Mann-Whitney *U* test and χ^2.

[d]Chi-square value.

[e]*z* score.

[f]POE: predict observe explain.

Figure 5. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. AR: augmented reality; IMI: Intrinsic Motivation Inventory; POE: predict observe explain.



The results of the normality test revealed a nonnormal distribution of the data (Table S3 in Multimedia Appendix 2). Consequently, the researchers conducted a paired-sample Wilcoxon rank sum test to compare the pre- and posttest findings of the intervention and control groups to assess any differences between the 2 groups. The results shown in Table 3 demonstrate notable fluctuations in both light transmission (z=−2.696; P=.008) and total POE test scores (z=−2.458; P=.01). Nevertheless, the results of the study indicate that there was no statistically significant advantage observed in animal vision (z=−0.847; P=.42) and color-light mixing POE test scores (z=−0.782; P=.46) as a result of the AR game intervention. It should be noted, however, that there was an improvement in scores following the intervention.

Table . Between-group differences between the intervention and control groups on each of predict-observe-explain (POE) test (pre- and posttests).

POE test score		Intervention group[a] (n=18), mean (SD)	Control group[b] (n=18), mean (SD)	Difference, mean (95% CI)	z score	P value[c]
Animal vision				0.36 (−0.71 to 1.43)	−.847	.42
	Pretest	1.83 (1.10)	1.94 (1.21)			
	Posttest	2.33 (1.14)	2.17 (0.99)			
Light transmission				0.97 (−0.37 to 2.31)	−2.696	.008
	Pretest	2.83 (1.72)	2.67 (2.2)			
	Posttest	4.44 (1.76)	3.00 (1.88)			
Color-light mixing				0.72 (−0.44 to 1.88)	−0.782	.46
	Pretest	1.50 (1.09)	1.94 (1.16)			
	Posttest	2.39 (1.24)	2.50 (1.04)			
Total				2.06 (−0.1 to 4.22)	−2.458	.01
	Pretest	6.17 (2.28)	6.56 (2.12)			
	Posttest	9.17 (2.48)	7.67 (1.71)			

[a]Augmented reality game.

[b]Non–augmented reality game.

[c]Mann-Whitney *U* test.

In this study, subjective IMI scale values acquired during the trial were statistically analyzed. It was observed that the different groups showed significant variability in levels of interest-enjoyment (z=−2.440; P=.02) and perceived competence (z=−2.170; P=.03; Table 4), whereas significant differences were not observed in levels of effort-importance (z=−1.310; P=.20) and tension-pressure (z=−0.733; P=.48).

Table . Comparison of Intrinsic Motivation Inventory (IMI) variability between the intervention and control groups.

IMI subscale	Intervention group[a] (n=18), mean (SD)	Control group[b] (n=18), mean (SD)	z score	P value[c]
Interest-enjoyment	20.28 (2.72)	18.00 (2.57)	−2.440	.02
Perceived competence	18.83 (3.05)	16.33 (3.34)	−2.170	.03
Effort-importance	12.89 (1.97)	11.83 (2.57)	−1.310	.20
Tension-pressure	10.44 (2.33)	9.77 (2.90)	−0.733	.48

[a]Augmented reality game.

[b]Non–augmented reality game.

[c]Mann-Whitney *U* test.

Discussion

Principal Findings

The integration of science education into the foundational education of children aims to systematically cultivate their abilities in inductive and deductive thinking [34]. Serious games have demonstrated efficacy in enhancing teaching and learning outcomes within the contemporary domain of children's science education [7]. AR technology has garnered growing interest in the realm of serious game design in recent times due to its ability to visually represent scientific processes that are not easily observable in real-life situations [35]. Further, the incorporation of AR technology into mobile devices has resulted in widespread adoption, facilitating the implementation of many apps [17]. Nevertheless, there is a lack of comprehensive study and experimentation to substantiate the efficacy of AR design in the realm of children's science education. Consequently, a series of AR science instructional games were developed, focusing on the comprehension of light principles. The objective was to assess the efficacy of the games and the degree of intrinsic motivation of the students. The results showed that children who participated in the AR science game had substantially higher POE test scores and conceptual understanding of light propagation than the control group.

The study revealed that children exhibited varying levels of comprehension in relation to light concepts across diverse themes. Reliable between-group differences were detected among the topics of light propagation. The rationale behind the use of AR lies in its inherent benefits, which include the ability for children to engage in physical activity while delving into a more comprehensive exploration as compared to 2D games. Additionally, AR technology facilitates the rendering of real-world light phenomena, as supported by previous studies [36,37]. Our game was developed with the purpose of creating a metaphorical representation of concepts connected to light orientation using gesture-touch interactions. It aims to alleviate the cognitive burden experienced by children when trying to connect abstract information about light with real-world light occurrences. This is achieved by incorporating 3D properties of lighting effects into the game. Furthermore, a notable increase in the game scores of the intervention group was noticed across all topics. This observation serves as evidence for the beneficial influence of AR games on children's conceptual transformation during the process of acquiring scientific knowledge. Our game also serves as a means of scientific investigation, necessitating active engagement from the children. Certain children had not before contemplated the underlying mechanisms responsible for commonplace visual occurrences. The stimulation of their drive to study and their high curiosity played a significant role in facilitating their conceptual shift and fostering the development of scientific thinking [38]. Nevertheless, the intervention group did not exhibit any noteworthy disparity compared to the control group in relation to the topics of animal vision and color-light mixing. The limited influence of the different interactive designs, specifically mixed camera mode and projected lighting mode, on children's cognition may account for this disparity when compared to the visual representation format in the 2D game.

Intrinsic motivation is a potent factor that influences performance, learning persistence, and productivity [39]. Children in the intervention group demonstrated greater interest and enjoyment in intrinsic motivation than those in the control group, and they demonstrated an ability to embrace and comprehend the causes of certain light phenomena.

AR imparts scientific information that challenges children's prior knowledge and stimulates their interest. Consequently, it can arouse interest in the principles and stimulate active thought [35]. During the test, we observed that participants had a keen interest in the game and avidly explored the interface's interactive elements. Moreover, there was a significant difference in perceptual ability between the intervention and control groups. We believe that this difference stems from the fact that AR games, created by adding 3D virtual objects to real-world images, can better facilitate children's understanding of complex concepts [15]. However, we also found that the intervention group showed some stress toward AR games. Children have a period of adjustment for things to which they are not accustomed to, as evidenced by their inattention and attempts to communicate with the observer when they encounter difficulties in the game [17]. Future research can therefore concentrate on how to provide prompts even when encountering difficulties.

The researchers observed that the children's engagement in gameplay facilitated their conscious observation of light occurrences in their daily lives, resulting in a modest improvement in their comprehension during the final phase of the tests. Furthermore, when the optical principles pertaining to linear propagation, reflection, and refraction became increasingly complex, it became more challenging for the children to comprehend, leading to confusion in certain preintervention participants regarding the distinctions between these concepts. It is important to acknowledge that when a child misinterprets the dynamic effects, animation, or creative expression of a game feature, the game can potentially facilitate the development of novel alternative understanding. Fortunately, the occurrence of this scenario was limited in the 2 assessment tests conducted during the formal experiment.

The strengths and weaknesses of our study in comparison with other studies is shown in Table S4 in Multimedia Appendix 2.

In summary, the integration of AR into educational games has the potential to enhance children's science education by offering a more immersive and engaging learning experience. This approach also may address the challenges associated with inadequate education and the lack of motivation among children to explore scientific subjects.

Conclusions and Limitations

The results suggest that the use of AR serious games can effectively motivate children to undergo conceptual shifts during the initial phases of science education. This, in turn, leads to an improved level of comprehension of scientific material. Furthermore, it is expected that these positive outcomes can be replicated in future preschool science education settings. This randomized controlled trial provides confirmation that the science education game we developed, using AR technology, has the potential to rectify children's misconceptions regarding scientific concepts and improve the overall efficacy of science teaching.

However, there are also some limitations. First, the sample size used in the study was limited, and the sample population was mainly from the more resource-rich region of Jiangsu Province, China. Consequently, it is challenging to ascertain the presence of regional variations in other geographical areas. Prospective studies with large samples are needed to further confirm the results, and the results can be improved by considering gender, family upbringing, and children's interest preferences in subsequent studies. Second, AR apps require a lot of attention and can be a distraction. It can cause students to ignore instructions or important stages of the experience. In addition, as the situation appeared in the pre-experiment, the game as a teaching tool may generate new misconceptions if the child misinterprets the content of the game. Finally, the existing game conveys scientific concepts mostly through voice-over prompts, which are insufficient to grab the children's attention, and children may be distracted and lose essential information during the voice-over prompts.

Editorial Notice

This randomized study was only retrospectively registered, as the authors had not considered it necessary to register prospectively. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears to be low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The designed augmented reality game.
[MP4 File, 18278 KB - games_v12i1e47807_app1.mp4]

Multimedia Appendix 2
Supplementary tables and figures.
[DOCX File, 7270 KB - games_v12i1e47807_app2.docx]

Checklist 1
CONSORT eHEALTH Checklist.
[PDF File, 1214 KB - games_v12i1e47807_app3.pdf]

References

1. Duit R, Treagust DF. Conceptual change: a powerful framework for improving science teaching and learning. Int J Sci Educ 2003 Jun;25(6):671-688. [doi: 10.1080/09500690305016]
2. Adey P, UNESCO International Bureau of Education. The science of thinking, and science for thinking: a description of cognitive acceleration through science education (CASE). UNESCO. 1999. URL: https://unesdoc.unesco.org/ark:/48223/pf0000124789 [accessed 2023-12-07]
3. Gopnik A. Scientific thinking in young children: theoretical advances empirical research, and policy implications. Science 2012 Sep 28;337(6102):1623-1627. [doi: 10.1126/science.1223416] [Medline: 23019643]
4. Magnusson S, Krajcik J, Borko H. Nature, sources, and development of pedagogical content knowledge for science teaching. In: Gess-Newsome J, Lederman NG, editors. Examining Pedagogical Content Knowledge: Kluwer Academic Publishers; 1999:95-132. [doi: 10.1007/0-306-47217-1_4]
5. Nazry NNM, Romano DM. Mood and learning in navigation-based serious games. Comput Hum Behav 2017 Aug;73:596-604. [doi: 10.1016/j.chb.2017.03.040]
6. Malaquias RF, Malaquias FFO, Hwang Y. Understanding technology acceptance features in learning through a serious game. Comput Hum Behav 2018 Oct;87:395-402. [doi: 10.1016/j.chb.2018.06.008]
7. Zhonggen Y. A meta-analysis of use of serious games in education over a decade. International Journal of Computer Games Technology 2019 Feb 3;2019. [doi: 10.1155/2019/4797032]
8. Lester JC, Spires HA, Nietfeld JL, Minogue J, Mott BW, Lobene EV. Designing game-based learning environments for elementary science education: a narrative-centered learning perspective. Inf Sci 2014 Apr;264:4-18. [doi: 10.1016/j.ins.2013.09.005]
9. Laine TH, Nygren E, Dirin A, Suk HJ. Science Spots AR: a platform for science learning games with augmented reality. Education Technology Research and Development 2016 Jun;64(3):507-531. [doi: 10.1007/s11423-015-9419-0]
10. Wang M, Zheng X. Using game-based learning to support learning science: a study with middle school students. Asia-Pacific Education Researcher 2021 Apr;30(2):167-176. [doi: 10.1007/s40299-020-00523-z]
11. Amin TG. Conceptual metaphor and the study of conceptual change: research synthesis and future directions. Int J Sci Educ 2015 Apr 1;37(5-6):966-991. [doi: 10.1080/09500693.2015.1025313]
12. Oranç C, Küntay AC. Learning from the real and the virtual worlds: educational use of augmented reality in early childhood. International Journal of Child-Computer Interaction 2019 Sep;21:104-111. [doi: 10.1016/j.ijcci.2019.06.002]
13. Arici F, Yildirim P, Caliklar Ş, Yilmaz RM. Research trends in the use of augmented reality in science education: content and bibliometric mapping analysis. Computers & Education 2019 Dec;142:103647. [doi: 10.1016/j.compedu.2019.103647]
14. Tarng W, Ou KL, Lu YC, Shih YS, Liou HH. A sun path observation system based on augment reality and mobile learning. Mobile Information Systems 2018 Mar 29;2018. [doi: 10.1155/2018/5950732]
15. Sahin D, Yilmaz RM. The effect of augmented reality technology on middle school students' achievements and attitudes towards science education. Computers & Education 2020 Jan;144:103710. [doi: 10.1016/j.compedu.2019.103710]
16. Fleer M. Affective imagination in science education: determining the emotional nature of scientific and technological learning of young children. Res Sci Educ 2013 Oct;43(5):2085-2106. [doi: 10.1007/s11165-012-9344-8]
17. Garzón J, Pavón J, Baldiris S. Systematic review and meta-analysis of augmented reality in educational settings. Virtual Reality 2019 Dec;23(4):447-459. [doi: 10.1007/s10055-019-00379-9]

18. di Serio Á, Ibáñez MB, Kloos CD. Impact of an augmented reality system on students' motivation for a visual art course. Computers & Education 2013 Oct;68:586-596. [doi: 10.1016/j.compedu.2012.03.002]
19. Nkadimeng M, Ankiewicz P. The affordances of Minecraft Education as a game-based learning tool for atomic structure in junior high school science education. J Sci Educ Technol 2022 Oct;31(5):605-620. [doi: 10.1007/s10956-022-09981-0]
20. Swacha J, Maskeliūnas R, Damaševičius R, et al. Introducing sustainable development topics into computer science education: design and evaluation of the Eco JSity game. Sustainability 2021 Apr 1;13(8):4244. [doi: 10.3390/su13084244]
21. Xiao Y, Jiang C. Conceptual change in preschool science education: evaluating a serious game designed with image schemas for teaching sound concept. In: Fang X, editor. HCI in Games. HCII 2020. Lecture Notes in Computer Science: Springer; 2020, Vol. 12211:503-520. [doi: 10.1007/978-3-030-50164-8_37]
22. Cai S, Liu C, Wang T, Liu E, Liang JC. Effects of learning physics using augmented reality on students' self ‐ efficacy and conceptions of learning. Brit J Educ Technol 2021 Jan;52(1):235-251. [doi: 10.1111/bjet.13020]
23. Lu SJ, Liu YC. Integrating augmented reality technology to enhance children's learning in marine education. Environ Educ Res 2015;21(4):525-541. [doi: 10.1080/13504622.2014.911247]
24. Wang Y, Ko E, Wang H. Augmented reality (AR) app use in the beauty product industry and consumer purchase intention. Asia Pacific Journal of Marketing and Logistics 2022 Jan 3;34(1):110-131. [doi: 10.1108/APJML-11-2019-0684]
25. Lu SJ, Liu YC, Chen PJ, Hsieh MR. Evaluation of AR embedded physical puzzle game on students' learning achievement and motivation on elementary natural science. Interactive Learning Environments 2020;28(4):451-463. [doi: 10.1080/10494820.2018.1541908]
26. van der Graaf J, Segers E, Verhoeven L. Discovering the laws of physics with a serious game in kindergarten. Computers & Education 2016 Oct;101:168-178. [doi: 10.1016/j.compedu.2016.06.006]
27. Jamonnak S, Cheng E. Little Botany: a mobile game utilizing data integration to enhance plant science education. International Journal of Computer Games Technology 2017 Mar 22;2017:3635061. [doi: 10.1155/2017/3635061]
28. Chen MHM, Tsai ST, Chang CC. Effects of game-based instruction on the results of primary school children taking a natural science course. Education Sciences 2019 Apr 12;9(2):79. [doi: 10.3390/educsci9020079]
29. Liew CW, Treagust DF. The effectiveness of predict-observe-explain tasks in diagnosing students' understanding of science and in identifying their levels of achievement. ERIC. 1998 Apr. URL: https://eric.ed.gov/?id=ED420715 [accessed 2023-12-07]
30. Bayrak BK. Using two-tier test to identify primary students' conceptual understanding and alternative conceptions in acid base. MIJE 2013 Aug 1;3(2):19-26 [FREE Full text] [doi: 10.13054/mije.13.21.3.2]
31. Li FY, Hwang GJ, Chen PY, Lin YJ. Effects of a concept mapping-based two-tier test strategy on students' digital game-based learning performances and behavioral patterns. Computers & Education 2021 Nov;173:104293. [doi: 10.1016/j.compedu.2021.104293]
32. Vos N, van der Meijden H, Denessen E. Effects of constructing versus playing an educational game on student motivation and deep learning strategy use. Computers & Education 2011 Jan;56(1):127-137. [doi: 10.1016/j.compedu.2010.08.013]
33. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
34. Kohlstedt SG. Teaching Children Science: Hands-On Nature Study in North America, 1890-1930: University of Chicago Press; 2010. [doi: 10.7208/chicago/9780226449920.001.0001]
35. Cheng KH, Tsai CC. The interaction of child–parent shared reading with an augmented reality (AR) picture book and parents' conceptions of AR learning. Br J Educ Technol 2016 Jan;47(1):203-222. [doi: 10.1111/bjet.12228]
36. Pellas N, Fotaris P, Kazanidis I, Wells D. Augmenting the learning experience in primary and secondary school education: a systematic review of recent trends in augmented reality game-based learning. Virtual Real 2019 Dec;23(4):329-346. [doi: 10.1007/s10055-018-0347-2]
37. Thomas BH. A survey of visual, mixed, and augmented reality gaming. Computers in Entertainment 2012 Oct;10(1):1-33. [doi: 10.1145/2381876.2381879]
38. Lan R, Sun X, Liu B. The use of augmented reality to motivate scientific learning for preschool children by fostering curiosity. In: Zaphiris P, Ioannou A, editors. Learning and Collaboration Technologies. Novel Technological Environments. HCII 2022. Lecture Notes in Computer Science, vol 13329: Springer; 2022:71-80. [doi: 10.1007/978-3-031-05675-8_7]
39. López-Martínez A, Meroño L, Cánovas-López M, García-de-Alcaraz A, Martínez-Aranda LM. Using gamified strategies in higher education: relationship between intrinsic motivation and contextual variables. Sustainability 2022 Sep 3;14(17):11014. [doi: 10.3390/su141711014]

Abbreviations

AR: augmented reality
IMI: Intrinsic Motivation Inventory
POE: predict observe explain

Edited by G Eysenbach, T Leung; submitted 02.04.23; peer-reviewed by C Beaudoin, K Kanno, S Hoermann; revised version received 09.10.23; accepted 06.11.23; published 01.02.24.

Please cite as:
Liu B, Wan X, Li X, Zhu D, Liu Z
An Augmented Reality Serious Game for Children's Optical Science Education: Randomized Controlled Trial
JMIR Serious Games 2024;12:e47807
URL: https://games.jmir.org/2024/1/e47807
doi:10.2196/47807

Original Paper

Effects of a Serious Smartphone Game on Nursing Students' Theoretical Knowledge and Practical Skills in Adult Basic Life Support: Randomized Wait List–Controlled Trial

Nino Fijačko[1,2*], MSc; Ruth Masterson Creber[3*], PhD; Špela Metličar[1,4*], RN; Matej Strnad[2,5,6*], MD, PhD; Robert Greif[7,8,9*], MME, MD; Gregor Štiglic[1,10*], PhD; Pavel Skok[2*], MD, PhD

[1]Faculty of Health Sciences, University of Maribor, Maribor, Slovenia
[2]Maribor University Medical Centre, Maribor, Slovenia
[3]School of Nursing, Columbia University, New York, NY, United States
[4]Medical Dispatch Centre Maribor, University Clinical Centre Ljubljana, Ljubljana, Slovenia
[5]Faculty of Medicine, University of Maribor, Maribor, Slovenia
[6]Community Healthcare Center Dr Adolfa Drolca Maribor, Maribor, Slovenia
[7]European Resuscitation Council Research Net, Niels, Belgium
[8]University of Bern, Bern, Switzerland
[9]School of Medicine, Sigmund Freud University Vienna, Vienna, Austria
[10]Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia
[*]all authors contributed equally

Corresponding Author:
Nino Fijačko, MSc
Faculty of Health Sciences
University of Maribor
Žitna 15
Maribor, 2000
Slovenia
Phone: 386 23004764
Email: nino.fijacko@um.si

Abstract

Background: Retention of adult basic life support (BLS) knowledge and skills after professional training declines over time. To combat this, the European Resuscitation Council and the American Heart Association recommend shorter, more frequent BLS sessions. Emphasizing technology-enhanced learning, such as mobile learning, aims to increase out-of-hospital cardiac arrest (OHCA) survival and is becoming more integral in nursing education.

Objective: The aim of this study was to investigate whether playing a serious smartphone game called MOBICPR at home can improve and retain nursing students' theoretical knowledge of and practical skills in adult BLS.

Methods: This study used a randomized wait list–controlled design. Nursing students were randomly assigned in a 1:1 ratio to either a MOBICPR intervention group (MOBICPR-IG) or a wait-list control group (WL-CG), where the latter received the MOBICPR game 2 weeks after the MOBICPR-IG. The aim of the MOBICPR game is to engage participants in using smartphone gestures (eg, tapping) and actions (eg, talking) to perform evidence-based adult BLS on a virtual patient with OHCA. The participants' theoretical knowledge of adult BLS was assessed using a questionnaire, while their practical skills were evaluated on cardiopulmonary resuscitation quality parameters using a manikin and a checklist.

Results: In total, 43 nursing students participated in the study, 22 (51%) in MOBICPR-IG and 21 (49%) in WL-CG. There were differences between the MOBICPR-IG and the WL-CG in theoretical knowledge (*P*=.04) but not in practical skills (*P*=.45) after MOBICPR game playing at home. No difference was noted in the retention of participants' theoretical knowledge and practical skills of adult BLS after a 2-week break from playing the MOBICPR game (*P*=.13). Key observations included challenges in response checks with a face-down manikin and a general neglect of safety protocols when using an automated external defibrillator.

Conclusions: Playing the MOBICPR game at home has the greatest impact on improving the theoretical knowledge of adult BLS in nursing students but not their practical skills. Our findings underscore the importance of integrating diverse scenarios into adult BLS training.

Trial Registration: ClinicalTrials.gov (NCT05784675); https://clinicaltrials.gov/study/NCT05784675

(JMIR Serious Games 2024;12:e56037) doi:10.2196/56037

KEYWORDS

serious smartphone game; adult basic life support; teaching; game; games; gaming; education; nurse; nursing; nurses; educational; mHealth; mobile health; app; apps; application; applications; smartphone; smartphones; RCT; randomized controlled trial; controlled trial; technology-enhanced learning; TEL; life support; knowledge retention; theoretical knowledge; practice; practical; resuscitation

Introduction

Sudden cardiac arrest is one of the leading causes of death in adults worldwide. It is responsible for over a million deaths annually [1]. Most deaths occur in the out-of-hospital setting, and the outcome possibly can be improved with proper adult basic life support (BLS) [2]. Effective implementation of adult BLS can double the chances of survival after a sudden cardiac arrest [3,4]. Reviews report poor cardiopulmonary resuscitation (CPR) by nursing students, despite the completion of adult BLS certification [5]. BLS knowledge and skills decline significantly within months of initial training [5,6]. For this reason, the European Resuscitation Council (ERC) and American Heart Association guidelines recommend shorter and more frequent adult BLS training as it helps retain adult BLS content longer and maintain competency levels [7,8]. Currently, adult BLS education in higher nursing education institutions traditionally imparts theoretical knowledge through a frontal approach and teaches practical skills using manikins and automated external defibrillators (AEDs), although the approach can vary significantly from one university to another [5,9].

A noticeable generational shift is evident in health care systems, both in Europe and abroad, characterized by the increasingly common employment of younger individuals. These younger future health care employees bring a higher proficiency in technology and information literacy [10,11], attributes cultivated from growing up in an era dominated by modern technology [12]. Technology-enhanced learning (TEL) approaches, developed to improve adult BLS knowledge and skill retention, ultimately aim to increase out-of-hospital cardiac arrest (OHCA) survival [8]. The most recent adult BLS guidelines highlight the integration of TEL into adult BLS courses [8,13,14]. This includes not only immersive technologies, such as extended reality [15], but also mobile learning (m-learning), which has increased dramatically in nursing education in recent years [16]. A recent meta-analysis indicates that serious smartphone games are a promising and effective tool for adult BLS education [17].

M-learning, by its definition, encompasses the use of mobile technology [18], with mobile apps on smartphones often serving as the educational platform [19]. Research has demonstrated m-learning's beneficial effects on fostering a variety of learning outcomes and competencies in the field of nursing [20,21]. Smartphone-based m-learning [21] seamlessly complements education through serious games and gamification [15]. Gamification involves applying game design elements to nongame contexts [22], such as educational content in higher education [23]. Conversely, serious games are crafted to use a specific type of game (eg, computer or mobile games) for the purpose of learning about significant subjects, such as adult BLS content education at the higher education level [24].

To the best of our knowledge, only a limited number of studies have explored the use of serious smartphone games for teaching adult BLS to health care students [25-29]. Among these, only 1 study demonstrated an improvement in both the theoretical knowledge and practical skills associated with adult BLS [28]. Other studies have reported enhancements in either theoretical knowledge [29] or practical skills related to adult BLS. The positive effects of a serious smartphone game can be seen as early as 2 weeks [25,26], as well as 1 month after the intervention [27-29]. Studies have compared different teaching methods, where the use of serious smartphone games seems to have better results than simulation-based learning but is less effective than virtual reality–based game learning [26,30]. Some studies have also shown improvements in practical skills, such as compression rate accuracy [27,28], although this tends to be inferior when compared to simulation-based methods [30]. In contrast, in 2 studies, serious smartphone games did not provide notable benefits and led to worse performance in theoretical and practical areas, although students showed a clear preference in favor of serious smartphone games [27,28].

The aim of the study was to evaluate whether playing a serious smartphone game called MOBICPR [31] at home can enhance nursing students' theoretical knowledge of and practical skills in adult BLS.

Methods

Study Protocol

The study was conducted at the Faculty of Health Sciences, University of Maribor (Maribor, Slovenia) between March and May 2023. The study was registered in ClinicalTrials.gov (NCT05784675). The study protocol was written in accordance with the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth (Multimedia Appendix 1) [32].

Ethical Considerations

Ethical approval was obtained from the Slovenian National Medical Ethics Committee (0120-157/2018), and permission to conduct the study on the faculty premises was obtained from the Faculty of Health Sciences, University of Maribor. During an oral presentation of the study, nursing students were informed about the research protocol, and written consent was obtained afterward. Data confidentiality and anonymity were maintained throughout the study. Participants were rewarded for their participation in the study with a free beverage from a vending machine and a copy of the *Game Changer* painting by street artist Banksy [33].

Participants

All nursing students enrolled in the first-degree nursing program at the Faculty of Health Sciences, University of Maribor, during the 2022-2023 academic year were invited to participate in the study. Inclusion criteria to participate in the study were written informed consent, an age of at least 18 years, and ability to perform adult BLS on a manikin (eg, without injury). Our study had no exclusion criteria.

Study Design and Randomization

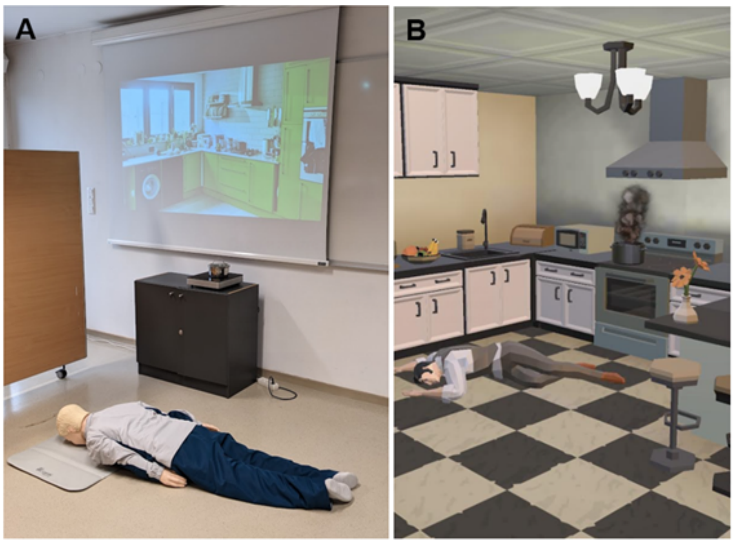
This study had a randomized wait list–controlled design, where nursing students were randomly assigned in a 1:1 ratio using a computer-generated list (Microsoft 365 Excel Enterprise) to either a MOBICPR intervention group (MOBICPR-IG) or a wait-list control group (WL-CG). The WL-CG was a group of nursing students who were assigned to a wait list and received the intervention (MOBICPR game for playing at home) 2 weeks after the MOBICPR-IG.

Interventions

All assessments of the participants' theoretical knowledge of and practical skills in adult BLS were conducted 3 time points: baseline, 2-week follow-up, and 4-week follow-up. At the baseline assessment, the investigators first collected demographic data from the participants. Additionally, the participants were questioned about their willingness to assist both family members and strangers during OHCA with CPR. Prior to practical skills in adult BLS, participants' the theoretical knowledge of adult BLS was assessed using a questionnaire with 33 single- and multiple-choice questions [25,34-36] on an open source online survey app called 1ka (Ljubljana, Slovenia); see Multimedia Appendix 2. A back-translation approach was used for translating the questionnaire into the Slovenian language.

Prior to the assessment of adult BLS practical skills, each participant was given a scenario based on OHCA to read (Multimedia Appendix 3). After reading the scenario, the participants were given a smartphone for calling emergency services at the time of performing adult BLS. Instead of dialing the actual emergency number, the participants used the contact stored on the smartphone as 112 (ie, the Slovenian emergency number). After the call was placed by each participant, the investigator answered the phone and conducted a simulated dispatcher conversation [37]. The investigator was a registered nurse working at the local medical dispatch center. Each participant performed 2-minute adult BLS without any help in a staged kitchen on a manikin (Resusci Anne Quality Cardiopulmonary Resuscitation [QCPR], Laerdal Medical) using an AED (Defibtech, Trainer AED). The staged kitchen was a space surrounded by mobile walls in the hospital's simulated room. A photo of a kitchen was projected onto the wall, and below it was an electric stove with a pot full of water (Figure 1). In each adult BLS scenario, the investigators turned on the electric stove, and the scenario began when the water started to boil, simulating a hazard. The kitchen was chosen because the majority of OHCAs occur there [38].

Figure 1. Staged kitchen with the Resusci Anne QCPR in the middle (A) and a cartoon person in the MOBICPR game lying on the floor in a kitchen (B). QCPR: Quality Cardiopulmonary Resuscitation.



After 2 minutes of performing adult BLS, each participant received assistance from an outside person bringing in an AED and taking over CPR. The adult BLS practical skills of each participant were recorded using a Sony Handycam 4K AX53 camera and an Apple iPad Pro 3rd generation tablet. Two investigators with a background in emergency medicine and teaching laypersons adult BLS assessed the participants' practical skills in adult BLS using a modified checklist [25,36,39] according to the ERC BLS guidelines of 2015 [40] and 2021 [34], with a total of 34 items (Multimedia Appendix 4). A back-translation approach was used for translating the checklist into the Slovenian language. Numerical data from the

SkillReporter for Tablet version 1.4.1 (Laerdal Medical) app installed on a Samsung Galaxy Tab S6 Lite tablet was also included in the evaluation of the participants' practical skills in adult BLS. Investigator debriefing was not conducted following the assessment of the participants' adult BLS theoretical knowledge and practical skills. Instead, each participant (from MOBICPR-IG at baseline and from WL-CG 2 weeks after baseline) first played the MOBICPR game [31] on a Samsung Galaxy A13 smartphone in front of the investigator and then received the same smartphone to play at home. The objective of the MOBICPR game is for participants to interact with a smartphone using gestures (eg, tapping) and actions (eg, talking) to help save the life of a virtual patient with OHCA by performing evidence-based adult BLS. The MOBICPR game is based on the 2021 ERC BLS guidelines [34], and the BLS content was developed using the Delphi process. The patient's chance of survival in the MOBICPR game reduced with each incorrect interaction by the participants. At the end of the MOBICPR game, each participant received a total score in the form of a gamification feature that corresponded to the risk of survival (score>50% meant the patient survived) [41]. Gamification, defined as "using game design elements in non-game contexts," has been introduced into nursing education to promote engagement using features such as leaderboards, rewards, badges, and avatars [22]. After playing the MOBICPR game as much as they wanted for 2 weeks, participants in the MOBICPR-IG returned the smartphones. Participants in the W-CG then received the smartphones and followed the same protocol as participants in the MOBICPR-IG, that is, they played the MOBICPR game in front of the investigator before taking the smartphone home. Participants in the W-CG also returned the smartphones after playing the MOBICPR game at home for 2 weeks. Additionally, at the study's conclusion, each participant was asked an open-ended question regarding the number of family members or friends with whom they shared the MOBICPR game for playing.

Outcome Measures

The primary outcomes were (1) assessment of the participants' theoretical knowledge of adult BLS using a questionnaire with a total maximum score of 33 points, where each correct answer was awarded 1 point (Multimedia Appendix 2), and (2) assessment of the participants' practical skills in adult BLS using a checklist with a total maximum score of 39 points (Multimedia Appendix 4).

The secondary outcome was a summary score of high-quality CPR components: (1) a chest compression (CC) rate of 100-120 beats per minute (bpm), (2) a CC depth of 50-60 mm, (3) CC fraction>80%, and (4) a rescue breath volume of 500-600 mL (Multimedia Appendix 4). All measures were taken as mentioned earlier [16,23,27,28]. A total QCPR score was also included, ranging from 0% to 100%. More detailed information about software scoring is available on the Laerdal Medical website [42]. Both primary and secondary outcomes were measured at 3 time points: baseline, 2-week follow-up, and 4-week follow-up.

Statistical Analysis

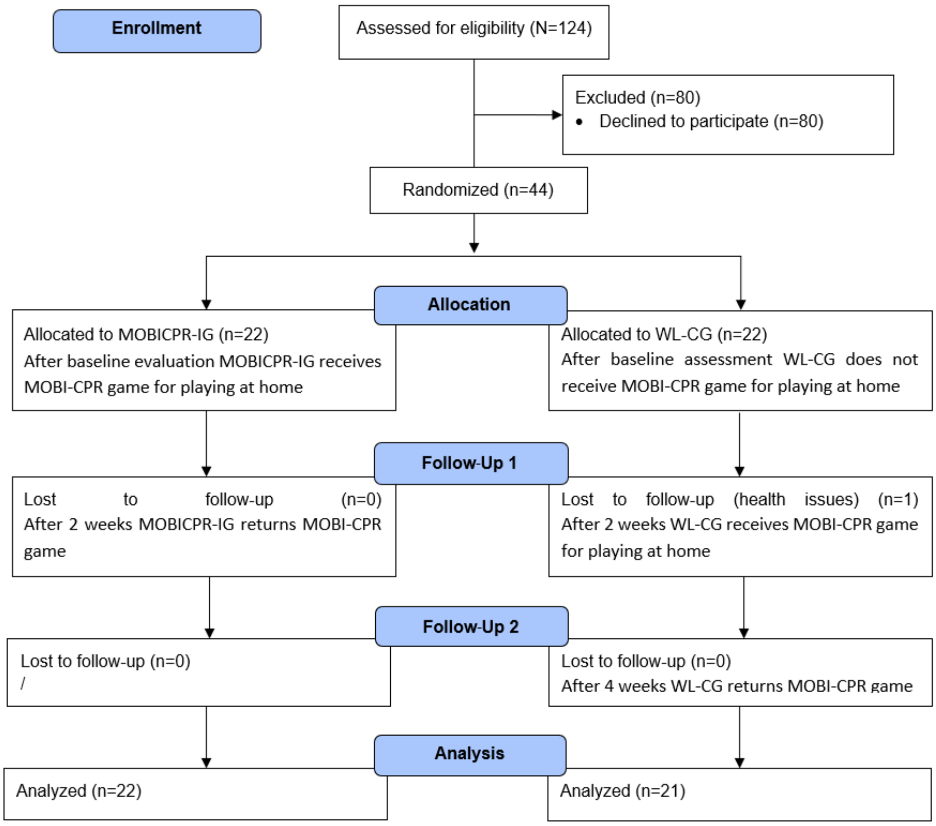
Statistical analyses were conducted in October and November 2023. Data were analyzed using the R statistical programming language (R Foundation for Statistical Computing). The data presented in the summary table were prepared using frequency analysis, which also included a chi-square test to assess the similarity of the distribution between the intervention and control groups. Theoretical knowledge and practical skill assessments were averaged at the item level and subsequently analyzed using nonparametric statistical tests (Wilcoxon paired-sample test and Mann-Whitney *U* test) as the normality of the distribution was violated. As nonnormal distribution might represent a problem when calculating mean values, violin plots were also used for the purpose of visualizing aggregated scores due to their ability to visualize the distribution of the data. $P<.05$ was considered statistically significant. Effect size (η^2) values >0.1 represented a small effect; 0.3, a moderate effect; and ≥0.5, a large effect. Continuous variables were analyzed according to the Gaussian distribution and reported as the mean (SD) or the median (IQR), whichever was appropriate.

Results

Participant Details

Of 124 nursing students, 80 (64.5%) declined to participate in the study and 44 (35.5%) were enrolled into the study. At follow-up, 1 (5%) of the 22 participants in the WL-CG dropped out. In the end, 43 (98%) of 44 participants were included in the final analysis (Figure 2).

Figure 2. Flow diagram of study participants. MOBICPR-IG: MOBICPR intervention group; WL-CG: wait-list control group.



The mean age of the participants was 19 (SD 0.6) years, 38 (88%) were female, 35 (81%) had a background in health care and nursing education, 32 (74%) had an iOS smartphone, and the self-reported mean daily time spent on the smartphone was 3.8 (SD 1.2) hours (Table 1).

Table 1. Baseline demographic characteristics.[a]

Characteristics	Total participants (N=43), n (%)	MOBICPR-IG[b] (n=22), n (%)	WL-CG[c] (n=21), n (%)	*P* value
Gender				.67
Male	5 (12)	3 (14)	2 (10)	—[d]
Female	38 (88)	19 (86)	19 (90)	—
Age (years)				.90
19	30 (70)	16 (73)	14 (67)	—
20	11 (26)	5 (23)	6 (29)	—
21	2 (5)	1 (5)	1 (5)	—
Education				.19
Health care and nursing	35 (81)	18 (82)	17 (81)	—
High school/gymnasium	5 (12)	4 (18)	1 (5)	—
Pharmacy	2 (5)	0	2 (10)	—
Economy	1 (2)	0	1 (5)	—
Operating system on smartphone				.001
Apple iOS	32 (74)	11 (50)	21 (100)	—
Android	11 (26)	11 (50)	0	—

[a]The percentages may exceed 100 because of rounding.

[b]MOBICPR-IG: MOBICPR intervention group.

[c]WL-CG: wait-list control group.

[d]Not applicable.

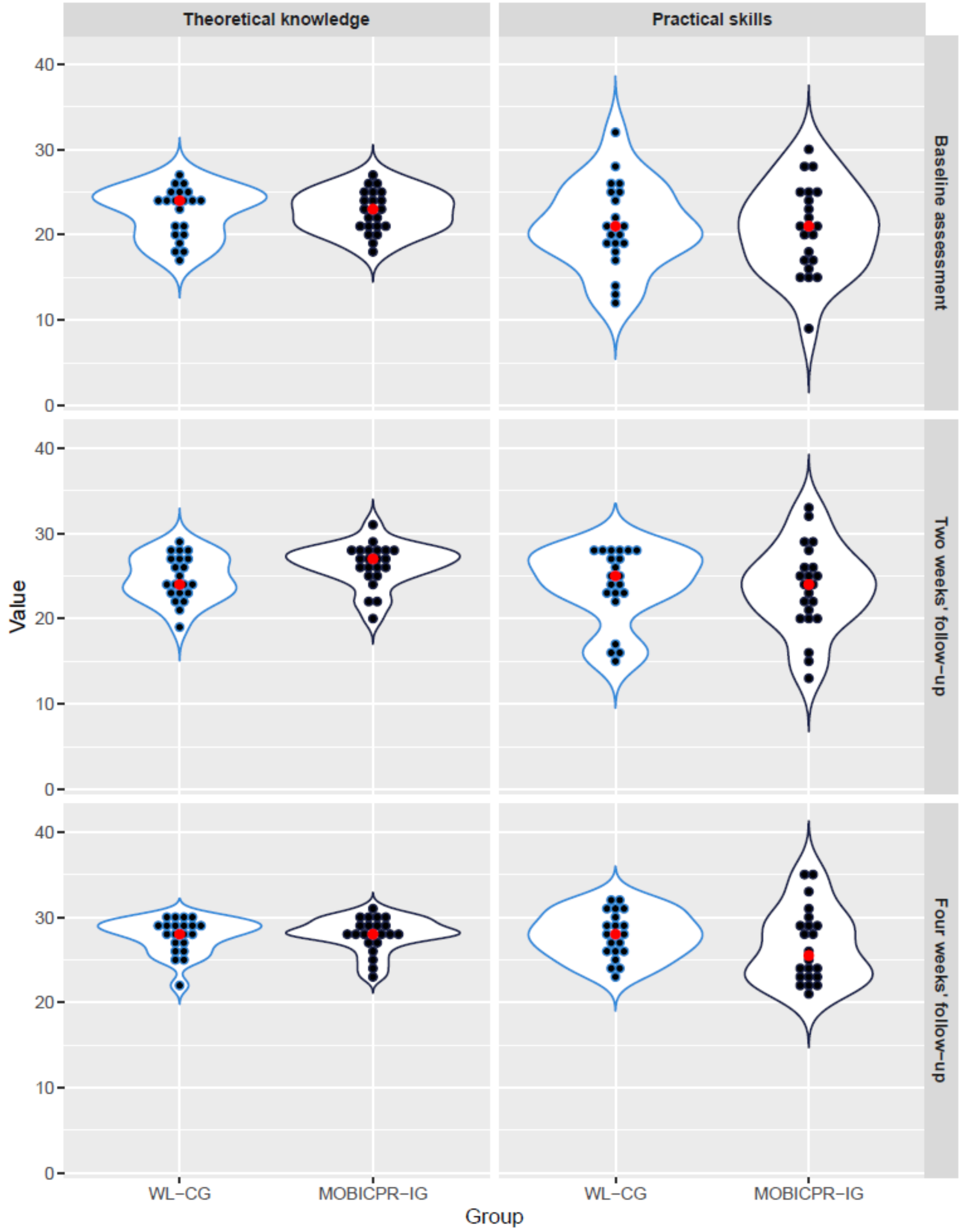
All participants had received some previous adult BLS training. However, only 2 (5%) had witnessed a cardiac arrest. Most of them (n=38, 88%) had already performed CCs on manikins, but only a few had also been giving rescue breaths (n=12, 28%) and used any kind of AED (n=13, 30%). All participants (n=43, 100%) expressed a willingness to assist a patient with OCHA and perform adult BLS. In addition, they all expressed a willingness to perform mouth-to-mouth resuscitation on a family member or acquaintance. However, only about half of them (n=19, 44%) were willing to do the same for a stranger. The predominant concern for not administering rescue breaths to unknown individuals was the uncertainty regarding the patient's medical history and the risk for infectious diseases, as cited by 22 (92%) of the 24 (56%) participants who expressed reluctance. On average, each participant introduced and shared the MOBICPR game with 3 (SD 2) family members or friends for trial and play.

Primary Outcomes

To assess the differences between the 2 groups at all 3 observed time points, we calculated the cumulative scores of adult BLS theoretical knowledge and practical skills for both groups.

Figure 3 shows that playing the MOBICPR game at home for 2 weeks improved the overall adult BLS theoretical knowledge (median gain of 4 points, IQR 3, η^2=0.113, *P*=.005) and practical skills (median gain of 4 points, IQR 7, η^2=0.05, *P*=.04). However, in the WL-CG, which waited for 2 weeks to play the MOBICPR game at home, the theoretical knowledge of adult BLS improved by 2 points (IQR 4, η^2=0.302, *P*=.001), whereas the practical skills in adult BLS increased by 3 points (IQR 3, η^2=0.018, *P*=.14). In the MOBICPR-IG, after 2 weeks of not playing the MOBICPR game at home, the retention of theoretical knowledge gained an additional 2 points (IQR 2, η^2=0.019, *P*=.13) and practical skills gained 3 points (IQR 3.75, η^2=0.122, *P*=.003) compared to the 2-week follow-up.

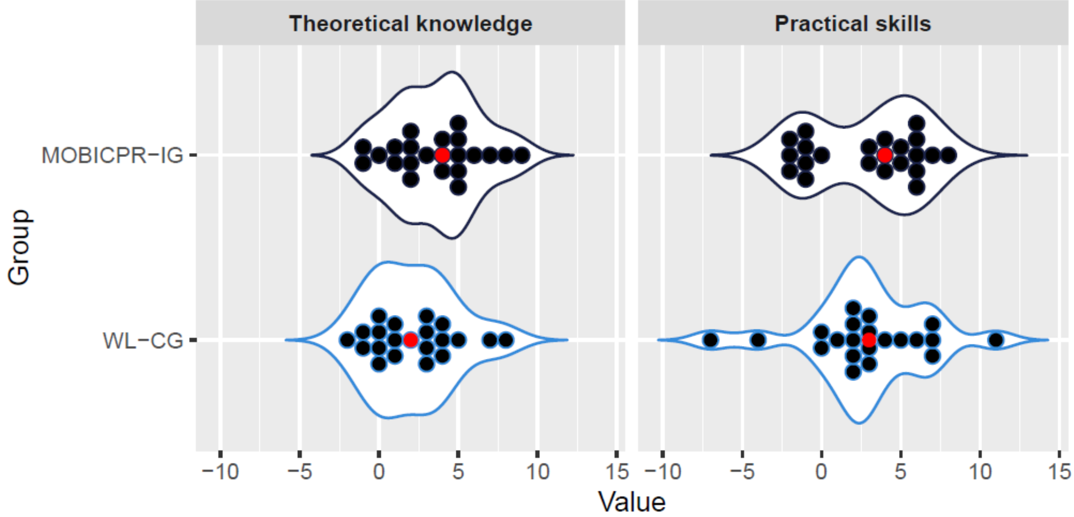
Figure 3. Comparison of theoretical knowledge and practical skill scores for all study participants at baseline, 2-week follow-up, and 4-week follow-up measurements. MOBICPR-IG: MOBICPR intervention group; WL-CG: wait-list control group.



To focus on the impact of playing the MOBICPR game on adult BLS theoretical knowledge and practical skills, we observed participants in both groups and calculated the difference in the cumulative points for both groups after they played the MOBICPR game at home for 2 weeks.

As demonstrated in Figure 4, in the WL-CG, only 3 (14%) participants improved their theoretical knowledge by ≥5 points and only 6 (29%) study participants who achieved this kind of improvement in the adult BLS practical skill score. In contrast, in the MOBICPR-IG, 9 (41%) participants improved their score by at least 5 points in both adult BLS theoretical knowledge and practical skills. The difference in improvement between the MOBICPR-IG and the WL-CG was not significant in practical skills (η^2=0.021, *P*=.45), while in theoretical knowledge, we observed a statistically significant difference (η^2=0.268, *P*=.04).

Figure 4. Difference in theoretical knowledge and practical skills score before and after playing the MOBICPR game. MOBICPR-IG: MOBICPR intervention group; WL-CG: wait-list control group.



To obtain more detailed insight into the improvements due to playing the MOBICPR game, we observed the differences in item-level scores before and after playing. Table 2 compares the participants' scores on questions used to test their theoretical knowledge. It is evident that there were notable differences in most items following engagement with the MOBICPR game. Of 33 scores, 13 (39%) decreased during MOBICPR game playing. For example, the score on question 3 (What is the second thing we check in a patient with cardiac arrest?) improved notably after MOBICPR game playing at home (*P*=.001). In contrast, the score on question 16 (You are alone. Will you go for the AED if it is 100 m away?) did not improve after MOBICPR game playing at home (*P*=.103).

Table 2. Question-level comparison of the mean scores for adult BLS[a] theoretical knowledge evaluation for MOBICPR-IG[b] and WL-CG[c] before and after MOBICPR game playing at home for 2 weeks (N=43).

Questions for evaluation of adult BLS theoretical knowledge	Score before playing the MOBICPR game at home, mean (SD)	Score after playing the MOBICPR game at home, mean (SD)	Difference (after – before)	*P* value
1. What is the first thing we check when we approach the patient?	2.51 (0.86)	1.72 (0.43)	–0.79	.04
2. On what kind of surface do we perform adult BLS?	2.88 (0.45)	2.09 (0.34)	–0.79	.001
3. What is the second thing we check in a patient with cardiac arrest?	2.14 (0.74)	2.93 (0.58)	0.79	.001
4. How many seconds do we need to assess consciousness?	1.74 (0.54)	1.95 (0.43)	0.21	.05
5. Before we assess breathing or perform CC[d], do we remove the patient's clothes?	1.26 (0.44)	1.77 (0.21)	0.51	.001
6. How do we open the airway?	2.95 (0.21)	1.05 (0.15)	–1.91	.001
7. What maneuver do we use to open the airway?	1.02 (0.15)	2.98 (0)	1.95	.001
8. How do we assess breathing?	2.95 (0.3)	1.00 (0)	–1.95	.001
9. How many seconds do we need to assess breathing?	1.51 (0.86)	3.00 (0)	1.49	.001
10. What is the most common breathing in a patient with cardiac arrest?	2.47 (0.83)	1.00 (0.63)	–1.47	.001
11. Who are you calling on the 112 number?	3.16 (0.37)	2.12 (0.32)	–1.05	.001
12. Will calling 911 in Slovenia or Europe reach emergency services?	1.72 (0.45)	3.12 (0.5)	1.40	.001
13. Who dials 112 in the case of cardiac arrest?	2.81 (0.55)	1.58 (0.3)	–1.23	.001
14. What do we need to tell the emergency medical dispatcher?	2.51 (0.51)	2.95 (0.51)	0.44	.001
15. What do you do with the phone after providing all the data?	1.91 (0.29)	2.70 (0.15)	0.79	.001
16. You are alone. Will you go for the AED[e] if it is 100 m away?	1.88 (0.32)	1.98 (0.35)	0.09	.10
17. You have help. Will you send it for the AED if it is 2 minutes away?	1.02 (0.15)	1.86 (0.15)	0.84	.001
18. Is this the sign for an AED?	2.00 (0)	1.02 (0)	–0.98	.001
19. Which picture shows the correct hand grip for CPR[f]?	2.00 (0)	2.00 (0.15)	0	.99
20. What is the right depth for CCs?	2.44 (0.77)	2.02 (0.35)	–0.42	.001
21. What is the correct body position for CCs?	1.98 (0.15)	2.86 (0)	0.88	.001
22. Where is the right place for CCs?	1.19 (0.59)	2.00 (0.78)	0.81	.001
23. What is the right frequency for CCs?	2.65 (1.41)	1.23 (0.48)	–1.42	.001
24. What is the CC-to-breath ratio for an adult?	1.98 (0.15)	1.09 (0.46)	–0.88	.001
25. How long can you interrupt CCs for rescue breaths?	1.30 (0.6)	1.93 (0.51)	0.63	.001
26. What is the volume of a rescue breath?	1.88 (0.59)	1.30 (0.26)	–0.58	.02
27. What do you do first if you have an AED?	2.58 (0.91)	2.93 (0.82)	0.35	.002
28. What do we do during AED rhythm analysis?	2.19 (0.59)	2.42 (0.46)	0.23	.17
29. What do we do during AED defibrillation?	1.00 (0)	2.07 (0)	1.07	.001

Questions for evaluation of adult BLS theoretical knowledge	Score before playing the MOBICPR game at home, mean (SD)	Score after playing the MOBICPR game at home, mean (SD)	Difference (after – before)	*P* value
30. Which of the following statements about the use of AEDs is false?	3.02 (1.3)	1.00 (1.05)	–2.02	.001
31. What do we do after the AED delivers an electric shock?	2.63 (0.69)	3.47 (0.65)	0.84	.001
32. When do we stop CPR?	1.74 (0.44)	2.91 (0.39)	1.16	.001
33. When is it recommended to replace someone during CPR?	1.72 (2.20)	1.81 (0.52)	0.09	.80
Cumulative score	2.08 (0.44)	2.06 (0.26)	0.03	.89

[a]BLS: basic life support.

[b]MOBICPR-IG: MOBICPR intervention group.

[c]WL-CG: wait-list control group.

[d]CC: chest compression.

[e]AED: automated external defibrillator.

[f]CPR: cardiopulmonary resuscitation.

Similarly, in the item-level score differences for practical skills, in 7 (21%) of 34 items, a significant increase was calculated (Table 3). For example, the score on item 1 (Approaches the patient safely) improved after MOBICPR game playing at home (*P*=.001). In contrast, the score on item 2 (Checks responsiveness: shouts and shakes the patient) did not improve after MOBICPR game playing at home (*P*=.81).

Table 3. Item-level comparison of the mean scores for adult BLS[a] practical skill evaluation for MOBICPR-IG[b] and WL-CG[c] before and after MOBICPR game playing at home for 2 weeks (N=43).

Items for evaluation of adult BLS practical skills	Score before playing the MOBICPR game at home, mean (SD)	Score after playing the MOBICPR game at home, mean (SD)	Difference (after – before)	*P* value
1. Approaches the patient safely	0.09 (0.29)	0.95 (0.46)	0.86	.001
2. Checks responsiveness: shouts and shakes the patient	0.67 (0.47)	0.70 (0.32)	0.02	.81
3. Opens the airway: head tilt–chin lift	0.40 (0.49)	0.88 (0.46)	0.49	.001
4. Performs look, listen, feel	0.74 (0.44)	0.70 (0.15)	–0.05	.62
5. Looks, listens, feels: time	0.53 (0.74)	0.98 (0.78)	0.44	.001
6. Calls 112 in the first minute	0.65 (0.48)	1.35 (0.26)	0.70	.001
7. Calls 112 at the right time	0.65 (0.48)	0.93 (0.29)	0.28	.003
8. Turns on the phone speaker and immediately starts CPR[d]	0.33 (0.47)	0.91 (0.49)	0.58	.001
9. Provides correct information to the dispatcher	0.33 (0.47)	0.37 (0.5)	0.05	.62
10. Provides information about the location	0.21 (0.41)	0.42 (0.5)	0.21	.05
11. Time to the first CCs[e]	0.51 (0.51)	0.56 (0.5)	0.05	.68
12. Corrects the body position for CCs	0.72 (0.45)	0.53 (0.32)	–0.19	.07
13. Corrects the CC location	0.88 (0.32)	0.88 (0.35)	0.00	.99
14. Corrects hand CCs	0.72 (0.45)	0.86 (0.41)	0.14	.11
15. Corrects the CC depth	1.49 (0.8)	0.79 (0.9)	–0.70	.001
16. Recoil of the chest	0.79 (0.41)	1.26 (0.21)	0.47	.002
17. Corrects the CC rate	1.42 (0.7)	0.95 (0.76)	–0.47	.001
18. Ratios CCs	0.91 (0.68)	1.37 (0.53)	0.47	.004
19. CC fraction	0.84 (0.43)	0.95 (0.38)	0.12	.17
20. Opens the airway: head tilt–chin lift	0.44 (0.5)	0.95 (0.45)	0.51	.001
21. Closes the nose and fits lips around the patient’s mouth	0.63 (0.49)	0.72 (0.37)	0.09	.32
22. Average pause of ventilation	0.70 (0.46)	0.84 (0.43)	0.14	.08
23. Opens the nose	0.02 (0.15)	0.77 (0.32)	0.74	.001
24. Looks for the chest to rise between 2 rescue breaths	0.37 (0.49)	0.12 (0.51)	–0.26	.003
25. Two rescue breaths	0.84 (0.37)	0.51 (0)	–0.33	.001
26. Volume of rescue breaths	0.53 (0.7)	1.00 (0.65)	0.47	.001
27. Switches on the AED[f] first at the right time	0.49 (0.51)	0.65 (0.46)	0.16	.20
28. Removes clothing	0.98 (0.15)	0.70 (0)	–0.28	.001
29. Position of the right AED pad	0.28 (0.45)	1.00 (0.45)	0.72	.001
30. Position of the left AED pad	0.47 (0.5)	0.28 (0.5)	–0.19	.103
31. Ensures nobody is touching the patient: analyzing	0.51 (0.51)	0.44 (0.5)	–0.07	.58
32. Ensures nobody is touching the patient: shock	0.09 (0.29)	0.58 (0.26)	0.49	.001
33. Presses the shock button at the right time	0.67 (0.47)	0.07 (0.44)	–0.60	.001
34. Immediately restarts CCs	0.95 (0.21)	0.74 (0.29)	–0.21	.002
Cumulative score	0.613 (0.14)	0.76 (0.17)	0.14	.04

[a]BLS: basic life support.
[b]MOBICPR-IG: MOBICPR intervention group.
[c]WL-CG: wait-list control group.
[d]CPR: cardiopulmonary resuscitation.
[e]CC: chest compression.
[f]AED: automated external defibrillator.

Secondary Outcomes

Table 4 shows a comparison of the high-quality CPR components between participants before and after MOBICPR game playing at home for 2 weeks. There were notable differences in the median (IQR) of the total QCPR score for MOBICPR game playing at home for 2 weeks for the MOBICPR-IG (before: median 41 (IQR 54); after: median 70 (IQR 41); *P*=.011). There was no difference for the MOBICPR-IG after not playing the MOBICPR game at home for 2 weeks.

Table 4. Results for high-quality CPR[a] components for the MOBICPR-IG[b] and the WL-CG[c].

High-quality CPR components	Baseline assessment[d], median (IQR)	Score after 2 weeks of playing the MOBICPR game at home[e], median (IQR)	*P* value[d,e]	Score after 2 weeks of not playing the MOBICPR game at home[c], median (IQR)	*P* value[e,f]
CC[g] rate (bpm[h])					
MOBICPR-IG	108 (18)	112 (18)	.24	112 (18)	.12
WL-CG	103 (22)	110 (10)	.38	—[i]	—
P value	.78	—	—	—	—
CC depth (mm)					
MOBICPR-IG	57 (7)	56 (7)	.56	57 (4)	.25
WL-CG	58 (6)	59 (2)	.27	—	—
P value	.16	—	—	—	—
CC fraction (%)					
MOBICPR-IG	70 (6)	72 (6)	.26	68 (9)	.15
WL-CG	68 (13)	70 (8)	.63	—	—
P value	.88	—	—	—	—
Volume of rescue breaths (mL)					
MOBICPR-IG	496 (369)	600 (463)	.54	473 (204)	.66
WL-CG	356 (147)	567 (270)	.45	—	—
P value	.64	—	—	—	—
Total QCPR[j] score (%)					
MOBICPR-IG	41 (54)	70 (41)	.01	77 (38)	.54
WL-CG	43 (42)	72 (46)	.24	—	—
P value	.62	—	—	—	—

[a]CPR: cardiopulmonary resuscitation.
[b]MOBICPR-IG: MOBICPR intervention group.
[c]WL-CG: wait-list control group.
[d]Measurement at baseline.
[e]Measurement after 2 weeks of playing the MOBICPR game at home.
[f]Measurement after 2 weeks of not playing the MOBICPR game at home.
[g]CC: chest compression.
[h]bpm: beats per minute.
[i]Not applicable.
[j]QCPR: Quality Cardiopulmonary Resuscitation.

Discussion

Principal Findings

In this study, playing the MOBICPR game at home for 2 weeks improved the theoretical knowledge of adult BLS in the participants but little their practical skills. These outcomes were expected, considering that the MOBICPR game was designed primarily to impart theoretical knowledge of adult BLS, rather than providing hands-on practice with an actual BLS manikin. To the best of our knowledge, only 2 studies have used data collected from manikin software to evaluate the practical parts of adult BLS as we did [27,28]. We observed in our study population that both the CC rate and the CC depth remain within the margins of the current ERC recommendation [34]; in comparison to our results, in the 2 studies [27,28], both the CC rate and the CC depth dropped below the margins after serious smartphone game playing. These 2 studies [27,28] also presented the total QCPR scores, and where our scores improved compared to theirs. Consequently, we recommend considering the MOBICPR game as a supplementary educational tool in future BLS course formats that incorporate immersive technologies [43,44] for retention of adult BLS knowledge.

In evaluating study participants performing adult BLS on a manikin, we observed 5 learning points (all reported in Tables 2 and 3), which could be useful for debriefing topics after BLS courses. Initially, a large number of participants struggled with checking the manikin's response as it lay face down. Some checked the response without turning the manikin onto its back, while others did so with the manikin still face down. After playing the MOBICPR game at home, only a minority checked the response after turning the manikin onto its back. Studies show that two-thirds of all patients are found in positions unsuitable for performing CCs, such as the recovery position [45]. The second learning point concerned the right time for chest exposure during CPR. Many participants removed the clothing before looking, listening, and feeling for signs of breathing, while others did so before applying AED electrodes to the manikin's bare chest. Studies indicate that exposing the chest during CPR can improve the rescuer's ability to locate the center of the patient's chest, leading to more effective CCs and reducing the risk of inaccurate compressions [46]. The third point was about shouting for help. A recent study revealed that almost all European BLS instructors teach laypersons to shout for help [47], even though it was removed from the ERC BLS guidelines [34,40]. Despite playing the MOBICPR game at home, the participants still tended to shout for help before calling emergency services. As a fourth point, we noticed that some participants attempted to multitask by calling the dispatcher and performing CCs simultaneously. This practice resulted in lower-quality CCs, as the focus was divided between providing information to the dispatcher and maintaining the 30:2 CC-to-rescue-breath ratio. Generation Z, like the participants in our study, tends to multitask and is more engaged in independent work [48]. Considering this insight, we are re-evaluating the recent ERC BLS guidelines, particularly their recommendation to activate the speakerphone or another hands-free feature on a mobile device before promptly initiating CPR [34]. Finally, we observed that almost all study participants failed to ensure safety before defibrillation when using an AED on the manikin. Issues arose before pressing the shock button, either because they did not check whether someone was touching the manikin or because they pressed the shock button prematurely. This highlights that using an AED is not intuitive for laypersons, as studies suggest, and special training should be considered [49].

The International Liaison Committee on Resuscitation provides a scientific statement on teaching laypersons adult BLS and suggests using TEL, such as serious smartphone games, to engage, motivate, and educate children and adolescents in saving more lives [34]. Several legitimate smartphone games have been identified as suitable for teaching adult BLS, but their content is questionable because it does not follow current BLS guidelines [50,51]. Moreover, most of them teach only hands-on CPR. Some also include ventilation and AED use [51]. However, the MOBICPR game was developed based on recent ERC BLS guidelines [34] and includes all the recommended BLS steps. In a recent MOBICPR study, students agreed that it was beneficial to play the MOBICPR game before practicing adult BLS on a manikin [41]. They also highly rated the usability of the MOBICPR game for providing adult BLS theoretical knowledge and practical skills. The results show that the MOBICPR game could be a novel, interactive, evidence-based BLS educational tool for playing at home after adult BLS training [41,52]. Moreover, our study revealed that the MOBICPR game could be an effective method for enhancing bystander willingness and awareness in performing CPR. This potential is demonstrated by the fact that all study participants introduced the MOBICPR game to their family members, relatives, or friends, as seen in similar studies where enhanced technology was used teaching adult BLS [53].

This gamified learning approach fits well with the educational theory heutagogy, also known as self-determined learning, where learners determine what they want to learn [8]. In the case of the MOBICPR game, learners can play it at any time to refresh their adult BLS knowledge without waiting for the next training session [54]. Moreover, the use of do-it-yourself manikins made from everyday items, such as plastic bottles, toilet paper, or even a pillow, for practicing CC techniques at home, especially in low-resource settings, coupled with the MOBICPR game, can potentially improve and solidify practical skills in adult BLS [55-58]. The MOBICPR game also includes gamification features, such as avatars, points, and various audio, textual, and graphical feedback. These gamification elements could motivate learners to engage with the game more frequently than they normally would [59]. Future educational tools, such as the MOBICPR game, should align with the 5 key messages outlined in the recent ERC BLS guidelines, ranging from recognizing cardiac arrest to learning the proper techniques for performing CPR [34]. This adherence is crucial for the effective education and retention of adult BLS skills, particularly following adult BLS courses in a home environment.

Limitations

This study has several limitations. First, because the study participants were only followed for 4 weeks, we were not able to show that the MOBICPR game improved their long-term retention of resuscitation knowledge and skills. Second, the sample size was small due to the lack of interest of participants in participating in the study and because only 1 generation of participants was able to be included at that time. Third, this was a single-faculty study, which limits the generalizability of the results. Fourth, in this study, participants were familiar with smartphone games. It is unclear how effective the MOBICPR game would be in children or older populations. Fifth, because this was a simulation-based study, the performance results may not be generalizable to real-life situations and could not present the impact on patient outcomes. Finaly, the content in the MOBICPR game was developed by researchers based on recent ERC BLS guidelines [34]. In the future, there are plans to introduce the MOBICPR game to the Slovenian National Resuscitation Council, with the goal of securing its certification, a process akin to that followed by the Italian Resuscitation Council for its smartphone-based serious games [60].

Conclusion

The home use of the MOBICPR game shows promise in enhancing the theoretical knowledge of adult BLS. Although there was no significant improvement in performing adult BLS or in retaining the related knowledge and skills, the study yielded important learning objectives for the enhancement of future adult BLS training. Further research is necessary to explore its lasting effects across various demographics and to determine the most effective use of the MOBICPR game in teaching adult BLS.

Acknowledgments

NF and GŠ are supported by Slovenian Research Agency (grants ARRS P2-0057, ARRS N3-0307, ARRS BI-US/22-24-138), NextGenerationEU, and the Ministry of Higher Education, Science and Innovation (C3330-22-953012). The authors would like to thank all nursing students who participated in this study.

Authors' Contributions

The study was carried out through collaboration among all authors. NF developed the study design and supervised the study. NF, GŠ, and ŠM drafted the manuscript. NF, GŠ, RG, and ŠM conducted data collection and analysis. NF, MS, PS, and RG interpreted results from the cardiopulmonary resuscitation point of view. RMC and RG conducted a comprehensive content review. All authors have read, revised, and approved the final manuscript.

Conflicts of Interest

NF is a member of the European Resuscitation Council (ERC) basic life support (BLS) Science and Education Committee. RG is the ERC director of guidelines and the International Liaison Committee on Resuscitation (ILCOR) and chair of the ILCOR Education, Implementation, and Teams Task Force. Other authors declare no conflicts of interest.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1235 KB - games_v12i1e56037_app1.pdf]

Multimedia Appendix 2
Adult BLS theoretical knowledge questionnaire. BLS: basic life support.
[DOCX File , 155 KB - games_v12i1e56037_app2.docx]

Multimedia Appendix 3
Out-of-hospital cardiac arrest scenario.
[DOCX File , 13 KB - games_v12i1e56037_app3.docx]

Multimedia Appendix 4
Adult BLS practical skills checklist. BLS: basic life support.
[DOCX File , 28 KB - games_v12i1e56037_app4.docx]

References

1. Rea T, Kudenchuk PJ, Sayre MR, Doll A, Eisenberg M. Out of hospital cardiac arrest: past, present, and future. Resuscitation 2021 Aug;165:101-109. [doi: 10.1016/j.resuscitation.2021.06.010] [Medline: 34166740]
2. Rajagopalan B, Shen WK, Patton K, Kutyifa V, Di Biase L, Al-Ahmad A, et al. Surviving sudden cardiac arrest—successes, challenges, and opportunities. J Interv Card Electrophysiol 2022 Sep 28;64(3):567-571. [doi: 10.1007/s10840-021-00969-1] [Medline: 33909223]
3. Merchant RM, Yang L, Becker LB, Berg RA, Nadkarni V, Nichol G, et al. Incidence of treated cardiac arrest in hospitalized patients in the United States*. Crit Care Med 2011;39(11):2401-2406. [doi: 10.1097/ccm.0b013e3182257459]

4. Meaney PA, Bobrow BJ, Mancini ME, Christenson J, de Caen AR, Bhanji F, CPR Quality Summit Investigators, the American Heart Association Emergency Cardiovascular Care Committee, the Council on Cardiopulmonary, Critical Care, Perioperative and Resuscitation. Cardiopulmonary resuscitation quality: [corrected] improving cardiac resuscitation outcomes both inside and outside the hospital: a consensus statement from the American Heart Association. Circulation 2013 Jul 23;128(4):417-435 [FREE Full text] [doi: 10.1161/CIR.0b013e31829d8654] [Medline: 23801105]
5. Dick-Smith F, Power T, Martinez-Maldonado R, Elliott D. Basic life support training for undergraduate nursing students: an integrative review. Nurse Educ Pract 2021 Jan;50:102957. [doi: 10.1016/j.nepr.2020.102957] [Medline: 33421680]
6. Riggs M, Franklin R, Saylany L. Associations between cardiopulmonary resuscitation (CPR) knowledge, self-efficacy, training history and willingness to perform CPR and CPR psychomotor skills: a systematic review. Resuscitation 2019 May;138:259-272. [doi: 10.1016/j.resuscitation.2019.03.019] [Medline: 30928504]
7. Cheng A, Magid DJ, Auerbach M, Bhanji F, Bigham BL, Blewer AL, et al. Part 6: resuscitation education science: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020 Oct 20;142(16_suppl_2):S551-S579 [FREE Full text] [doi: 10.1161/CIR.0000000000000903] [Medline: 33081527]
8. Greif R, Lockey A, Breckwoldt J, Carmona F, Conaghan P, Kuzovlev A, et al. European Resuscitation Council guidelines 2021: education for resuscitation. Resuscitation 2021 Apr;161:388-407. [doi: 10.1016/j.resuscitation.2021.02.016] [Medline: 33773831]
9. Junli A, Isa SNI, Ibrahim FS. Factors of cardiopulmonary resuscitation skills retention among healthcare providers: a scoping review. Nurse Educ Pract 2023 May;69:103617. [doi: 10.1016/j.nepr.2023.103617] [Medline: 36996552]
10. Cheskes S. High school CPR training: it's only an app away!!. Resuscitation 2017 Nov;120:A9-A10. [doi: 10.1016/j.resuscitation.2017.08.241] [Medline: 28974343]
11. Cetinkaya F, Akbulut Z, Dur N, Eryalcin O, Korkmaz M. Analysis of job satisfaction and burnout level of nurses in different generations. Int J Caring Sci 2017;10(3):1507-1513.
12. Tan SHE, Chin GF. Generational effect on nurses' work values, engagement, and satisfaction in an acute hospital. BMC Nurs 2023 Mar 30;22(1):88 [FREE Full text] [doi: 10.1186/s12912-023-01256-2] [Medline: 36997911]
13. Pei L, Wu H. Does online learning work better than offline learning in undergraduate medical education? A systematic review and meta-analysis. Med Educ Online 2019 Dec;24(1):1666538 [FREE Full text] [doi: 10.1080/10872981.2019.1666538] [Medline: 31526248]
14. Issleib M, Kromer A, Pinnschmidt HO, Süss-Havemann C, Kubitz JC. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students: a randomized controlled trial. Scand J Trauma Resusc Emerg Med 2021 Feb 01;29(1):27 [FREE Full text] [doi: 10.1186/s13049-021-00836-y] [Medline: 33526042]
15. Fijačko N, Metličar Š, Kleesiek J, Egger J, Chang TP. Virtual reality, augmented virtuality, or mixed reality in cardiopulmonary resuscitation: which extended reality am I using for teaching adult basic life support? Resuscitation 2023 Nov;192:109973. [doi: 10.1016/j.resuscitation.2023.109973] [Medline: 37730097]
16. Yalcinkaya T, Cinar Yucel S. Mobile learning in nursing education: a bibliometric analysis and visualization. Nurse Educ Pract 2023 Aug;71:103714. [doi: 10.1016/j.nepr.2023.103714] [Medline: 37552905]
17. Cheng P, Huang Y, Yang P, Wang H, Xu B, Qu C, et al. The effects of serious games on cardiopulmonary resuscitation training and education: systematic review with meta-analysis of randomized controlled trials. JMIR Serious Games 2024 Feb 06;12:e52990 [FREE Full text] [doi: 10.2196/52990] [Medline: 38319697]
18. Kumar Basak S, Wotto M, Bélanger P. E-learning, m-learning and d-learning: conceptual definition and comparative analysis. E-Learn Digital Media 2018 Jul 04;15(4):191-216 [FREE Full text] [doi: 10.1177/2042753018785180]
19. Yoo I, Lee Y. The effects of mobile applications in cardiopulmonary assessment education. Nurse Educ Today 2015 Feb;35(2):e19-e23. [doi: 10.1016/j.nedt.2014.12.002] [Medline: 25537170]
20. Chen B, Wang Y, Xiao L, Xu C, Shen Y, Qin Q, et al. Effects of mobile learning for nursing students in clinical education: a meta-analysis. Nurse Educ Today 2021 Feb;97:104706. [doi: 10.1016/j.nedt.2020.104706] [Medline: 33360012]
21. Kim JH, Park H. Effects of smartphone-based mobile learning in nursing education: a systematic review and meta-analysis. Asian Nurs Res (Korean Soc Nurs Sci) 2019 Feb;13(1):20-29 [FREE Full text] [doi: 10.1016/j.anr.2019.01.005] [Medline: 30659927]
22. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". 2011 Sep Presented at: MindTrek '11: 15th International Academic MindTrek Conference: Envisioning Future Media Environments; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
23. Krath J, Schürmann L, von Korflesch HF. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comp Hum Behav 2021 Dec;125:106963. [doi: 10.1016/j.chb.2021.106963]
24. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev 2019 Aug 15;32(1):77-112. [doi: 10.1007/s10648-019-09498-w]
25. Farsi Z, Yazdani M, Butler S, Nezamzadeh M, Mirlashari J. Comparative effectiveness of simulation versus serious game for training nursing students in cardiopulmonary resuscitation: a randomized control trial. Int J Comp Games Technol 2021 Mar 24;2021:1-12. [doi: 10.1155/2021/6695077]

26. Aksoy E. Comparing the effects on learning outcomes of tablet-based and virtual reality–based serious gaming modules for basic life support training: randomized trial. JMIR Serious Games 2019 May 01;7(2):e13442 [FREE Full text] [doi: 10.2196/13442] [Medline: 31042153]
27. de Sena DP, Fabrício DD, da Silva VD, Bodanese LC, Franco AR. Comparative evaluation of video-based on-line course versus serious game for training medical students in cardiopulmonary resuscitation: a randomised trial. PLoS One 2019 Apr 8;14(4):e0214722 [FREE Full text] [doi: 10.1371/journal.pone.0214722] [Medline: 30958836]
28. Demirtas A, Basak T, Sahin G, Sonkaya. The serious game and integrated simulator for cardiopulmonary resuscitation training in nursing students. Simulat Gaming 2022 Feb 28;53(2):97-110. [doi: 10.1177/10468781211073162]
29. Arif A, Arif A, Fasciglione K, Jafri F. The iterative development and testing of an interactive mobile application for skill retention of cardiopulmonary resuscitation among high school students: a pilot study. Simulat Gaming 2021 Jul 14;52(5):650-663. [doi: 10.1177/10468781211032692]
30. Yunanto RA, Wihastuti TA, Rachmawati SD. Comparison of CPR training with mobile application and simulation to knowledge and skill of CPR. NLJ 2017 Nov 29;2(2):183. [doi: 10.19184/nlj.v2i2.5943]
31. Interactive learning of basic CPR procedures. MOBICPR. URL: https://mobicpr.si/ [accessed 2024-03-21]
32. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
33. Painting for saints. Wikipedia. URL: https://en.wikipedia.org/w/index.php?title=Painting_for_Saints&oldid=1151677300 [accessed 2024-03-21]
34. Olasveengen TM, Semeraro F, Ristagno G, Castren M, Handley A, Kuzovlev A, et al. European Resuscitation Council guidelines 2021: basic life support. Resuscitation 2021 Apr;161:98-114. [doi: 10.1016/j.resuscitation.2021.02.009] [Medline: 33773835]
35. Papalexopoulou K, Chalkias A, Dontas I, Pliatsika P, Giannakakos C, Papapanagiotou P, et al. Education and age affect skill acquisition and retention in lay rescuers after a European Resuscitation Council CPR/AED course. Heart Lung 2014 Jan;43(1):66-71. [doi: 10.1016/j.hrtlng.2013.09.008] [Medline: 24246726]
36. Hernández-Padilla JM, Suthers F, Granero-Molina J, Fernández-Sola C. Effects of two retraining strategies on nursing students' acquisition and retention of BLS/AED skills: a cluster randomised trial. Resuscitation 2015 Aug;93:27-34. [doi: 10.1016/j.resuscitation.2015.05.008] [Medline: 26026776]
37. Fink A, Čander D, Kelebuda D, Gorjup D, Strnad M, Špindler M, et al. Slovenski indeks za nujno medicinsko pomoč. 1. izd. Ljubljana: Ministrstvo za zdravje, Služba za razvoj in organizacijo dejavnosti nujne medicinske pomoči; 2017.
38. Kiyohara K, Nishiyama C, Matsuyama T, Sado J, Kitamura T, Shimamoto T, et al. Out-of-hospital cardiac arrest at home in Japan. Am J Cardiol 2019 Apr 01;123(7):1060-1068. [doi: 10.1016/j.amjcard.2018.12.038] [Medline: 30654927]
39. Whitfield RH, Newcombe RG, Woollard M. Reliability of the Cardiff test of basic life support and automated external defibrillation version 3.1. Resuscitation 2003 Dec;59(3):291-314. [doi: 10.1016/s0300-9572(03)00246-6] [Medline: 14659599]
40. Perkins GD, Handley AJ, Koster RW, Castrén M, Smyth MA, Olasveengen T, Adult basic life supportautomated external defibrillation section Collaborators. European Resuscitation Council guidelines for resuscitation 2015: section 2. Adult basic life support and automated external defibrillation. Resuscitation 2015 Oct;95:81-99. [doi: 10.1016/j.resuscitation.2015.07.015] [Medline: 26477420]
41. Fijačko N, Metličar Š, Vinojčić D, Greif R, Masterson Creber R. Learn how to save the life of a victim of out-of-hospital cardiac arrest by playing a serious smartphone game. Resuscitation 2023 Jul;188:109859. [doi: 10.1016/j.resuscitation.2023.109859] [Medline: 37270089]
42. CPR scoring explained. Laerdal Medical. URL: https://laerdal.com/us/support/scoring/ [accessed 2024-02-15]
43. Alcázar Artero PM, Pardo Rios M, Greif R, Ocampo Cervantes AB, Gijón-Nogueron G, Barcala-Furelos R, et al. Efficiency of virtual reality for cardiopulmonary resuscitation training of adult laypersons: a systematic review. Medicine (Baltimore) 2023 Jan 27;102(4):e32736 [FREE Full text] [doi: 10.1097/MD.0000000000032736] [Medline: 36705392]
44. Kuyt K, Park S, Chang TP, Jung T, MacKinnon R. The use of virtual reality and augmented reality to enhance cardio-pulmonary resuscitation: a scoping review. Adv Simul (Lond) 2021 Apr 12;6(1):11 [FREE Full text] [doi: 10.1186/s41077-021-00158-0] [Medline: 33845911]
45. Wagner P, Schloesser S, Braun J, Arntz H, Breckwoldt J. In out-of-hospital cardiac arrest, is the positioning of victims by bystanders adequate for CPR? A cohort study. BMJ Open 2020 Sep 23;10(9):e037676 [FREE Full text] [doi: 10.1136/bmjopen-2020-037676] [Medline: 32967879]
46. Jo CH, Cho GC, Ahn JH, Ryu JY. The importance of victim chest exposure during cardiopulmonary resuscitation: a simulation study. J Emerg Med 2015 Feb;48(2):165-171. [doi: 10.1016/j.jemermed.2014.09.040] [Medline: 25453862]
47. Fijačko N, Greif R, Metličar Š, Gómez CA, Strnad M, Creber RMM. 344 Shout for help: still part of adult basic life support education, but not recommended by European resuscitation council guidelines. Resuscitation 2023 Nov;192:S137-S138. [doi: 10.1016/s0300-9572(23)00663-9]

48. Liu Y, Sun X, Zhang P, Han P, Shao H, Duan X, et al. Generation Z nursing students' online learning experiences during COVID-19 epidemic: a qualitative study. Heliyon 2023 Apr;9(4):e14755 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e14755] [Medline: 36968654]
49. Abelairas-Gómez C, Carballo-Fazanes A, Chang TP, Fijačko N, Rodríguez-Núñez A. Is the AED as intuitive as we think? Potential relevance of "the sound of silence" during AED use. Resusc Plus 2022 Dec;12:100323 [FREE Full text] [doi: 10.1016/j.resplu.2022.100323] [Medline: 36386767]
50. Kalz M, Lenssen N, Felzen M, Rossaint R, Tabuenca B, Specht M, et al. Smartphone apps for cardiopulmonary resuscitation training and real incident support: a mixed-methods evaluation study. J Med Internet Res 2014 Mar 19;16(3):e89 [FREE Full text] [doi: 10.2196/jmir.2951] [Medline: 24647361]
51. Fijačko N, Masterson Creber R, Gosak L, Štiglic G, Egan D, Chaka B, et al. Evaluating quality, usability, evidence-based content, and gamification features in mobile learning apps designed to teach children basic life support: systematic search in app stores and content analysis. JMIR Mhealth Uhealth 2021 Jul 20;9(7):e25437 [FREE Full text] [doi: 10.2196/25437] [Medline: 34283034]
52. Fijačko N, Greif R, Štiglic G, Metličar Š, Creber R. 322 Using a serious smartphone game to improve and retain knowledge of hand placement for adult chest compressions: a randomized wait-list controlled trial. Resuscitation 2023 Nov 1;192:S127. [doi: 10.1016/s0300-9572(23)00641-x]
53. Nas J, Thannhauser J, Konijnenberg LSF, van Geuns RM, van Royen N, Bonnes JL, et al. Long-term effect of face-to-face vs virtual reality cardiopulmonary resuscitation (CPR) training on willingness to perform CPR, retention of knowledge, and dissemination of CPR awareness: a secondary analysis of a randomized clinical trial. JAMA Netw Open 2022 May 02;5(5):e2212964 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.12964] [Medline: 35587346]
54. Abelairas-Gómez C, Martinez-Isasi S, Barcala-Furelos R, Varela-Casal C, Carballo-Fazanes A, Pichel-López M, et al. Training frequency for educating schoolchildren in basic life support: very brief 4-month rolling-refreshers versus annual retraining-a 2-year prospective longitudinal trial. BMJ Open 2021 Nov 30;11(11):e052478 [FREE Full text] [doi: 10.1136/bmjopen-2021-052478] [Medline: 34848519]
55. Nehra A, Ravindra P, Bhat R, Nagesh SK, Alok Y, Nisarg S, et al. Comparison between a low-cost model (CPR Pillow) and a mannequin in training hands only cardiopulmonary resuscitation (CPR): a randomised trial. Resusc Plus 2024 Mar;17:100518 [FREE Full text] [doi: 10.1016/j.resplu.2023.100518] [Medline: 38089839]
56. Nakagawa NK, Salles IC, Costa MM, Pinho JL, Böttiger BW. The "four-finger ruler" as a novel, simple and easy technique for hands placement during CPR training. Resuscitation 2023 Jan;182:109683. [doi: 10.1016/j.resuscitation.2022.109683] [Medline: 36586504]
57. Ohle R, Moskalyk M, Boissonneault E, Bilgasem A, Tissot E, McIsaac S. Is a homemade cardiopulmonary resuscitation (CPR) trainer non-inferior to a commercially available CPR mannequin in teaching high-quality CPR? A non-inferiority randomized control trial. Resusc Plus 2021 Jun;6:100134 [FREE Full text] [doi: 10.1016/j.resplu.2021.100134] [Medline: 34223390]
58. Agario Y, Tanaka H, Takyu H. Cost-effective cardiopulmonary resuscitation practice device for maintaining chest compression quality: a semi-randomized controlled clinical trial. J EMS Med 2023 May;2(2):55-63. [doi: 10.35616/jemsm.2022.00024]
59. Ratinho E, Martins C. The role of gamified learning strategies in student's motivation in high school and higher education: a systematic review. Heliyon 2023 Aug;9(8):e19033 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e19033] [Medline: 37636393]
60. Semeraro F, Imbriaco G, Del Giudice D, Antognoli M, Celin D, Cuttitta M, Collaborators. Empowering the next generation: an innovative “kids save lives" blended learning programme for schoolchildren training. Resuscitation 2024 Jan;194:110088. [doi: 10.1016/j.resuscitation.2023.110088] [Medline: 38101506]

Abbreviations

AED: automated external defibrillator
BLS: basic life support
bpm: beats per minute
CC: chest compression
CPR: cardiopulmonary resuscitation
ERC: European Resuscitation Council
m-learning: mobile learning
MOBICPR-IG: MOBICPR intervention group
OHCA: out-of-hospital cardiac arrest
QCPR: Quality Cardiopulmonary Resuscitation
TEL: technology-enhanced learning
WL-CG: wait-list control group

Edited by G Eysenbach; submitted 03.01.24; peer-reviewed by L Wei; comments to author 14.02.24; revised version received 19.02.24; accepted 10.03.24; published 05.04.24.

Please cite as:
Fijačko N, Masterson Creber R, Metličar Š, Strnad M, Greif R, Štiglic G, Skok P
Effects of a Serious Smartphone Game on Nursing Students' Theoretical Knowledge and Practical Skills in Adult Basic Life Support: Randomized Wait List–Controlled Trial
JMIR Serious Games 2024;12:e56037
URL: https://games.jmir.org/2024/1/e56037
doi:10.2196/56037
PMID:38578690

Original Paper

Adoption of Augmented Reality in Educational Programs for Nurses in Intensive Care Units of Tertiary Academic Hospitals: Mixed Methods Study

Suyoung Yoo[1*], BA; Sejin Heo[2*], MD, PhD; Soojin Song[3], BA; Aeyoung Park[3], BA; Hyunchung Cho[1], BA; Yuna Kim[3], PhD; Won Chul Cha[1,2], MD, PhD; Kyeongsug Kim[4], PhD; Meong Hi Son[1,2], MD

[1]Department of Digital Health, Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University, Seoul, Republic of Korea

[2]Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Gangnam-gu, Seoul, Republic of Korea

[3]Department of Nursing Education, Samsung Medical Center, Seoul, Republic of Korea

[4]Graduate School of Clinical Nursing Science, Sungkyunkwan University, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Meong Hi Son, MD
Department of Emergency Medicine
Samsung Medical Center
Sungkyunkwan University School of Medicine
115 Irwon-ro
Gangnam-gu, Seoul, 06355
Republic of Korea
Phone: 82 2 3410 2053
Fax: 82 2 3414 2832
Email: meonghison@gmail.com

Abstract

Background: In the wake of challenges brought by the COVID-19 pandemic to conventional medical education, the demand for innovative teaching methods has surged. Nurse training, with its focus on hands-on practice and self-directed learning, encountered significant hurdles with conventional approaches. Augmented reality (AR) offers a potential solution to addressing this issue.

Objective: The aim of this study was to develop, introduce, and evaluate an AR-based educational program designed for nurses, focusing on its potential to facilitate hands-on practice and self-directed learning.

Methods: An AR-based educational program for nursing was developed anchored by the Kern six-step framework. First, we identified challenges in conventional teaching methods through interviews and literature reviews. Interviews highlighted the need for hands-on practice and on-site self-directed learning with feedback from a remote site. The training goals of the platform were established by expert trainers and researchers, focusing on the utilization of a ventilator and extracorporeal membrane oxygenation system. Intensive care nurses were enrolled to evaluate AR education. We then assessed usability and acceptability of the AR training using the System Usability Scale and Technology Acceptance Model with intensive care nurses who agreed to test the new platform. Additionally, selected participants provided deeper insights through semistructured interviews.

Results: This study highlights feasibility and key considerations for implementing an AR-based educational program for intensive care unit nurses, focusing on training objectives of the platform. Implemented over 2 months using Microsoft Dynamics 365 Guides and HoloLens 2, 28 participants were trained. Feedback gathered through interviews with the trainers and trainees indicated a positive reception. In particular, the trainees mentioned finding AR particularly useful for hands-on learning, appreciating its realism and the ability for repetitive practice. However, some challenges such as difficulty in adapting to the new technology were expressed. Overall, AR exhibits potential as a supplementary tool in nurse education.

Conclusions: To our knowledge, this is the first study to substitute conventional methods with AR in this specific area of critical care nursing. These results indicate the multiple principal factors to take into consideration when adopting AR education in hospitals. AR is effective in promoting self-directed learning and hands-on practice, with participants displaying active engagement and enhanced skill acquisition.

Trial Registration: ClinicalTrials.gov NCT05629663; https://clinicaltrials.gov/study/NCT05629663.

(JMIR Serious Games 2024;12:e54188) doi:10.2196/54188

KEYWORDS

augmented reality; AR; clinical skills education; nurse education; technology-based education; education; nurse; nursing; allied health; technology-enhanced learning; interview; training; usability; acceptability; educational; teaching; ICU; intensive care unit; self-guided; self-directed; hands-on; adoption; TAM; Technology Acceptance Model; skill; acquisition

Introduction

In recent years, conventional education, especially in the medical field, has been challenged by the introduction of new technologies [1]. The COVID-19 pandemic further highlighted the limitations of conventional teaching methods [2]. Nurse training, with its emphasis on hands-on practice and self-directed learning, was particularly affected by the pandemic, making it evident that conventional training methods could not sustain the demands of the situation [3,4]. Given these constraints, the search for alternative, technology-driven educational methods intensified, aiming to address both physical resource and time challenges without compromising education quality [5,6]. In this context, an immersive learning environment, based on a computer-generated environment enabling real-time user interactions [7], has emerged as a promising solution.

Such an immersive platform merges augmented reality (AR) and virtual reality (VR), offering a dynamic 3D space for learners. This integration not only enhances the learning experience by providing a rich, immersive environment [8,9] but also reshapes the boundaries between reality and the virtual realm, paving the way for innovative learning avenues [10,11]. Exploring AR's potential reveals that its uses surpass merely aiding in remote education. AR also introduces real-time feedback mechanisms, empowering trainees to obtain immediate insights about their actions and performance through virtual aids [12,13]. This immediacy in feedback is invaluable, as it allows errors to be addressed promptly, fostering continuous improvement in learning [14,15].

In the field of critical care, there has been an exploration of the use of AR and VR in educational applications [16]. VR-enhanced training for tracheostomy care in the intensive care unit (ICU) setting has demonstrated the potential of education in an immersive learning environment [17,18]. Studies have been conducted for training mechanical ventilator settings and central line insertion, showing improvements in self-efficacy, increased familiarity, confidence, and reduced anxiety compared to conventional methods [13,17,19]. However, while the advantages are evident, existing research into AR and VR remains limited. These studies are usually one-time or short-term investigations, mainly focusing on the effectiveness of the immersive learning environment [20]. Moreover, integrating these technologies into a nursing curriculum represents an area yet to be fully explored [21].

In this study, we aimed to derive key considerations for each phase of implementation based on our experience of introducing an AR nursing program within an ICU in a tertiary hospital setting.

Methods

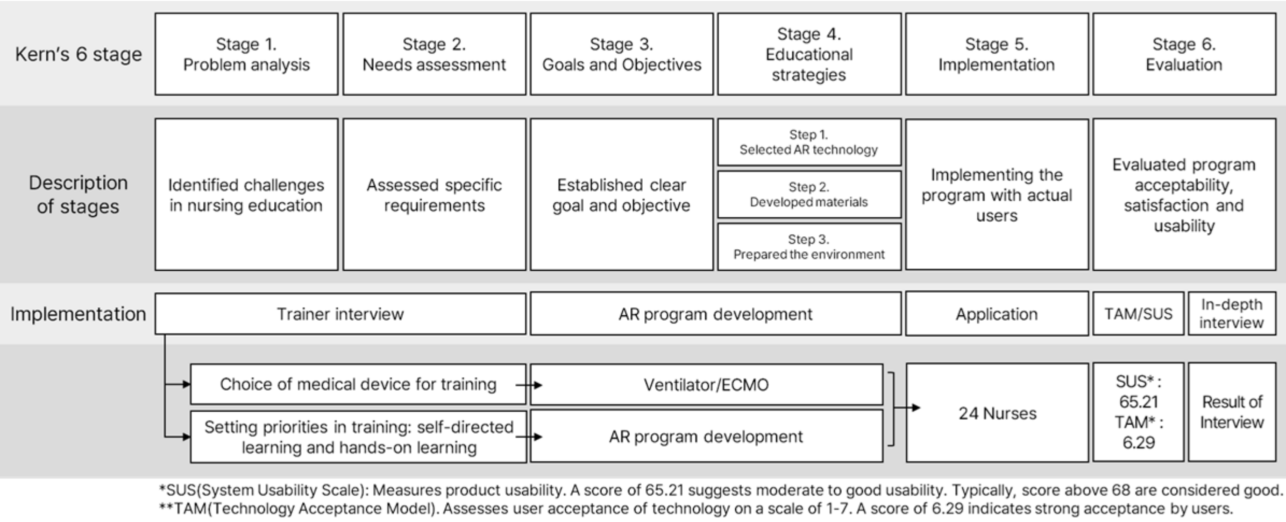
Experimental Design

Overview

Our methodology was refined based on the Kern six-step approach to transition nursing education into an immersive AR-based format [22], focusing on the following key stages: (1) problem identification, (2) needs assessment, (3) setting goals and objectives, (4) choosing educational strategies, (5) implementation, and (6) evaluation [23] (Figure 1).

Our initial steps involved conducting interviews with trainer nurses to discern existing issues and identify a procedure amenable to transition into an AR format [24]. We then developed and implemented an AR-based educational program encapsulating two distinct procedures. We surveyed and interviewed trainees, focusing on technology acceptance and usability.

Figure 1. Overall process of the adoption of augmented reality into nursing education based on the Kern 6-step framework [22]. AR: augmented reality; ECMO: extracorporeal mechanical oxygenation.



Stage 1: Problem Identification

In this foundational stage, we endeavored to identify core problems through a general needs assessment, employing two primary approaches: engaging in interviews with training nurses and reviewing the extant literature [24-27].

The interviews, conducted with nurses from the nursing education department, were based on a semistructured format and were held online or offline depending on participant preference. Voice recordings ensured the precise capture of data shared during these interactions.

In addition to the interviews, we explored previous studies with the aim of harvesting insights and identifying common issues found within the existing research landscape, thereby anchoring our findings in a robust context of existing knowledge [28].

Stage 2: Needs Assessment

Upon conducting the interviews in stage 1, we recognized the challenges endemic to conventional educational methods as identified by educators. We investigated possible solutions to these challenges, and found the need for feedback through remote supervision, especially regarding certain devices. We further identified key elements that should be considered in the development of the educational program. This stage also served to validate our problem identification process.

Stage 3: Goals And Objectives

We defined our goals through collaboration between trainers and researchers, focusing on improving access and proficiency with complex medical equipment.

Stage 4: Educational Strategies

Step 1: Selecting Appropriate Technology

With the imperative for hands-on practice in nursing education, AR was chosen to enable nurses to interact with virtual medical devices within a realistic clinical setting [29]. AR's ability to superimpose digital models onto the physical world allows for a highly interactive and immersive learning experience without the traditional constraints of location, time, or physical resources [30]. This aligns with our goal to empower self-directed learning, permitting nurses to engage in practical education at their own pace and convenience [31].

Step 2: Developing Educational Material

Navigating through the lens of self-directed learning and hands-on practice, we considered AR options that could facilitate tangible interaction with 3D objects. The development of materials required a detailed comprehension of the unique needs of nurses and the incorporation of AR content to support self-directed learning.

Step 3: Preparation of the Educational Environment

To facilitate a high-fidelity learning experience, our AR-based educational program was set within the hospital's simulation laboratory. A designated area within the lab was prepared, encompassing a minimum of 3×3 meters to provide trainees with sufficient room to maneuver and interact with the virtual elements without spatial constraints. To ensure uninterrupted delivery of our AR-based educational program, we utilized five HoloLens 2 devices. This approach was adopted to mitigate against battery and overheating issues that could disrupt the learning process. This setup was optimized to allow multiple nurses to receive training at the same time, promoting efficient learning throughout while maintaining an individualized, hands-on experience.

Stage 5: Implementation

We recruited the participants through an advertisement posted on the hospital's internal internet network. Our goal was to enroll a minimum of 10 participants for each session to ensure a dynamic and interactive learning environment while still allowing for personalized instruction. We designed the sessions to accommodate up to five nurses at a time, which was determined as the optimal number for both effective learning and space utilization within our AR setup. This small-group approach not only facilitated focused attention from the instructors but also ensured that each participant could engage deeply with the AR modules.

To maintain a high standard of education and safety, we appointed two experienced supervisors for each training session. These supervisors were selected based on their expertise in intensive care procedures and their familiarity with AR technology. Their role was to provide immediate assistance and feedback, ensuring that any technical issues could be addressed without disrupting the learning process. They were also tasked with observing the sessions to gather informal feedback, contributing to the continuous improvement of the program.

Stage 6: Evaluation

This study was conducted at a large academic tertiary hospital in Seoul, Korea, which accommodates more than 3100 nurses and 2000 inpatient beds. The research provided education to intensive care nurses and included a postsession survey and semistructured interview.

Ethical Considerations

The Institutional Review Board of Samsung Medical Center approved the study design (SMC-2022-08-058 and SMC-2022-08-079), and all trainers and trainees provided written informed consent before participating in the study, ensuring ethical adherence throughout the research. To protect the participants' privacy and confidentiality, all data collected during this study were anonymized or deidentified. Stringent data protection measures are in place, including the use of secure, encrypted data storage systems accessible only by authorized personnel. These precautions are designed to safeguard sensitive information and maintain the integrity of the data. Participants were compensated for their time and contribution. Each participant received 30,000 KRW (~US $22) upon completion of their involvement in the study. This compensation was intended to acknowledge their valuable time and effort and to offset any inconvenience associated with participation. The compensation structure was clearly communicated to all participants prior to their enrollment and was administered transparently to ensure fairness.

Outcome Measures of the Survey

Upon completion of the educational program, participants were asked to fill out a questionnaire evaluating their user experience. This evaluation was based on the theories of self-directed learning and hands-on practice, including questions on personal characteristics, job satisfaction, and appropriateness. To evaluate the AR program's acceptability and usability, we employed the System Usability Scale (SUS) and the Technology Acceptance Model (TAM) [20,32,33].

We chose the SUS for its proven reliability and efficiency across various technologies. The SUS is widely used across various domains, including software, websites, and medical devices, to assess overall user experience. Moreover, it has been validated in hospital environments and shown effectiveness with small sample sizes. The SUS consists of 10 simple questions presented in a 5-point Likert-scale format, assessing both positive and negative aspects of the system, with total scores ranging from 0 to 100 [34,35]. The TAM was selected for its emphasis on understanding user acceptance of information technology [16,36,37]. We adapted the TAM-based survey questions to fit the context of AR nurse education, informed by various relevant studies [24,26,27,38]. In contrast, the SUS was employed in its unmodified form. In addition, we conducted a correlation analysis between the TAM and SUS elements [37,39].

Outcome Measures of Interviews

Nurses who responded to the questionnaire were selectively screened for their willingness to participate in further interviews. These interviews were semistructured and guided by nursing education theories from previous studies. The format allowed flexibility, permitting up to two additional questions based on the responses of the interviewees.

Statistical Analysis

The statistical analysis was performed using R software (version 4.3.2). Continuous variables are expressed as either mean (SD) or median (IQR), depending on their distribution, while nominal variables are expressed as counts (n) and percentages (%). We performed a correlation analysis to examine the relationship between the TAM and SUS using survey data.

Results

Stage 1: Problem Identification

Overview

We obtained interview results from four nurses who are trainers and operators in the nursing education department. The insights garnered from the interviews are summarized below and detailed in Table 1.

Table 1. Key considerations in augmented reality (AR) education development as expressed during trainer interviews.

Category	Key details	Core implications
Educational needs and challenges	Training requirements necessary for handling advanced medical equipment; aligning AR educational content to complement the features of specific medical devices; adapting training modules to meet the unique demands precipitated by the COVID-19 pandemic; implementing streamlined training processes for the rapid acclimatization of new nurses; addressing the limitations inherent in conventional training techniques; tackling the deficit of hands-on training equipment in nursing training	Imperative for AR solutions in bridging training disparities and responding to progressive requirements
Program development and strategies	Tailoring educational programs to align with the diverse experience levels of nursing professionals; standardizing the phases of training to ensure uniformity and consistency in educational outcomes; formulating well-structured and strategic plans for nursing training; ensuring efficient and effective distribution and management of training equipment and resources	Criticality of a holistic design and meticulous implementation in AR training for optimal efficacy
Challenges and future concerns	Addressing trainees' physical challenges, such as the necessity to wear glasses or masks, in the training environment; guaranteeing the safety and appropriateness of both the devices and venues utilized for training; modifying AR training methods to be inclusive and effective for older nursing personnel; integrating strategies within training programs to manage and reduce trainee fatigue effectively	Recognition of and addressing present and prospective hurdles for continuous advancement in AR training

Requirement of Education in Difficult-to-Use Devices

Using difficult-to-use medical devices in health care can pose a significant challenge for medical staff due to the increased risk of errors, negatively impacting patient outcomes. Proper education on these devices is essential to ensure that medical staff can use them safely and effectively.

Lack of Resources: Space, Instructor, Time, and Cost

The lack of education resources in health care can be a significant challenge for health care organizations and medical staff [15]. Education is essential to ensure that health care professionals possess the necessary knowledge, skills, and abilities to provide safe and effective care to patients. However, the interviewees mentioned that many health care organizations face barriers in providing adequate education resources to their staff, which can negatively impact patient outcomes.

Another challenge is the lack of adequate education time. Health care professionals are often required to work long hours, and finding time to attend education sessions and complete the necessary course work can be difficult. This lack of time can make it difficult for health care organizations to provide education tailored to the specific needs of their staff.

Stage 2: Needs Assessment

Following the interviews, we identified crucial factors to consider when selecting educational topics. Educators highlighted the significant challenges of limited access to educational devices and instructors. Additionally, they emphasized the necessity for education in technically demanding skills. Trainers expressed a preference for educational topics that required hands-on practice. Their reasoning is grounded in the knowledge that complex devices are frequently used in treating patients with critical illnesses. The competence of nurses in operating these devices directly impacts patient outcomes. The responses from trainers and operators related to needs are summarized in Table 1.

Stage 3: Goals and Objectives

Selected Objectives

Based on the interview results, extracorporeal membrane oxygenation (ECMO) machines and ventilators were selected as the objectives for training owing to their complexity and difficulty of use. Ventilators were selected as important yet challenging devices to master. The complexity of ventilators, compounded by the multitude of lines and connections involved, can pose challenges for nurses with limited experience.

By contrast, an ECMO machine is a high-risk medical device that is essential for patients with COVID-19. When the alarm of an ECMO machine sounds, nurses must promptly find a solution. However, given its rarity, even experienced nurses may not have encountered this situation. Nevertheless, as this could pose a risk to the patient, appropriate education was deemed necessary, and therefore use of the ECMO machine was selected as the problem scenario for this evaluation.

Selected Goals

Our aim was to develop a sustainable AR-based educational program that could offer numerous benefits to trainers and trainees. These benefits include enhanced engagement and motivation, interactive and immersive learning experiences, and the facilitation of personalized learning. Key considerations for developing such a sustainable AR-based educational program encompass designing for scalability and accessibility, and incorporating user feedback to enhance the program in its maintenance and operation over time.

Stage 4: Educational Strategies

Step 1: Development of an AR-Based Educational Program

The ECMO machine educational program comprises 45 slides divided into four parts, each detailing the operation of the machine, managing machine disruptions, responding to "low battery" alarms, and addressing the loss of flow signal ("SIG") alarm. Each part includes approximately 8-10 steps, guiding trainees on how to manage each situation effectively.

The ventilator program is composed of three parts encompassing a total of 46 slides. Each part involves 26 steps related to ventilator settings and preuse inspections, as well as seven steps for application, alarm configuration, and educational content evaluation. The education process involves checking supplies, power sources, wall oxygen and medical gas connections, and exhalation cassette connections; powering on the device; performing preuse checks; connecting test tubes; selecting the target and application method; turning on the humidifier; configuring the mode and parameters; connecting the patient to the system; monitoring after patient application; and setting alarms.

Step 2: Adoption of Innovative Technology

The AR-based education was performed with a Maquet Servo-I mechanical ventilator and the RotaFlow II System Permanent Life Support ECMO machine from Getinge. We attempted to incorporate a 3D guide for hands-on practice and used videos to enhance understanding. The AR-based educational program was developed using the Microsoft Dynamics 365 Guides program. The program's content was delivered to users through a Microsoft HoloLens 2 device.

Step 3: Operation Plan

The previously designated simulation laboratory was successfully used during the AR-based educational program. The allocated space for the program proved sufficient, with each trainee having access to the minimum 3×3 meters space as planned [40]. This spatial arrangement allowed for unimpeded movement and interaction with the AR components, which was critical for the immersive learning experience.

In practice, the ventilator and ECMO machine simulations were conducted without any spatial constraints, enabling a total of 28 trainees to complete the training per the session schedules. The effective use of space was evidenced by the trainees' ability to perform the necessary tasks and their reported comfort level during the training sessions.

Stage 5: Implementation

The AR-based educational program platform was operational for a period of 2 months, with education sessions scheduled from 9 AM to 5 PM. This schedule allowed nurses to select their preferred date and time within this interval. To facilitate the program's implementation, we used five HoloLens 2 devices, along with two laptops for supervisor screen connections and two large screens for the research environment. Throughout the research, a total of 22 trainees actively engaged in the education sessions. The trainees' screens, as viewed through the HoloLens 2 devices, were immediately visible to the trainer, enabling real-time progress monitoring. Additionally, trainees were encouraged to request assistance if they encountered any difficulties during the session.

Stage 6: Evaluation

Participants

Training sessions were conducted by two trainers and two operators for the 28 nurses in the ICU from January 1 to February 3, 2022. Twenty-four nurses participated in the survey, 11 of whom took part in an in-depth interview. They were trained either in ventilator or ECMO machine usage with an even distribution across both groups. The participants' baseline characteristics are presented in Table 2. The median work experience was 3 (IQR 0-6.25) years with a mean of 3.75 (SD 3.90) years. There was a predominance of female participants (17/24, 71%). All participants belonged to the general nursing field with a slight majority working in the medical ICU compared to the surgical ICU (Table 2).

Additionally, all participants (24/24, 100%) owned smartphones and the majority (23/24, 95.83%) possessed either a tablet PC or laptop. Prior to the instruction, 13 (54%) nurses had previous experience with a head-mounted display.

Table 2. Demographic and clinical characteristics of the surveyed nurse trainees (N=24).

Characteristics	Trainees, n (%)
Method trained on	
ECMO[a]	12 (50)
Ventilator	12 (50)
Sex	
Male	7 (29)
Female	17 (71)
Medicine specialty	
Internal medicine	15 (63)
Surgical department	9 (38)
Experience (years)	
<1	5 (25)
1-2	3 (15)
3-4	2 (10)
5-6	2 (10)
≥7	3 (15)

[a]ECMO: extracorporeal membrane oxygenation.

Comparison of SUS and TAM Scores

In the usability test, the items "I think that I would like to use this system frequently" and "I don't think the system is unnecessarily complex" received the highest rating of 4.38 out of a possible 5, while the lowest-rated item, "I thought there was too much inconsistency in this system," received an average score of 1.83. The responses concerning technology acceptance were categorized into four areas according to the TAM: perceived usefulness (PU), perceived ease of use (PEU), perceived enjoyment (PE), and intention to use (IU). The survey included 15 questions scored on a 7-point scale. The item with the highest score was "It is fun to use," scoring 6.71, while the lowest-rated item was "It is easy to use," scoring 5.17. In further survey results, factors such as age, sex, department of work, and years of work did not impact satisfaction with the education or usability. All response results for the survey are provided in Multimedia Appendix 1.

Correlation of Usability and Acceptance

Our correlation analysis revealed varying degrees of association between SUS and TAM factors. For instance, there was a strong correlation between PU and IU and a moderate correlation between PE and PU. However, the correlation between PEU and IU was not significant (Table 3).

Table 3. Correlation between the Technology Acceptance Model–based survey items and System Usability Scale (SUS).

Variable	UE[a] (SUS)	PU[b]	PEU[c]	PE[d]	IU[e]
UE (SUS)					
r	1	0.2	0.34	0.15	0.22
P value	—[f]	.50	.12	.45	.30
PU					
r	0.2	1	0.2	0.69	0.76
P value	.50	—	.50	<.001	<.001
PEU					
r	0.34	0.20	1		0.31
P value	.12	.50	—	.34	.33
PE					
r	0.15	0.69	0.32	1	0.72
P value	.45	<.001	.34	—	<.001
IU					
r	0.22	0.76	0.31	0.72	1
P value	.30	<.001	0.33	<.001	—

[a]UE: user experience.
[b]PU: perceived usefulness.
[c]PEU: perceived ease of use.
[d]PE: perceived enjoyment.
[e]IU: intention to use.
[f]Not applicable.

Insights From Participant Interviews

Four participants completed interviews related to their experiences with the AR-based educational program. Key insights from these interviews have been collated and are summarized in Table 4. We present a curated selection of interview responses that most effectively capture the key insights. These selections were thoroughly chosen for their relevance and ability to represent the broader findings of the study.

Overall, the evaluation of the AR-based education was positive, with participants indicating that AR could enhance their actual clinical performance. AR technology is particularly well-suited for individuals interested in self-directed or hands-on learning theories. Nurses were found to be open to education using innovative technology. When asked if they needed assistance with the curriculum, no participant responded negatively regarding the content. However, some participants did express a need for help in adapting to new devices and technologies.

Table 4. Trainees' feedback after augmented reality (AR) implementation in the training program.

Category and subcategory	Details
Motivation	
Intrinsic motivation	Interest in the integration of AR into educational settings; desire among learners for practical experience with medical devices; expectations that AR technology will significantly improve learning efficacy
Reasons for participation	Influence of colleague recommendations; curiosity about AR teaching methods; specific needs related to own job
Existing issues	Noticing varying standards in educational quality; underscoring the need for improvement; necessity to establish standardized training procedures and protocols; need for training programs to be customized to individual learning styles and needs
Learning preferences	
Preferred learning method	Balancing traditional (54.5%) and self-directed (45.5%) learning approaches; valuing feedback and interaction in traditional learning; preference for flexibility and pace in self-directed learning
Face-to-face versus nonface-to-face	Equilibrium between face-to-face (54.5%) and remote learning (45.5%); diverse preferences shaped by feedback, comfort, and flexibility
Feedback	
Practical use	Majority opinion holding that AR technology is beneficial for skill development; mixed opinions regarding the real-world applicability of AR in professional settings; varied levels of expectation regarding the use of AR devices in educational contexts
Training experience	Recognizing the benefits of AR in providing realistic scenarios, allowing for self-directed learning, and enabling repeated practice; challenges include unfamiliarity with AR, focus on operation over content, and limited interaction
Content and support	General satisfaction with AR content amid comparisons to traditional methods; requirement for technical support and assistance in AR training
Comparative analysis and outlook	AR's superiority in learning pace, error identification, and training repetition over conventional methods; challenges in mastering AR operation and content depth; mixed perspectives on AR replacing traditional methods (viewed as supplementary); AR's efficacy in specific scenarios; considered resource-intensive for broad implementation; potential for enhancing self-directed and iterative learning
Future considerations	
Target demographics for AR training	Target new nurses, individuals lacking device experience, and department transfers; however, limited relevance for experienced nurses
Benefits of self-learning with AR	Reduced pressure, time efficiency, review flexibility; utility in learning uncommon scenarios and repeatable sessions
Concerns with self-learning	Limitations of AR training in actual clinical settings; lack of communal learning opportunities in AR environments; concerns over system errors and device quantity limitations

Discussion

Principal Results

The principal findings of this study provide valuable insights into the strategic translation of conventional critical nursing education to AR-based education platforms in the use of difficult-to-use medical devices [41]. Through interviews conducted with trainers before program development, the study successfully identified the specific needs and requirements of trainers in critical care nursing education. The study employed AR-based educational technology to enhance self-directed learning and hands-on practice.

In the educational strategies employed, the study leveraged the unique features of AR to facilitate self-directed learning. By offering interactive and self-controlled learning experiences, AR empowered trainees to take ownership of their learning process. The program incorporated instructional materials and modules that allowed learners to explore and acquire knowledge at their own pace, fostering a sense of autonomy and self-guided learning. The use of AR also enabled real-time feedback and assessment, allowing learners to track their progress and identify areas for improvement [29].

Through the overlay of 3D objects and virtual models onto real-world settings, trainees engaged in simulated scenarios closely resembling authentic ICU environments. This hands-on component of the program enabled learners to apply their knowledge and skills in realistic contexts, promoting an understanding of the subject matter and the development of critical thinking and problem-solving abilities.

This study highlights how AR technology significantly contributes to the success of self-directed learning and hands-on practice. The utilization of AR technology facilitates active engagement, learner-centeredness, and skill development, thereby enhancing the overall effectiveness of critical care nursing education. Moreover, we provide useful insights based on the perspectives of trainers and operators of the platform. The inherent nature of education often necessitates a lower number of educators compared to learners. The main strength of our study thus lies in presenting an infrequent perspective of educators, a viewpoint seldom encountered within the large-scale hospital setting.

Comparison With Prior Work

AR technology has been extensively explored in areas related to nursing education [42] such as surgical simulation [43], anatomy education [20], and patient safety education [44]. However, it is worth noting that this study represents the first investigation into the use of AR for replacing a conventional educational program in the use of difficult-to-use devices such as an ECMO machine and mechanical ventilator specifically within the critical care nursing field. By incorporating AR into these fields, this study pioneers the integration of innovative approaches in nursing education.

Limitations

This study, being characteristic of a pilot study to identify and apply new educational methods, has the limitation of a restricted number of participants. In further research, a larger sample size could be recruited to identify factors influencing user acceptability and to enhance usability, leveraging insights for more effective implementation.

Implications

This study highlights the implications of AR in future research and practice. The findings suggest the need for longitudinal studies to assess AR's long-term impact on clinical performance and patient outcomes, and to explore its scalability and cost-effectiveness compared to traditional training. Practically, the results of our study indicate that institutions adopting AR should invest in technical support and training and consider integrating AR as a supplementary tool in curricula for a blended learning approach.

Conclusions

This study provides insights on the development, launch, and operation of an AR-based medical educational program. The study suggests that an AR-based educational program can be an alternative to compensate for insufficient resources for conventional critical care nursing education. Further research can be conducted to compare the effectiveness and feasibility of this program with other AR-based educational programs and traditional nursing educational programs.

Acknowledgments

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI23C038700). We would like to express our sincere gratitude to all the individuals who contributed to the successful completion of this study. Primarily, we would like to thank the nursing education chiefs for their valuable insights and participation in the interviews. We are also grateful to the nurses who willingly participated in the study, sharing their experiences and providing valuable feedback. Additionally, we extend our appreciation to the operators who assisted in the implementation of the augmented reality–based educational program. The authors would like to thank Seungheon Choo for his invaluable assistance with the English proofreading of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
All response results for the System Usability Scale and survey based on the Technology Acceptance Model.
[DOCX File , 16 KB - games_v12i1e54188_app1.docx]

References

1. Lockee BB. Online education in the post-COVID era. Nat Electron 2021 Jan 25;4(1):5-6. [doi: 10.1038/s41928-020-00534-0]
2. Gause G, Mokgaola IO, Rakhudu MA. Technology usage for teaching and learning in nursing education: an integrative review. Curationis 2022 Jun 15;45(1):e1-e9 [FREE Full text] [doi: 10.4102/curationis.v45i1.2261] [Medline: 35792609]
3. Robbins DE. The effect of hands-on practice on the skills proficiency of nursing home nurses. Int J Stud Nurs 2020;5(1):39. [doi: 10.20849/ijsn.v5i1.717]
4. Örs M. The self-directed learning readiness level of the undergraduate students of midwife and nurse in terms of sustainability in nursing and midwifery education. Sustainability 2018 Oct 07;10(10):3574. [doi: 10.3390/su10103574]
5. Johnsen HM, Briseid HS, Brodtkorb K, Slettebø Å, Fossum M. Nursing students' perceptions of combining hands-on simulation with simulated patients and a serious game in preparing for clinical placement in home healthcare: a qualitative study. Nurse Educ Today 2021 Feb;97:104675 [FREE Full text] [doi: 10.1016/j.nedt.2020.104675] [Medline: 33302184]
6. Noble C, Billett S, Hilder J, Teodorczuk A, Ajjawi R. Enriching medical trainees' learning through practice: a video reflexive ethnography study protocol. BMJ Open 2019 Aug 22;9(8):e031577 [FREE Full text] [doi: 10.1136/bmjopen-2019-031577] [Medline: 31444194]
7. Moro C. Utilizing the metaverse in anatomy and physiology. Anat Sci Educ 2023 Jan 05;16(4):574-581. [doi: 10.1002/ase.2244] [Medline: 36545794]

8. Patel ST, Shah S, Sood RP, Siddiqui Z, McKay-Davies I. The implementation of virtual clinical skills teaching in improving procedural confidence in ENT trainees. Adv Med Educ Pract 2021;12:965-969 [FREE Full text] [doi: 10.2147/AMEP.S322965] [Medline: 34475794]
9. Ordov K, Madiyarova A, Ermilov V, Tovma N, Murzagulova M. New trends in education as the aspect of digital technologies. Int J Mech Eng Technol 2019;10(2):1319-1330 [FREE Full text]
10. Dwivedi YK, Hughes L, Baabdullah AM, Ribeiro-Navarrete S, Giannakis M, Al-Debei MM, et al. Metaverse beyond the hype: multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. Int J Inf Manage 2022 Oct;66:102542. [doi: 10.1016/j.ijinfomgt.2022.102542]
11. Zhang X, Chen Y, Hu L, Wang Y. The metaverse in education: definition, framework, features, potential applications, challenges, and future research topics. Front Psychol 2022 Oct 11;13:1016300 [FREE Full text] [doi: 10.3389/fpsyg.2022.1016300] [Medline: 36304866]
12. Dinh A, Tseng E, Yin AL, Estrin D, Greenwald P, Fortenko A. Perceptions about augmented reality in remote medical care: interview study of emergency telemedicine providers. JMIR Form Res 2023 Mar 28;7:e45211 [FREE Full text] [doi: 10.2196/45211] [Medline: 36976628]
13. Rochlen LR, Levine R, Tait AR. First-person point-of-view-augmented reality for central line insertion training: a usability and feasibility study. Simul Healthc 2017 Feb;12(1):57-62 [FREE Full text] [doi: 10.1097/SIH.0000000000000185] [Medline: 27930431]
14. Lateef F. Simulation-based learning: just like the real thing. J Emerg Trauma Shock 2010 Oct;3(4):348-352 [FREE Full text] [doi: 10.4103/0974-2700.70743] [Medline: 21063557]
15. Real FJ, Whitehead M, Rosen BL, Meisman A, Crosby LE, Klein MD, et al. Assessment of competency-based behavioral health anticipatory guidance skills among pediatric residents: the role of virtual reality. J Technol Behav Sci 2021 Mar 18;7(1):115-124. [doi: 10.1007/s41347-021-00202-w]
16. Samadbeik M, Aslani N, Maleki M, Garavand A. Acceptance of mobile health in medical sciences students: applying technology acceptance model. Inform Med Unlocked 2023;40:101290. [doi: 10.1016/j.imu.2023.101290]
17. Chiang D, Huang C, Cheng S, Cheng J, Wu C, Huang S, et al. Immersive virtual reality (VR) training increases the self-efficacy of in-hospital healthcare providers and patient families regarding tracheostomy-related knowledge and care skills: a prospective pre-post study. Medicine 2022 Jan 14;101(2):e28570 [FREE Full text] [doi: 10.1097/MD.0000000000028570] [Medline: 35029229]
18. Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 2020 Apr;147:103778. [doi: 10.1016/j.compedu.2019.103778]
19. Heo S, Moon S, Kim M, Park M, Cha WC, Son MH. An augmented reality-based guide for mechanical ventilator setup: prospective randomized pilot trial. JMIR Serious Games 2022 Jul 22;10(3):e38433 [FREE Full text] [doi: 10.2196/38433] [Medline: 35867382]
20. Baashar Y, Alkawsi G, Ahmad WNW, Alhussian H, Alwadain A, Capretz LF, et al. Effectiveness of using augmented reality for training in the medical professions: meta-analysis. JMIR Serious Games 2022 Jul 05;10(3):e32715 [FREE Full text] [doi: 10.2196/32715] [Medline: 35787488]
21. Aebersold M, Dunbar D. Virtual and augmented realities in nursing education: state of the science. Annu Rev Nurs Res 2020 Dec 01;39(1):225-242. [doi: 10.1891/0739-6686.39.225] [Medline: 33431644]
22. Heidari H, Beni ZHM, Deris F. Using Kern model to design, implement, and evaluate an infection control program for improving knowledge and performance among undergraduate nursing students: a mixed methods study. BMC Med Educ 2023 Oct 25;23(1):795 [FREE Full text] [doi: 10.1186/s12909-023-04775-3] [Medline: 37880731]
23. Sweet LR, Palazzi DL. Application of Kern's six-step approach to curriculum development by global health residents. Educ Health 2015;28(2):138-141. [doi: 10.4103/1357-6283.170124] [Medline: 26609014]
24. Uymaz P, Uymaz AO. Assessing acceptance of augmented reality in nursing education. PLoS One 2022 Feb 17;17(2):e0263937 [FREE Full text] [doi: 10.1371/journal.pone.0263937] [Medline: 35176073]
25. Elo S, Kääriäinen M, Kanste O, Pölkki T, Utriainen K, Kyngäs H. Qualitative content analysis. SAGE Open 2014 Feb 11;4(1):215824401452263. [doi: 10.1177/2158244014522633]
26. Dumpit DZ, Fernandez CJ. Analysis of the use of social media in Higher Education Institutions (HEIs) using the Technology Acceptance Model. Int J Educ Technol High Educ 2017 Mar 1;14(1):5. [doi: 10.1186/s41239-017-0045-2]
27. Papakostas C, Troussas C, Krouska A, Sgouropoulou C. Exploring users' behavioral intention to adopt mobile augmented reality in education through an extended Technology Acceptance Model. Int J Hum Comput Interact 2022 Apr 17;39(6):1294-1302. [doi: 10.1080/10447318.2022.2062551]
28. Carroll JK, Yancey AK, Spring B, Figueroa-Moseley C, Mohr DC, Mustian KM, et al. What are successful recruitment and retention strategies for underserved populations? Examining physical activity interventions in primary care and community settings. Transl Behav Med 2011 Jun 12;1(2):234-251 [FREE Full text] [doi: 10.1007/s13142-011-0034-2] [Medline: 24073048]
29. Chaghari M, Saffari M, Ebadi A, Ameryoun A. Empowering education: a new model for in-service training of nursing staff. J Adv Med Educ Prof 2017 Jan;5(1):26-32 [FREE Full text] [Medline: 28180130]

30. Al-Ansi AM, Jaboob M, Garad A, Al-Ansi A. Analyzing augmented reality (AR) and virtual reality (VR) recent development in education. Soc Sci Humanit Open 2023;8(1):100532. [doi: 10.1016/j.ssaho.2023.100532]
31. Lau ST, Siah CJR, Loh WL, Rusli KDB, Schmidt LT, Lim FP, et al. Enhancing professional competency in clinical procedures using head-mounted display virtual reality - a mixed method study. Med Educ Online 2023 Dec 05;28(1):2232134 [FREE Full text] [doi: 10.1080/10872981.2023.2232134] [Medline: 37406175]
32. Davis FD. A technology acceptance model for empirically testing new end-user information systems: theory and results. Doctoral dissertation. MIT Sloan School of Management, MIT Libraries. 1986. URL: https://dspace.mit.edu/handle/1721.1/15192 [accessed 2024-04-17]
33. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart 1989 Sep;13(3):319. [doi: 10.2307/249008]
34. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press, Taylor & Francis Group; 1996:189-194.
35. Cheah W, Mat Jusoh N, Aung MMT, Ab Ghani A, Mohd Amin Rebuan H. Mobile technology in medicine: development and validation of an adapted System Usability Scale (SUS) questionnaire and modified Technology Acceptance Model (TAM) to evaluate user experience and acceptability of a mobile application in MRI safety screening. Indian J Radiol Imaging 2023 Jan 07;33(1):36-45 [FREE Full text] [doi: 10.1055/s-0042-1758198] [Medline: 36855734]
36. Wahdain EA, Ahmad MN. User acceptance of information technology: factors, theories and applications. J Inf Syst Res Innov 2014 Jan;6(1):17-25 [FREE Full text]
37. Kim S, Garrison G. Investigating mobile wireless technology adoption: an extension of the technology acceptance model. Inf Syst Front 2008 Apr 9;11(3):323-333. [doi: 10.1007/s10796-008-9073-8]
38. Martínez-Caro E, Cegarra-Navarro J, Solano-Lorente M. An extension of the technology acceptance model in hospital-in-the-home units. In: Cruz-Cunha MM, Miranda IM, Gonçalves P, editors. Handbook of Research on ICTs and Management Systems for Improving Efficiency in Healthcare and Social Care. Hershey, PA: IGI Global; 2013:1191-1207.
39. Chuenyindee T, Montenegro LD, Ong AKS, Prasetyo YT, Nadlifatin R, Ayuwati ID, et al. The perceived usability of the learning management system during the COVID-19 pandemic: integrating system usability scale, technology acceptance model, and task-technology fit. Work 2022;73(1):41-58. [doi: 10.3233/WOR-220015] [Medline: 35871387]
40. Zorzal ER, Campos Gomes JM, Sousa M, Belchior P, da Silva PG, Figueiredo N, et al. Laparoscopy with augmented reality adaptations. J Biomed Inform 2020 Jul;107:103463 [FREE Full text] [doi: 10.1016/j.jbi.2020.103463] [Medline: 32562897]
41. AlGerafi MAM, Zhou Y, Oubibi M, Wijaya TT. Unlocking the potential: a comprehensive evaluation of augmented reality and virtual reality in education. Electronics 2023 Sep 20;12(18):3953. [doi: 10.3390/electronics12183953]
42. Tang YM, Chau KY, Kwok APK, Zhu T, Ma X. A systematic review of immersive technology applications for medical practice and education - Trends, application areas, recipients, teaching contents, evaluation methods, and performance. Educ Res Rev 2022 Feb;35:100429. [doi: 10.1016/j.edurev.2021.100429]
43. Williams MA, McVeigh J, Handa AI, Lee R. Augmented reality in surgical training: a systematic review. Postgrad Med J 2020 Sep 30;96(1139):537-542. [doi: 10.1136/postgradmedj-2020-137600] [Medline: 32229513]
44. Chen P, Liou W. The effects of an augmented reality application developed for paediatric first aid training on the knowledge and skill levels of nursing students: An experimental controlled study. Nurse Educ Today 2023 Jan;120:105629. [doi: 10.1016/j.nedt.2022.105629] [Medline: 36375385]

Abbreviations

AR: augmented reality
ECMO: extracorporeal membrane oxygenation
ICU: intensive care unit
IU: intention to use
PE: perceived enjoyment
PEU: perceived ease of use
PU: perceived usefulness
SUS: System Usability Scale
TAM: Technology Acceptance Model
VR: virtual reality

Edited by A McDaniel; submitted 01.11.23; peer-reviewed by G Alkawsi, G Evangelinos; comments to author 01.12.23; revised version received 22.12.23; accepted 31.03.24; published 23.05.24.

Please cite as:
Yoo S, Heo S, Song S, Park A, Cho H, Kim Y, Cha WC, Kim K, Son MH
Adoption of Augmented Reality in Educational Programs for Nurses in Intensive Care Units of Tertiary Academic Hospitals: Mixed Methods Study
JMIR Serious Games 2024;12:e54188
URL: https://games.jmir.org/2024/1/e54188
doi:10.2196/54188
PMID:38780998

Original Paper

Exploring the Impact of a Persuasive Serious Video Game (Farmily) on Promoting Home Gardening Among Novices: Design and Randomized Controlled Trial

Carlos Alberto García de Alba-Chávez[1], MSCS; Ismael Edrein Espinosa-Curiel[2], PhD; Rosa María Michel-Nava[1], PhD

[1]Departamento de Sistemas y Computación, Instituto Tecnológico de Ciudad Guzmán, Tecnológico Nacional de México, Ciudad Guzmán, Jalisco, Mexico

[2]Unidad de Transferencia Tecnológica Tepic, Centro de Investigación Científica y de Educación Superior de Ensenada, Tepic, Nayarit, Mexico

Corresponding Author:
Ismael Edrein Espinosa-Curiel, PhD
Unidad de Transferencia Tecnológica Tepic
Centro de Investigación Científica y de Educación Superior de Ensenada
Andador 10, entre Calles 3 y 4
Ciudad del Conocimiento
Tepic, Nayarit, 63173
Mexico
Phone: 52 3111295930 ext 28607
Email: ecuriel@cicese.edu.mx

Abstract

Background: Home gardens worldwide provide sustenance, economic support, and access to fresh produce and promote household well-being, health, self-sufficiency, and food security. However, they face significant challenges worldwide and necessitate innovative promotion approaches. Serious video games have proven effective in promoting agricultural knowledge. However, more research is needed on the persuasive potential of agriculture games to influence players' thoughts, attitudes, and behaviors. This provides an opportunity to examine the impact of persuasive games on promoting home gardening among novices.

Objective: This study aims to describe the design and development of Farmily, a persuasive video game promoting home gardening among novices. In addition, it evaluated the effectiveness of Farmily and compared its impact with that of a traditional home gardening workshop. Furthermore, the study explored how game enjoyment relates to the game's outcomes.

Methods: A randomized controlled trial with 50 novice gardening participants aged 20 to 50 years was carried out. Participants were randomly assigned to a control group (1.5-hour workshop) or an experimental group (1.5-hour Farmily session). Pre- and postintervention assessments were conducted. We evaluated Farmily's impact on knowledge, attitudes, perceived self-efficacy, and intentions regarding initiating home gardens. In addition, we investigated the user enjoyment and its relationship with the game's effect outcomes.

Results: The experimental group significantly improved their knowledge (t_{24}=4.26; *P*<.001), attitude (z_{24}=2.98; *P*=.003), self-efficacy (t_{24}=2.6; *P*=.02), and intention to initiate home gardens (z_{24}=4.15; *P*<.001). The experimental group showed similar effectiveness in knowledge transfer (t_{24}=–1.71; *P*=.09) and a more significant impact on attitude (z_{24}=2.73; *P*=.006), self-efficacy (t_{24}=2.21; *P*=.03), and intention to start a home garden (t_{24}=–5.33; *P*<.001) than the control group. Farmily was well received by the intervention group, generating high enjoyment. Furthermore, user enjoyment substantially correlated with user attitudes (r_{23}=0.72; *P*<.001) and self-efficacy (r_{23}=0.67; *P*<.001), yet no discernible association was observed among user enjoyment, knowledge (r_{23}=0.26; *P*=.20), and intention (r_{23}=0.06; *P*=.77).

Conclusions: Evidence indicates that Farmily appears to be a viable tool for promoting home gardening among novices in the short term. Farmily demonstrated similar effects in knowledge improvement to those of a traditional workshop and had a more significant impact on the other variables. In addition, we found that the player's gaming experience positively relates to the player's attitudes and self-efficacy. A well-powered randomized controlled trial with more diverse samples and extended follow-up periods will be conducted to establish the long-term efficacy of Farmily and gain a deeper understanding of the influence of enjoyment on game outcomes.

(JMIR Serious Games 2024;12:e60771) doi:10.2196/60771

KEYWORDS
serious video game; persuasive game; home gardening; knowledge; attitude; self-efficacy; intention

Introduction

Promotion of Home Gardening

Home gardens, a long-standing tradition of small-scale food cultivation near human settlements [1], remain a cornerstone of agriculture, providing essential sustenance and economic support for families worldwide [2]. These gardens, cultivated for centuries, serve multifaceted purposes within local food systems. Alongside economic benefits such as financial savings and income generation, they provide year-round nutrition and access to fresh produce and contribute to biodiversity conservation [2-4]. In addition, home gardens contribute to ecological production, enhance esthetics, and promote sustainable urban development [5]. Moreover, maintaining and using home gardens contributes significantly to household well-being by promoting psychological and physical health and fostering familiar and social connections and bonds [6]. Another significant benefit is fostering self-sufficiency as individuals cultivate plants or food crops at home to meet their sustenance needs, thereby reducing dependence on external sources [7]. This characteristic plays a crucial role in bolstering food security [8], a role that was particularly emphasized during the COVID-19 pandemic [9]. Due to these advantages, home gardening initiatives have expanded to include elementary schools [10], urban areas, and household environments [11]. They remain a valuable pursuit even for economically active individuals who can afford groceries.

Despite the global significance of home gardens, many people are not interested or hesitant to start one [12], contributing to their decline [13], a trend particularly concerning in countries such as Mexico [14]. Various strategies have been proposed to promote home gardening, ranging from campaigns highlighting its benefits to educational initiatives offering guidance through manuals, courses, workshops, and guides [15]. However, these initiatives often focus on individuals already motivated to start gardening, potentially missing the chance to inspire and engage newcomers. Furthermore, these strategies encounter challenges inherent to the multifaceted nature of home gardening, including the necessity to develop skills; address concerns; overcome motivation issues [16,17]; acquire interdisciplinary knowledge in fields such as biology, chemistry, and economics [10]; and achieve a balance between theoretical learning and hands-on activities [18].

To bridge the gap between promoting home gardening and its adoption, innovative strategies are essential to attract a diverse audience, motivate and empower individuals, and provide practical learning opportunities. It is crucial that these strategies mainly focus on novices—individuals who are open to learning about gardening but may not be strongly motivated to pursue it actively. These individuals often lack extensive knowledge, experience, and confidence in gardening and may not fully grasp its benefits. This situation presents an opportunity to develop accessible and engaging tools that inspire, educate, build confidence, and foster sustained interest in home gardening. Persuasive video games provide a promising platform to enhance users' knowledge, skills, and experiences through interactive platforms [19] and foster positive changes in perceptions, attitudes, or behaviors [20,21].

Persuasive Video Games

Persuasive video games, also known as serious games for change or persuasive games, aim to actively engage players while promoting positive changes in their thoughts, attitudes, feelings, actions, or behaviors [20-22]. They integrate interactive gameplay, narrative elements, educational content, and persuasive strategies to foster attitude and behavior change. Despite their serious intent, these games prioritize enjoyment and engagement, leveraging the immersive nature of video games to captivate players and sustain their interest [23]. They often include informative content relevant to the issue, offering players opportunities for learning and reflection during gameplay [24]. These games may use procedural rhetoric through interactive gameplay dynamics to convey persuasive messages effectively [22] and can use exocentric or endocentric approaches to engage and motivate players [24]. Successful play encourages players to develop a deeper understanding beyond the game, potentially influencing their behavior afterward. However, the effectiveness of these games hinges on game features, player characteristics, use context [25], and the balance of persuasive strategies [21].

Persuasive video games are a powerful tool for driving social change [26]. They are studied extensively for their ability to promote attitude and behavior changes across various issues such as politics, society, environment, and health [21]. They address social problems such as attitudes toward homelessness [27], humanitarian aid willingness [28], and workloads [20] while also tackling health issues such as physical activity, nutrition, and disease management [21]. Specific games target smoking cessation [29] and medication adherence [30]. Despite their widespread application, research shows mixed results on their impact, with some studies confirming effectiveness and others not, leading to inconsistent findings [20,21,31]. A systematic review indicated positive or partially positive outcomes [21], but effectiveness varied across studies, underscoring the need for rigorous evaluation across diverse contexts and user profiles [20]. Current research often focuses on overall impacts rather than specific game or user characteristics [31], necessitating further studies on how game features affect effectiveness, especially the effect of user enjoyment [32].

Persuasive Video Games for Agriculture

Most serious video games for agriculture have been designed as educational tools [33]. Explicitly designed persuasive games are scarce; only a few studies have evaluated their persuasive effects. The educational games cover various aspects of agriculture and vegetable cultivation, including genres such as role-playing games and farm simulators. Examples include "Farmtasia" [34], "Little Botany" [35], "Herbopolis" [36], and Serious Game for Agroecology Learning [37]. They offer valuable learning experiences in farming practices, sustainable agriculture, herbal medicine agriculture, and ecological awareness, leading to knowledge enhancement. However, there need to be more games specifically focusing on home gardening [33]. Another group of studies aims to raise awareness of agriculture-related issues. "AgriVillage" [38] focuses on environmental concerns in agriculture, including the effects of fertilizers and deforestation on water sources and weather patterns. "RebEarth" [39] promotes awareness of hydroponics. Furthermore, individuals who are agricultural novices are often overlooked by current serious video games for agriculture [33].

Among the studies evaluating the persuasive effects of serious agricultural games, an example is "Game of Piglets" [40], a virtual pig farm that allows students to practice external biosecurity and farrowing management procedures. This simulation and adventure game emphasizes core competencies such as farrowing aid, identifying unwell sows after farrowing, maintaining aseptic conditions during surgeries, and ensuring an optimal piglet environment. These competencies are developed through tasks closely resembling real-life scenarios in pig farms. Evaluation results indicated improved perceived self-efficacy among players. Another instance is "MahindiMaster" [41], a serious game simulating crop yields based on farmers' choices from various input options. These yields are customized using crop model outputs leveraging plot-level soil samples and historical weather data. The game allows farmers to experiment with 3 different fertilizers. Evaluation outcomes revealed positive shifts in players' beliefs and fertilizer use on their crops. While the evaluation results indicate positive impacts and the viability of using agriculture-focused serious video games to change players' attitudes and behaviors, the persuasive potential of agricultural video games remains unexplored. It prompts questions about their effectiveness in promoting home gardening among novices.

Inspired by the need for innovative strategies to promote home gardening among novices and recognizing the potential of persuasive video games to influence user knowledge, attitudes, and behaviors, this study addressed existing research gaps in persuasive serious games for home gardening. In total, 3 primary research questions guided our investigation. The first 2 questions focused on the effect of a persuasive game on promoting home gardening among novices:

- Can a persuasive video game encourage novices to engage in home gardening?
- Does a persuasive video game have a more significant impact on changing novices' knowledge, attitudes, self-efficacy, and intentions than a traditional course?

In addition, our investigation was guided by the following primary research question that sought to understand how the enjoyment derived from the game influenced its effectiveness:

- How does the enjoyment experienced by players of an agriculture-focused persuasive video game relate to the game's effects?

Objectives and Hypotheses

This study aimed to design and develop a persuasive video game promoting home gardening among novices called Farmily (Farm+Family) and evaluate its effects.

On the basis of the potential of persuasive video games to promote home gardening, we formulated the hypotheses that Farmily players would exhibit superior outcomes in the following areas after playing (hypothesis 1): home gardening knowledge (hypothesis 1A), attitudes toward home gardens (hypothesis 1B), self-efficacy in home gardening (hypothesis 1C), and intention to start a home garden (hypothesis 1D).

In addition, based on the idea that a persuasive video game offers more effective elements for promoting home gardening among novices compared to a traditional course, we established the hypotheses that Farmily players would demonstrate superior outcomes compared to traditional course attendees in the following areas (hypothesis 2): home gardening knowledge (hypothesis 2A), attitudes toward home gardens (hypothesis 2B), self-efficacy in home gardening (hypothesis 2C), and intention to start a home garden (hypothesis 2D).

Finally, based on the premise that game features, particularly user enjoyment, significantly relate to game outcomes, we formulated the hypotheses that the player's gaming enjoyment is positively related to (hypothesis 3) home gardening knowledge (hypothesis 3A), attitudes toward home gardens (hypothesis 3B), self-efficacy in home gardening (hypothesis 3C), and intention to start a home garden (hypothesis 3D).

Methods

Study Design

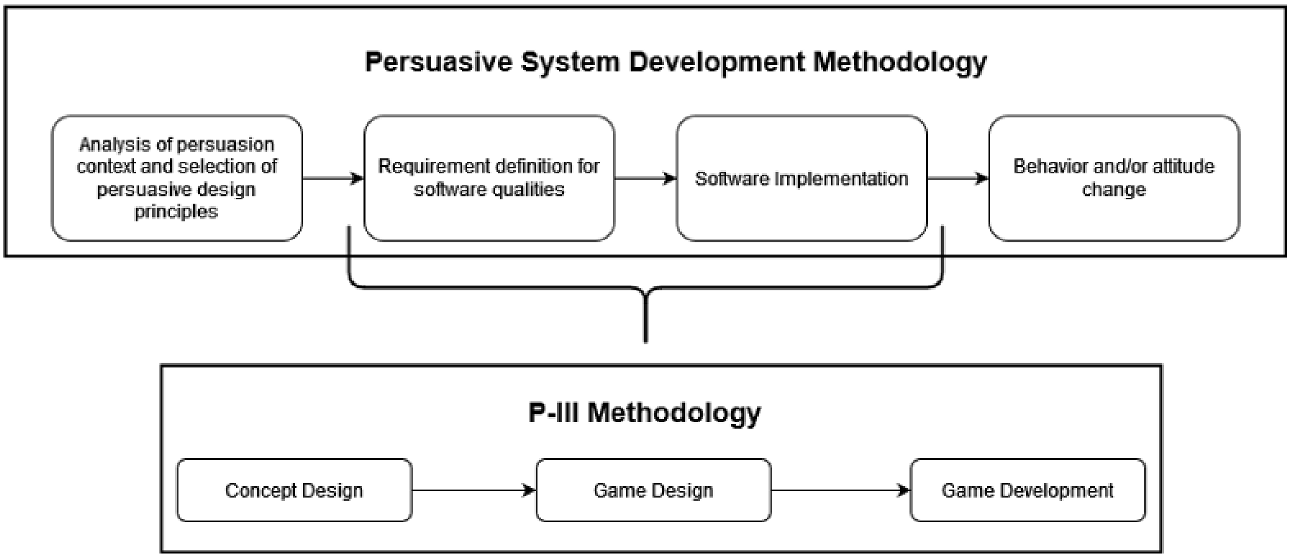
This study was divided into 2 main parts. The first part focused on designing and developing the video game Farmily. The second part involved conducting a randomized controlled trial (RCT) to assess the impact of Farmily and validate the formulated hypotheses. In the following sections, we provide a detailed description of the Farmily video game.

Farmily Video Game

Design and Development

To create Farmily, we integrated 2 methodologies (Figure 1): the persuasive system design (PSD) methodology [42] and the player-centered, iterative, interdisciplinary, and integrated (P-III) methodology [43]. The PSD methodology guides the design of persuasive systems [42], and accordingly, we followed four key steps aligned with persuasive principles. In step 1, we analyzed the context of persuasion and selected appropriate principles and techniques for Farmily. Steps 2 and 3 focused on defining and developing the game, whereas step 4 involved experimentation to evaluate its effectiveness.

Figure 1. Methodology for the design and development of Farmily. P-III: player-centered, iterative, interdisciplinary, and integrated.



During steps 2 and 3 of PSD, given the need to develop a specific persuasive system—a persuasive video game—we used the P-III methodology [43] tailored for serious game design. The P-III methodology emphasizes player-centered design, iterative development, and interdisciplinary collaboration. Following the P-III methodology, we engaged experts in human-computer interaction, agriculture, horticulture, education, and software engineering. In stage 1, the concept design phase, extensive research on home gardening informed the definition of user tasks through co-design sessions and expert consultations. Progressing to stage 2, the video game design phase, we used storyboards, focus groups, and paper prototypes, refining the game iteratively based on continuous feedback. Stage 3, the development phase, involved iterative prototyping and testing. We conducted 21 work sessions, incorporating multiple refinement cycles and expert testing sessions.

Theoretical Base and Persuasive Techniques

Farmily is based on the social cognitive theory by Bandura [44,45], which provides valuable insights into individual learning, development, and behavior maintenance in social contexts. This theory highlights the dynamic interplay of personal factors (cognition, beliefs, skills, and affect), environmental factors (social norms, institutions, and cultural influences), and behavioral factors (observable actions and responses) [44,45]. This theory proposes that individuals shape their environments through behavior, and vice versa, environments influence behavior and cognition. Learning, as per this theory, happens through direct experience and observing others [45]. Actions and outcomes offer crucial feedback on behavior appropriateness, with rewards encouraging desirable behaviors and punishments discouraging them. Furthermore, this theory emphasizes self-regulatory processes, which involve activating and sustaining behaviors, cognitions, and emotions directed toward achieving goals. It also highlights various motivational processes, including goals, progress evaluations, outcome expectations, values, and social comparisons. A key concept in the theory by Bandura [44,45] is self-efficacy, which pertains to one's belief in their capacity to achieve desired outcomes through specific actions. Self-efficacy significantly impacts motivation, goal setting, and persistence, with individuals who possess high self-efficacy more inclined to pursue challenging objectives and persist despite obstacles.

Drawing from the social cognitive theory by Bandura [44,45] and the PSD principles by Oinas-Kukkonen and Harjumaa [42] and Fogg [46], Farmily integrates persuasion techniques. Farmily's persuasion techniques are classified into primary support, dialogue support, credibility support, and social support, as detailed in Table 1. These strategies aim to engage players' intellect and motivation, guide behavior, and encourage their interest in starting home gardens. They have proven effective in persuasive video games such as "PowerHouse" [47] and "Smoke?" [21,48].

Table 1. Persuasive techniques of Farmily.

Name	Description
Primary support techniques	
Reduction	The game incorporates a virtual environment for family garden ownership and management that simplifies home garden maintenance.
Tunneling	The game offers a structured progression system with increasing complexity across levels, introducing new variables for users to understand.
Tailoring	Tailored knowledge by dividing levels into 3 difficulty tiers: easy, medium, and hard. Users unlock progressively more advanced and comprehensive information tailored to their learning needs as they advance.
Self-monitoring	The score and recommendation system helps players monitor their performance and receive guidance for improvement. This tangible evidence motivates them to persist in learning and practicing.
Simulation	Includes a home garden simulation mirroring real-life scenarios, helping players grasp cause-effect relationships between their actions and virtual outcomes.
Rehearsal	Provides a platform for simulating various home garden–related behaviors and activities. Through repetitive practice and self-monitoring, players enhance their self-efficacy and proficiency in gardening skills.
Dialogue support techniques	
Praise	Implements a praise system that offers positive reinforcement to players upon completing tasks or levels. Through motivational messages and rewards, the game aims to encourage continued gameplay.
Rewards	Incorporates rewards and motivators such as unlockable content and points to encourage continued play. For instance, users earn a trophy for completing the story mode and stars for each level, enhancing motivation and engagement.
Suggestion	The in-game recommendation system guides players in enhancing their gardening performance and achieving higher scores. Tailored suggestions empower players to improve and meet game objectives, boosting their self-efficacy and intention to apply gained knowledge to real-world gardening.
Liking	Meticulously recreates a home garden's visual and sensory experience, immersing players in a realistic virtual environment. This attention to detail enhances players' liking and connection to the game.
Credibility support techniques	
Trustworthiness	The game's gardening knowledge base is built on accurate and authentic information from reliable sources, such as books and manuals recognized and developed by Mexico's agriculture department and scientific papers sourced from reliable journals, ensuring its trustworthiness and reliability.
Surface credibility	The game's visual design and user interface are meticulously crafted to offer players a professional and credible experience.
Social support	
Social comparison	Including a leaderboard in the video game enables players to compare their scores and monitor the progress of other users. This feature fosters a competitive and social atmosphere that encourages engagement and interaction among players.
Competition	The video game allows players to compete against each other, striving to reach the top of the leaderboard and achieve the highest score. This competitive element enhances motivation and drives players to improve their performance.
Recognition	The game recognizes the top 5 players through the leaderboard, highlighting their achievements and establishing a sense of accomplishment. This acknowledgment rewards players and motivates them to continue their efforts and aim for higher rankings.

Description

Farmily is a 3D simulation single-player game available on desktop and Android platforms. It was developed using the Unity game engine (version 2021.3.8f1; Unity Technologies). The players guide and assist a virtual family in managing a home garden. The target users are individuals aged between 20 and 50 years who are novices in home gardening, chosen for suitability. This age range aligns with Mexico's economically active population (ages of 15 to 64 years) [49] and the gaming demographic [50], capturing a significant portion of gamers interested in home gardening. Variations in the age range do not significantly impact the study's relevant aspects. The game aims to convey messages such as gaining knowledge about home gardening, recognizing its importance and the enjoyment it generates for individuals and their families, understanding commitments and trade-offs, realizing accessibility and rewards, and acknowledging the multidisciplinary knowledge and risks involved.

Farmily consists of 3 sections: the "Main menu" (Figure 2), "Education" (Figure 3), and "Simulation" (Figure 4). In the main menu, players can register, access the leaderboard, adjust sound settings, and start a gaming session. The educational section is seamlessly integrated into the gameplay, offering tutorials on game mechanics and guidance on executing home gardening tasks both in the game and in real life. The simulation section encompasses the home garden simulation and all associated gameplay mechanics.

Figure 2. Screenshots of the main menu (top) and level selection (bottom). The game is presented in Spanish.

The game's difficulty gradually increases to enhance its educational and persuasive aspects. Farmily offers 20 levels divided into beginner (levels 1-8), medium (levels 9-14), and advanced (levels 15-20) tiers (Figure 2, bottom). Each tier presents challenges in crop yield, finances, and simulation complexity. The level goals include growing and harvesting crops, ensuring family participation, practicing sustainability, maximizing savings, and providing a comprehensive understanding of home gardening implications and benefits.

The game educates players on 12 recommended crops for daily consumption [51], empowering informed decisions and gardening optimization. In Farmily, each level follows a systematic progression—players start by preparing compost, clearing undergrowth, and watering and then plow an area spanning 1 to 40 m^2. Next, they plant seeds in nurseries, nurture them, and transplant them. Subsequently, players water, fertilize, and apply pesticides as needed to ensure crop health for harvesting.

Figure 3. Screenshots of the education section; home gardening phases (top) and tutorial menu (bottom). The game is presented in Spanish.

The game provides indicators for garden health, including a progress bar for the current month and phase, meters for overall garden health and sustainability, and progress bars for vegetable quantity (Figure 4, top). Detailed plot information, such as humidity and macronutrient levels, is also available. After harvesting all crops, players receive a rating from 1 to 5 stars based on factors such as harvested vegetables, family involvement, sustainable practices, financial economy, and self-sufficiency (Figure 4, bottom). These ratings unlock new scenarios, motivating players to improve and enhancing replay value. A home garden manual is a central hub for information on crops, fertilizers, pesticides, and game mechanics tutorials (Figure 3, top). Personalized recommendations based on player actions help optimize gardening strategies (Figure 4, bottom). A leaderboard fosters competitive excitement by allowing players to compare scores and aim for the top spot (Figure 2, top).

Figure 4. Screenshots of the simulation section; cultivation (top) and level evaluation with recommendations (bottom). The game is presented in Spanish.

Evaluation

Overview

In 2023, an RCT was conducted at a Mexican university, with written permission from the school administrators to conduct the trial at the school facilities.

Participant Recruitment

We used a convenience sampling approach by inviting 180 students from 9 academic groups to participate based on their availability and interest in the study. Participants were invited without previous information about their backgrounds or characteristics, ensuring a diverse representation across different scholar groups. On the day of the experiment, we verified that participants who agreed to take part met the inclusion criteria, which were (1) being aged between 20 and 50 years; (2) belonging to a family nucleus comprising at least 2 members; (3) having an interest in home gardening; (4) residing in Ciudad Guzmán, Jalisco, Mexico; and (5) having essential elementary-level reading and writing skills in Spanish. Exclusion criteria were applied to individuals who (1) currently or previously had a home garden, (2) had undergone a home gardening course, and (3) had vision or motor impairments that hindered their use of the video game on a PC.

Structure and Procedure

Individual participants were randomly assigned to either the control group (CG) or the experimental group (EG). Before the intervention (T_0), both groups completed 2 pretest questionnaires measuring their knowledge, attitude, self-efficacy, and intention to start a home garden (see the *Measured Variables* section). During the intervention, participants in the CG engaged in a 1.5-hour home gardening workshop facilitated by an expert in agriculture. The workshop provided participants with the same valuable information regarding the benefits, requirements, and activities associated with maintaining a home garden as that provided in the EG and Farmily. The expert presented the knowledge in a lecture format without hands-on experience with a home garden or interactive activities. This workshop mirrors the traditional way of fostering home gardening. In contrast, participants in the EG played Farmily for 1.5 hours. Finally, after the intervention period (T_1), the CG and EG participants completed the same questionnaires applied during the pretest. This posttest measurement served as a valuable indicator of any changes or improvements in these domains following the intervention. In addition to the posttest questionnaires, participants in the EG were also asked to provide feedback on their enjoyment of the video game using a questionnaire.

Measured Variables

The following variables were measured before (T_0) and after (T_1) the intervention: (1) knowledge about home gardening, (2) attitude toward home gardening, (3) self-efficacy in home gardening, and (4) intention to start a home garden. In addition, in T_1, we measured the users' enjoyment. Textbox 1 describes the measured variables. In the following paragraphs, we describe the methods used for their measurement.

Textbox 1. Description of the measured variables.

Knowledge about home gardening

- This is the comprehension level of a home garden's concepts and procedures.

Attitude toward home gardening

- Fishbein and Ajzen [52] define attitude as "an individual's positive or negative feelings towards realizing an objective behavior." In line with this, we defined *attitude toward home gardening* as "an individual's positive or negative thoughts about having a home garden."

Self-efficacy on home gardens

- Self-efficacy is "people's beliefs about their abilities to produce a certain level of performance that influences events that affect their lives" [53]. Given this concept, *self-efficacy in home gardening* was defined as "people's beliefs about their abilities to have a home garden with a certain level of performance that influences the events that affect their lives."

Intention to start a home garden

- Intention is "a person's motivation in the sense of their conscious plan or decision to exert effort to perform a behavior" [52]. On the basis of this concept, *intention to start a home garden* was defined as "the conscious motivation of a person to start a home garden."

Game enjoyment

- It is defined as "a person's perception and response as a result of using or anticipating using a product, system, or service" [54].

To evaluate participants' knowledge levels effectively, we developed a 23-item questionnaire with a home gardening expert, ensuring accuracy and relevance. The questionnaire items were strategically structured into 3 difficulty levels (easy, medium, and hard) designed to assess participants' understanding comprehensively. Each item provided 4 answer options, with only 1 correct answer. For example, sample questions included an easy query, such as "How many days does it take for a plant to germinate?"; a medium-level question, such as "Which of the following statements best describes the advantages of sustainable products?"; and a hard question, such as "What is a critical factor for achieving success in home gardening?"

To assess attitudes toward home gardens, self-efficacy in home gardening, and the intention to start a home garden, we developed a 14-item questionnaire. This questionnaire draws upon the attitude, self-efficacy, and intention scales from the Unified Theory of Acceptance and Use of Technology (UTAUT) [55]. The UTAUT questionnaire's validity has been established through rigorous testing and has been shown to account for 70% of the variance in intention [55]. Furthermore, the scale's reliability was demonstrated with a Cronbach α of 0.70 across all scales [55]. The attitude toward home gardens scale comprises 4 items, the self-efficacy scale includes 7 items, and the intention to start a home garden scale consists of 3 items. Each item is rated on a 5-point Likert scale ranging from "1—Totally disagree" to "5—Totally agree." Sample questions include "Are home gardens a beneficial idea?" (attitude toward home gardens), "Compared to other people, can I perform my home gardening activities correctly?" (self-efficacy), and "I plan to start a home garden in the next six months." (intention to start a home garden).

We used the EGameFlow questionnaire [56], a validated tool of 42 items, to assess user enjoyment. Comprehensive testing confirmed the questionnaire's validity, which was robust enough to account for 74% of the variance in learner enjoyment [56]. Moreover, the EGameFlow questionnaire exhibited strong reliability, boasting a Cronbach α of 0.942, signifying excellent internal consistency. In addition, it demonstrated good test-retest reliability [56]. The questionnaire items are categorized into 8 dimensions: concentration (6 items), clear goal (4 items), feedback (5 items), challenge (6 items), autonomy (3 items), immersion (7 items), social interaction (6 items), and knowledge improvement (5 items). Each item is rated on a 5-point Likert scale ranging from "1—Totally disagree" to "5—Totally agree."

Data Analysis

The data analysis was performed using the R software (version 4.3.0; R Foundation for Statistical Computing) on the Windows platform. A significance level of 5% (P<.05) was used to determine statistical significance. The normality of the numerical data was assessed using the Shapiro-Wilk test. Parametric tests were used for normally distributed data, while nonparametric tests were used for nonnormal data. Changes in knowledge within the same group were analyzed using a 1-tailed paired *t* test. A 2-tailed independent-sample *t* test was conducted to compare knowledge between groups. The Wilcoxon signed rank test was applied for questions with ordinal values, such as attitude, self-efficacy, intention, and user enjoyment. The 2-tailed paired *t* test was used to analyze differences from T_0 to T_1 in ordinal question scales within the same group assuming that the data followed a normal distribution. Alternatively, the Wilcoxon signed rank test was used. For comparisons between groups and when the data followed a normal distribution, a 2-tailed independent-sample *t* test was used. In the absence of normal distribution, a Mann-Whitney *U* test was conducted. The internal consistency of the user enjoyment scales was assessed using the Cronbach α. Finally, the bilateral Pearson correlation test was conducted to examine the relationships between user enjoyment and posttest scores on knowledge, attitude, self-efficacy, and intention to start a home garden.

Ethical Considerations

The institutional bioethics review board of the Centro de Investigación Científica y de Educación Superior de Ensenada approved all study procedures (BIOETICA_HUM_2023_03). Participants who expressed interest in joining the study provided written informed consent. The study ensured that all collected data were anonymized or de-identified to protect personal information, implementing strict protective measures for any data that could not be fully anonymized. Participation in the study was voluntary and uncompensated.

Results

Participant Characteristics

Of the 180 students and teachers approached, 50 (27.8%) agreed to participate in the study. The participants had an average age of 21.5 (SD 6.8) years, with 66% (33/50) being male and 34% (17/50) being female. Among the participants, the EG consisted of 25 individuals with an average age of 20.2 (SD 0.7) years, comprising 11 (44%) female and 14 (56%) male individuals. The CG also consisted of 25 participants with an average age of 22.7 (SD 9.5) years, including 6 (24%) female and 19 (76%) male individuals.

Knowledge of Home Gardening

The knowledge of participants (Table 2) in the EG significantly increased from T_0 to T_1 (mean difference 3.2, SD 0.18; t_{24}=4.26; *P*<.001). Similarly, the CG showed a significant increase in knowledge (mean difference 5, SD 0.3; t_{24}=6.78; *P*<.001). Comparing the differences in knowledge increases between groups from T_0 to T_1, no statistically significant differences were found (t_{24}=–1.71; *P*=.09). However, the CG had a slightly higher mean increase in knowledge (mean 5, SD 3.68) compared to the EG (mean 3.2, SD 3.75). Analyzing the results based on question difficulty, the EG showed a significant knowledge increase for easy (*P*=.02), medium (*P*=.03) and hard (*P*<.001) questions. Similarly, the CG showed a significant knowledge increase for easy (*P*<.001) medium (*P*=.001) and hard (*P*<.001) questions. There were no significant differences between the groups in the knowledge increase for the questions of easy (t_{24}=–0.68; *P*=.09) and medium (t_{24}=–0.73; *P*=.09) difficulty. However, a significant difference was observed for hard questions (t_{24}=0.64; *P*=.005), with the CG achieving a higher mean score (mean 5.76, SD 1.27) than the EG (mean 4.76, SD 1.39).

Table 2. Results for knowledge on home gardening.

Question	CG[a]			EG[b]			Intergroups, *P* value[c]
	T_0[d] scores, mean (SD)	T_1[e] scores, mean (SD)	*P* value[c]	T_0 scores, mean (SD)	T_1 scores, mean (SD)	*P* value[c]	
Easy (n=8)	4.28 (1.40)	5.68 (1.03)	<.001	4.04 (1.21)	5.08 (2.06)	.02	.09
Medium (n=8)	3.64 (1.38)	4.80 (1.55)	.001	3.32 (1.77)	4.16 (1.62)	.03	.09
Hard (n=7)	3.08 (1.32)	5.76 (1.27)	<.001	3.44 (1.23)	4.76 (1.39)	<.001	.005
All (n=23)	11.12 (2.85)	16.12 (3.15)	<.001	10.88 (3.22)	14.08 (3.40)	<.001	.09

[a]CG: control group.

[b]EG: experimental group.

[c]*P*<.05 was considered statistically significant.

[d]T_0: before the intervention.

[e]T_1: after the intervention.

Attitude Toward Home Gardening

The attitude toward home gardening of participants (Table 3) in the EG significantly increased from T_0 to T_1 (mean difference 0.8, SD –0.01; z_{24}=2.98; *P*=.003). When analyzing the results by question, significant differences (*P*<.05) were found for all items. In contrast, there was no significant difference in the attitude toward home gardening of the CG from T_0 to T_1 (mean difference –0.05, SD –0.09; z_{24}=−0.54; *P*=.59) including individual questions. A significant difference was observed in attitude toward home gardening increase between groups from T_0 to T_1 (z_{24}=2.73; *P*=.006). Furthermore, the EG exhibited a higher mean increase in attitude toward home gardening (mean 0.8, SD 1.21) than the CG (mean –0.05, SD 1.30).

Table 3. Results for attitude toward home gardening.

Question	CG[a]			EG[b]			Intergroups, *P* value[c]
	T_0[d] scores, mean (SD)	T_1[e] scores, mean (SD)	*P* value[c]	T_0 scores, mean (SD)	T_1 scores, mean (SD)	*P* value[c]	
1	4.44 (1.08)	4.52 (0.71)	.97	4.08 (1.19)	4.60 (1.07)	.04	—[f]
2	4.20 (1.12)	4.40 (0.76)	.58	3.72 (1.21)	4.60 (1.12)	.009	—
3	4.32 (0.94)	4.20 (0.87)	.56	3.56 (1.23)	4.48 (1.12)	.005	—
4	4.32 (1.07)	3.96 (1.17)	.22	3.36 (1.22)	4.24 (1.16)	.008	—
All	4.32 (0.93)	4.27 (0.84)	.59	3.68 (1.10)	4.48 (1.09)	.003	.006

[a]CG: control group.
[b]EG: experimental group.
[c]*P*<.05 was considered statistically significant.
[d]T_0: before the intervention.
[e]T_1: after the intervention.
[f]Not applicable.

Self-Efficacy in Home Gardening

The self-efficacy in home gardening of participants (Table 4) in the EG significantly increased from T_0 to T_1 (mean difference 0.75, SD –0.01; t_{24}=2.6; *P*=.02). When analyzing the results by question, all items showed significant differences (*P*<.05). In contrast, the CG did not show a significant increase in self-efficacy in home gardening from T_0 to T_1 (mean difference –0.09, SD –0.17; t_{24}=–0.36; *P*=.72) or in individual questions. Comparing the differences in self-efficacy in home gardening increase between the 2 groups revealed a significant difference (t_{24}=2.21; *P*=.03). Furthermore, the EG exhibited a higher mean increase in self-efficacy in home gardening (mean 0.75, SD 1.45) than the CG (mean –0.09, SD 1.25).

Table 4. Results for self-efficacy in home gardening.

Question	CG[a]			EG[b]			Intergroups, *P* value[c]
	T_0[d] scores, mean (SD)	T_1[e] scores, mean (SD)	*P* value[c]	T_0 scores, mean (SD)	T_1 scores, mean (SD)	*P* value[c]	
1	3.68 (1.25)	3.84 (0.75)	.62	3.48 (0.87)	4.24 (0.97)	.02	—[f]
2	3.60 (1.19)	3.60 (0.82)	.94	3.32 (0.90)	4.12 (0.97)	.02	—
3	4.04 (1.14)	3.76 (0.88)	.22	3.52 (1.08)	4.32 (0.94)	.02	—
4	4.00 (0.91)	3.76 (0.88)	.45	3.56 (1.16)	4.20 (0.87)	.03	—
5	3.72 (1.02)	3.52 (1.00)	.53	3.60 (0.96)	4.24 (0.97)	.046	—
6	3.64 (1.04)	3.64 (0.99)	.89	3.44 (1.23)	4.28 (1.02)	.02	—
7	3.76 (1.01)	3.68 (0.90)	.78	3.44 (1.16)	4. 24 (0.88)	.03	—
All	3.78 (0.93)	3.69 (0.76)	.72	3.48 (0.87)	4.23 (0.88)	.02	.03

[a]CG: control group.
[b]EG: experimental group.
[c]*P*<.05 was considered statistically significant.
[d]T_0: before the intervention.
[e]T_1: after the intervention.
[f]Not applicable.

Intention to Start a Home Garden

The intention to start a home garden of participants (Table 5) in the EG significantly increased from T_0 to T_1 (mean difference 2.02, SD 0; z_{24}=4.15; *P*<.001). When analyzing the results by question, all items showed significant differences (*P*<.05). In contrast, the CG did not show a significant increase in intention to start a home garden from T_0 to T_1 (mean difference –0.18, SD 1.73; t_{24}=–0.64; *P*=.53) or in individual questions. Comparing the differences in intention to start a home garden increase between the 2 groups revealed a significant difference (t_{24}=–5.33; *P*<.001). Furthermore, the EG exhibited a higher mean increase in intention to start a home garden (mean 2.01, SD 1.46) than the CG (mean –0.19, SD 1.46).

Table 5. Results for intention to start a home garden.

Question	CG[a]			EG[b]			Intergroups, *P* value[c]
	T_0[d] scores, mean (SD)	T_1[e] scores, mean (SD)	*P* value[c]	T_0 scores, mean (SD)	T_1 scores, mean (SD)	*P* value[c]	
1	2.96 (0.84)	2.76 (1.42)	.52	2.00 (1.08)	4.04 (0.89)	<.001	—[f]
2	3.04 (0.84)	2.72 (1.40)	.32	1.80 (0.91)	3.88 (1.09)	<.001	—
3	3.04 (0.98)	3.00 (1.44)	.88	1.88 (1.05)	3.80 (1.15)	<.001	—
All	3.01 (2.34)	2.83 (4.07)	.58	1.89 (0.95)	3.91 (0.95)	<.001	<.001

[a]CG: control group.
[b]EG: experimental group.
[c]*P*<.05 was considered statistically significant.
[d]T_0: before the intervention.
[e]T_1: after the intervention.
[f]Not applicable.

User Enjoyment

Results showed (Table 6) a significant difference between the participants' answers and the neutral value in all categories: concentration (z_{24}=4.11; *P*<.001; Cronbach α=0.90), goal clarity (z_{24}=4.15; *P*<.001; Cronbach α=0.97), feedback (z_{24}=3.92; *P*<.001; Cronbach α=0.96), challenge (z_{24}=4.03; *P*<.001; Cronbach α=0.93), autonomy (z_{24}=3.89; *P*<.001; Cronbach α=0.89), immersion (z_{24}=3.83; *P*<.001; Cronbach α=0.93), social interaction (z_{24}=3.82; *P*<.001; Cronbach α=0.93), and knowledge improvement (z_{24}=3.92; *P*<.001; Cronbach α=0.96). All categories exhibited a Cronbach α of >0.70, which indicates a high internal consistency in the questionnaire [57]. These results state that users positively evaluated the video game in all categories.

Table 6. Results for user enjoyment.

Measure	Questions, N	T_1[a] score, mean (SD)	Neutral value	*P* value[b]	Range	Cronbach α
Concentration	6	4.45 (0.82)	3	<.001	1.67-5	0.90
Goal clarity	6	6.77 (0.81)	3	<.001	1-5	0.97
Feedback	5	4.60 (0.80)	3	<.001	1-5	0.96
Challenge	6	4.52 (0.79)	3	<.001	1.67-5	0.93
Autonomy	3	4.48 (0.84)	3	<.001	1-5	0.89
Immersion	7	4.34 (0.90)	3	<.001	1-5	0.93
Social interaction	6	4.26 (0.91)	3	<.001	1-5	0.93
Knowledge improvement	5	4.61 (0.81)	3	<.001	1-5	0.96

[a]T_1: after the intervention.
[b]*P*<.05 was considered statistically significant.

Relationship Between User Enjoyment and Efficacy Outcomes

Results showed (Table 7) no statically significant relationship between user enjoyment scores and posttest knowledge (EGameFlow scale: r_{23}=0.26; *P*=.20) and intention (EGameFlow scale: r_{23}=0.06; *P*=.77) scores. In addition, the results showed a positive statically significant relationship between user enjoyment scores and posttest attitude (EGameFlow scale: r_{23}=0.72; *P*<.001) and self-efficacy (EGameFlow scale: r_{23}=0.67; *P*<.001) scores.

Table 7. Correlations of user enjoyment scores with posttest efficacy results.

Category	*r*	*P* value[a] (bilateral)
Player enjoyment and knowledge	0.26	.20
Player enjoyment and attitude	0.72	<.001
Player enjoyment and self-efficacy	0.67	<.001
Player enjoyment and intention	0.06	.77

[a]$P<.05$ was considered statistically significant.

Hypothesis Evaluation

On the basis of the aforementioned results, in this subsection, we summarize the validation of each hypothesis using appropriate statistical methods. In addition, Table 8 summarizes the statistical significance, indicating whether each hypothesis was confirmed based on these results.

Table 8. Summary of hypothesis evaluation.

Hypothesis	Statistical analysis results	*P* value	Hypothesis confirmed?
1A	$t_{24}=4.26$	<.001	Yes
1B	$z_{24}=2.98$	.003	Yes
1C	$t_{24}=2.6$	.02	Yes
1D	$t_{24}=2.21$	.03	Yes
2A	$t_{24}=–1.71$	.09	No
2B	$z_{24}=2.73$	.006	Yes
2C	$t_{24}=2.21$	.03	Yes
2D	$t_{24}=–5.33$	<.001	Yes
3A	$r_{23}=0.26$	.20	No
3B	$r_{23}=0.72$	<.001	Yes
3C	$r_{23}=0.67$	<.001	Yes
3D	$r_{23}=0.06$	.77	No

Discussion

Principal Findings

Overview

The findings of this study support the effectiveness of Farmily in enhancing knowledge of home gardening, promoting a positive attitude toward home gardening, increasing self-efficacy in home gardening, and fostering the intention to start a home garden among novice individuals aged 20 to 50 years. In addition, Farmily demonstrated comparable effectiveness in knowledge transfer and a more significant impact on the other variables than a home gardening workshop facilitated by an expert in agriculture. Finally, Farmily yielded favorable results in user enjoyment, and we identified a significant positive relationship between user enjoyment and attitudes and self-efficacy and no relationship with knowledge and intention. The following sections comprehensively analyze each hypothesis, presenting the corresponding findings in detail.

Knowledge of Home Gardening

The EG demonstrated improved knowledge of home gardening after the treatment. Thus, our data support hypothesis 1A (confirmed). These results are consistent with those of previous studies on the effectiveness of video games in agricultural education [36-38,58]. Similarly, the CG also showed improved knowledge after the treatment. Regarding hypothesis 2A (rejected), there were no significant differences between the groups regarding the increase in knowledge overall and for easy and medium questions. However, a significant difference was found for difficult questions. These findings suggest that persuasive video games can be comparably effective to expert-led workshops in certain aspects but they may have limitations in more complex scenarios. It is important to acknowledge that workshops offer the advantage of direct interaction with experts, providing specific answers and in-depth knowledge. On the other hand, video games are constrained by the information provided within them. However, our video game demonstrated its potential effectiveness. Simulation-based experimentation has been successful in various fields, enabling competence development through trial-and-error experiences [59]. The discovery learning theory supports that knowledge is constructed through an iterative discovery process facilitated by educators [60]. Our results indicate that a serious persuasive video game can be valuable for reinforcing knowledge and complementing traditional educational techniques, particularly in noncomplex knowledge activities.

Attitude Toward Home Gardening

Regarding hypothesis 1B (confirmed), the participants in the EG demonstrated a significant improvement in their attitude toward home gardens. These results are consistent with findings of studies such as that on MahindiMaster [41], supporting our results. In contrast, the CG showed a nonsignificant decrease in attitude. Regarding hypothesis 2B (confirmed), the EG exhibited a higher mean increase in attitude than the CG. The divergence in results can be attributed to the unique firsthand experimentation experience provided by the simulation environment in Farmily. Actively engaging in virtual home gardening activities likely influenced participants' perception of the garden's interest, importance, and enjoyment, leading to a positive attitude toward real-life gardening. This finding aligns with the literature suggesting that providing new information can change attitudes toward an object [61]. The persuasive messages embedded in the video game may have also contributed to the favorable attitude change as persuasive technologies have been shown to influence attitudes and behaviors effectively [62,63].

Self-Efficacy in Home Gardening

Regarding hypothesis 1C (confirmed), participants in the EG demonstrated a significant improvement in their self-efficacy. These results are consistent with those of previous studies reporting positive effects on self-efficacy after intervention in video games [64]. Similarly, video games such as "Game of Piglets" [40] have increased players' self-efficacy in agriculture-related topics following interventions, supporting our expectations for our video game. In contrast, the CG showed a nonsignificant decrease. Regarding hypothesis 2C (accepted), the EG exhibited a higher mean increase in self-efficacy than the CG. As described by Bandura [53], self-efficacy is an integral part of the self-system, encompassing a person's attitude, ability, and cognitive abilities. According to Bandura [44], individuals develop self-efficacy beliefs through previous performance outcomes, vicarious experiences, social persuasion, and emotional and physiological states. The direct experience of EG users with the simulation of the home garden had a profound impact on their perception of their capabilities. The feedback and recommendation systems within the video game allowed players to analyze their previous performances, reinforcing their sense of competence. In addition, the recognition received when completing a level served as a form of social persuasion, offering congratulations and encouragement. In contrast, the workshop provided only vicarious experiences, lacking the opportunity for hands-on experimentation, which may explain its limited impact on self-efficacy.

Intention to Start a Home Garden

Regarding hypothesis 1D (confirmed), participants in the EG showed a significant improvement in their intention. These results were expected and explained through the Unified Theory of Behavior proposed by Jaccard et al [65], which states that 2 key factors influencing an individual's intention to engage in a behavior are their attitude toward the behavior and their self-efficacy. In the case of the EG, attitude and self-efficacy showed significant increases, which directly contributed to enhancing their intention. Conversely, the CG did not exhibit significant differences in these variables, resulting in no notable change in intention. Regarding hypothesis 2D (confirmed), the EG exhibited a higher mean increase in intention than the CG. These results can also be explained by the UTAUT proposed by Venkatesh et al [55]. According to this theory, participants in the EG may have realized that the perceived complexity and effort associated with engaging in a home garden were lower than initially expected (expectation of effort). Moreover, their direct experimentation with the video game enabled them to discover their capabilities and potential for performing well in a real garden (performance expectancy). Furthermore, it is plausible that the video game impacted other variables such as social norms and emotions by including persuasive messages. However, it is important to note that no specific data or evidence regarding these phenomena were collected in the study. We did not identify previous studies on the intention to start a home garden, but given that a change in intention leads to a small to medium change in behavior [66], several studies that achieved a positive behavior change [21] may have successfully influenced intention. However, there is a need for more data to corroborate that assumption.

User Enjoyment

Farmily received highly positive ratings across various user enjoyment categories, including concentration, clarity of goals, feedback, challenge, autonomy, immersion, social interaction, and knowledge increase. Participants expressed satisfaction with the gameplay, level of challenge, clarity of objectives, and overall video game mechanics. The success of Farmily in providing positive user enjoyment can be attributed to the mixed methodology used during its design. The iterative development process and continuous feedback from home garden experts ensured the creation of a video game with a robust knowledge base. The interdisciplinary composition of the team, including a video game designer, an interaction expert, and an education expert, played a significant role in crafting clear, immersive, and enjoyable video game mechanics, pedagogical techniques, and interactive elements within the final product.

Relationship Between User Enjoyment and Efficacy Outcomes

Regarding hypotheses 3B (confirmed) and 3C (confirmed), the findings revealed a positive statistical relationship between user enjoyment scores and attitude and self-efficacy posttest scores. These results suggest that players who enjoy the game also have a positive attitude and can master the game. These results are consistent with those of previous studies on the relationship between enjoyment and attitudes (eg, environmental sustainability attitude [67] and attitude toward real-life refugees [68]) and the relationship between enjoyment and self-efficacy (eg, genocide awareness [69] and drinking refusal self-efficacy [70]). In addition, regarding hypotheses 3A (rejected) and 3D (rejected), there was no statistically significant correlation between user enjoyment scores and posttest scores measuring both knowledge and intention, suggesting that enjoyment does not appear to be associated with knowledge acquisition or the propensity to engage in home gardening. These results are particularly surprising given the contrast with previous research examining the connection between enjoyment and learning outcomes in serious gaming contexts, as evidenced by studies investigating food knowledge [71] or mathematical proficiency [72]. However, it is worth noting that some studies have also failed to establish a relationship between enjoyment and learning [73]. Furthermore, these findings diverge from those of studies exploring the relationship with behavioral intentions, such as those examining intentions related to reducing alcohol consumption [70].

Limitations and Future Work

This study's findings should be cautiously interpreted due to the small sample size (n=50). Future research would benefit from a larger sample to enhance statistical reliability. Despite this limitation, the results provide valuable insights into video games as a tool for promoting home gardening. The participant pool, composed solely of students and teachers, may not fully represent the broader adult population aged 20 to 50 years. Moreover, participants reported high previous experience with video games and technology, which may not represent the general population. The short treatment and evaluation period (90 minutes) only captures immediate effects, not longer-term impacts. In addition, the knowledge questionnaire was developed specifically for this study in collaboration with a home garden expert, potentially influencing the results. Finally, leveraging a serious video game for promoting home gardening in areas of a low socioeconomic status can empower individuals with practical skills, foster sustainability, and build community resilience. However, it must be implemented with careful consideration of literacy rates, technological access, cultural relevance, and ongoing engagement strategies to maximize its impact. An enhanced version of Farmily is planned, incorporating elements such as climate simulation, disease management, and crop rotation and association. Future investigations are underway to use a more diverse sample and extend follow-up periods, thereby enabling comprehensive analysis. Furthermore, these efforts seek to deepen our understanding of the intricate relationship between enjoyment and persuasive outcomes within gaming environments.

Conclusions

Farmily, a 3D simulation single-player serious persuasive video game, was developed to educate and promote home gardening in novice individuals. The evaluation study demonstrated Farmily's effectiveness in improving participants' knowledge, attitude, self-efficacy, and intention to start their home gardens. Farmily demonstrated similar effectiveness in knowledge transfer to that of a traditional promotion workshop and had a more significant impact on other variables. The findings highlight Farmily's potential to empower individuals and promote sustainable practices by teaching the required home gardening knowledge and persuading them to start a real home garden. Furthermore, player enjoyment substantially correlates with user attitudes and self-efficacy, yet no discernible association was observed between enjoyment and knowledge and intention. Future research could expand on the impact of Farmily by conducting studies with more extensive and diverse samples over an extended time frame, including a follow-up period. These studies would allow for a deeper understanding of the medium- and long-term effects of Farmily and the sustainability of its impact over time. Moreover, additional research endeavors have the potential to advance our knowledge of the intricate interplay between enjoyment and persuasive outcomes in gaming contexts.

Acknowledgments

The authors thank the National Council of Science and Technology of Mexico for their support (scholarship 806581 awarded to the first author) and the teachers and students who participated in this study. They appreciate the support and cooperation of the authorities at the Instituto Tecnológico de Ciudad Guzmán. Special thanks to engineers Enriqueta Sánchez Bernal, Teresa Licea González, and Dr Elia Cruz Crespo for their valuable contributions to designing and reviewing Farmily's agricultural content. The authors also acknowledge Enriqueta Sánchez Bernal for conducting the home gardening workshop during the evaluation.

Authors' Contributions

All authors contributed to the study's conception and design. CAGAC, IEEC, and RMMN performed material preparation and data collection and analysis. The first draft of the manuscript was written by CAGAC, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2571 KB - games_v12i1e60771_app1.pdf]

References

1. Niñez V. Household gardens: theoretical and policy considerations. Agric Syst 1987;23(3):167-186. [doi: 10.1016/0308-521X(87)90064-3]
2. Galhena DH, Freed R, Maredia KM. Home gardens: a promising approach to enhance household food security and wellbeing. Agric Food Secur 2013 May 31;2(1):8. [doi: 10.1186/2048-7010-2-8]
3. Some basic facts about home gardens. Food and Agriculture Organization of the United Nations (FAO). Rome, Italy: FAO; 2004. URL: https://www.fao.org/4/y5112e/y5112e03.htm#bm3 [accessed 2024-01-05]
4. Castañeda-Navarrete J. Rural urbanisation and home gardening in southern Mexico: agrobiodiversity loss and alternative pathways. J Peasant Stud 2022 Jul 30;50(3):1227-1248. [doi: 10.1080/03066150.2022.2101100]
5. Saroinsong FB, Ismail Y, Gravitiani E, Sumantra K. Utilization of home gardens as a community empowerment-based edible landscape to combat stunting. IOP Conf Ser Earth Environ Sci 2021;940:012093. [doi: 10.1088/1755-1315/940/1/012093]
6. Eng S, Khun T, Jower S, Murro MJ. Healthy lifestyle through home gardening: the art of sharing. Am J Lifestyle Med 2019 Apr 15;13(4):347-350. [doi: 10.1177/1559827619842068] [Medline: 31285714]
7. Kohima N, Ansari MA, Panwar AS. Self-reliant in year-round vegetable production through kitchen garden in Indo-Gangetic plains. Indian J Agric Sci 2022;91(12):1773-1777. [doi: 10.56093/ijas.v91i12.120804]
8. Strategic framework 2022-31. Food and Agriculture Organization of the United Nations (FAO). 2021 Oct. URL: https://openknowledge.fao.org/server/api/core/bitstreams/29404c26-c71d-4982-a899-77bdb2937eef/content [accessed 2024-01-06]
9. Lal R. Home gardening and urban agriculture for advancing food and nutritional security in response to the COVID-19 pandemic. Food Secur 2020;12(4):871-876. [doi: 10.1007/s12571-020-01058-3] [Medline: 32837634]
10. DeMarco LW, Relf D, McDaniel A. Integrating gardening into the elementary school curriculum. Horttechnology 1999 Jan;9(2):276-281. [doi: 10.21273/HORTTECH.9.2.276]
11. Grebitus C. Small-scale urban agriculture: drivers of growing produce at home and in community gardens in Detroit. PLoS One 2021 Sep 07;16(9):e0256913. [doi: 10.1371/journal.pone.0256913] [Medline: 34492063]
12. Ayoni VD, Ramli NN, Shamsudin MN, Abdul Hadi AH. Non-growers' perspectives on home gardening: exploring for future attraction. J Regional City Plan 2023 Apr 11;34(1):16-34. [doi: 10.5614/jpwk.2023.34.1.2]
13. Bhatti M, Church A. Home, the culture of nature and meanings of gardens in late modernity. Hous Stud 2004 Jan;19(1):37-51. [doi: 10.1080/0267303042000152168]
14. González-Jacóme A. Del huerto a los jardines y vecindades: procesos de cambio en un agroecosistema de origen antiguo. In: Mariaca Méndez R, editor. El Huerto Familiar del Sureste de México. Lerma, México: Secretaría de Recursos Naturales y Protección Ambiental del Estado de Tabasco, El Colegio de la Frontera Sur; 2012:487-522 URL: https://es.scribd.com/document/140086536/EL-HUERTO-FAMILIAR-DEL-SURESTE-DE-MEXICO
15. Guía rápida para huertos urbanos familiares, sin salir de casa. Secretaria de Medio Ambiente del Gobierno de la Ciudad de México. 2020. URL: https://www.sedema.cdmx.gob.mx/storage/app/media/GuiaHuertosUrbanosFamiliares.pdf [accessed 2024-01-21]
16. Basarir A, Al Mansouri NM, Ahmed ZF. Householders attitude, preferences, and willingness to have home garden at time of pandemics. Horticulturae 2022 Jan 08;8(1):56. [doi: 10.3390/horticulturae8010056]
17. Muñoz-Rodríguez M, Fernández-González C, Aguilar-Gallegos N, González-Santiago MV. The primacy of politics in public food security policies: the case of home gardens. Sustainability 2020 May 25;12(10):4316. [doi: 10.3390/su12104316]
18. Knobloch NA. Is experiential learning authentic? J Agric Educ 2003 Dec 31;44(4):22-34. [doi: 10.5032/jae.2003.04022]
19. Tsekleves E, Cosmas J, Aggoun A. Benefits, barriers and guideline recommendations for the implementation of serious games in education for stakeholders and policymakers. Brit J Educ Technol 2014 Oct 24;47(1):164-183. [doi: 10.1111/bjet.12223]
20. Jacobs RS. Play to win over: effects of persuasive games. Psychol Pop Media Cult 2018;7(3):231-240. [doi: 10.1037/ppm0000124]
21. Ndulue C, Orji R. Games for change—a comparative systematic review of persuasive strategies in games for behavior change. IEEE Trans Games 2023 Jun;15(2):121-133. [doi: 10.1109/tg.2022.3159090]
22. Bogost I. Persuasive Games: The Expressive Power of Videogames. Cambridge, MA: MIT Press; 2007.
23. Soekarjo M, van Oostendorp H. Measuring effectiveness of persuasive games using an informative control condition. Int J Serious Games 2015 Jun 17;2(2):37-56. [doi: 10.17083/ijsg.v2i2.74]
24. De la Hera Conde-Pumpido T. Persuasive gaming: identifying the different types of persuasion through games. Int J Serious Games 2017 Mar 22;4(1):31-39. [doi: 10.17083/ijsg.v4i1.140]
25. de la Hera Conde-Pumpido T, Jansz J, Jacobs R, Schouten B, Raessens J, Kors M. Persuasive gaming: from theory-based design to validation and back . An introduction. In: de la Hera Conde-Pumpido T, Jansz J, Raessens J, Schouten B, editors. Persuasive Gaming in Context. Amsterdam, The Netherlands: Amsterdam University Press; 2021. [doi: 10.1515/9789048543939-001]
26. Lee D, Ryu S, Jeong EJ. An overview of serious game mechanism for social changes. J Korean Game Soc 2013 Apr 30;13(2):81-98. [doi: 10.7583/jkgs.2013.13.2.81]
27. Ruggiero D. The effect of a persuasive social impact game on affective learning and attitude. Comput Hum Behav 2015 Apr;45:213-221. [doi: 10.1016/j.chb.2014.11.062]

28. Peng W, Lee M, Heeter C. The effects of a serious game on role-taking and willingness to help. J Commun 2010 Nov 04;60(4):723-742. [doi: 10.1111/j.1460-2466.2010.01511.x]
29. Khaled R, Fischer R, Noble J, Biddle R. A qualitative study of culture and persuasion in a smoking cessation game. In: Proceedings of the Third International Conference on Persuasive Technology. 2008 Presented at: PERSUASIVE 2008; June 4-6, 2008; Oulu, Finland. [doi: 10.1007/978-3-540-68504-3_20]
30. Maurin KD, Girod C, Consolini JL, Belzeaux R, Etain B, Cochet B, et al. Use of a serious game to strengthen medication adherence in euthymic patients with bipolar disorder following a psychoeducational programme: a randomized controlled trial. J Affect Disord 2020 Feb 01;262:182-188. [doi: 10.1016/j.jad.2019.10.008] [Medline: 31668996]
31. Kolek L, Martinková P, Vařejková M, Šisler V, Brom C. Is video games' effect on attitudes universal? results from an empirical study comparing video games' impact on the attitude change of players with different backgrounds. Comput Assist Learn 2023 Nov 22;40(2):667-684. [doi: 10.1111/jcal.12911]
32. Orji R, Mandryk RL, Vassileva J. Improving the efficacy of games for change using personalization models. ACM Trans Comput Hum Interact 2017 Oct 06;24(5):1-22. [doi: 10.1145/3119929]
33. Espinosa-Curiel IE, de Alba-Chávez CA. Serious video games for agricultural learning: scoping review. IEEE Trans Learn Technol 2024 Feb 08;17:1155-1169. [doi: 10.1109/tlt.2024.3364086]
34. Cheung KK, Jong MS, Lee FL, Lee JH, Luk ET, Shang J, et al. FARMTASIA: an online game-based learning environment based on the VISOLE pedagogy. Virtual Real 2008 Feb 19;12(1):17-25. [doi: 10.1007/s10055-008-0084-z]
35. Jamonnak S, Cheng E. Little botany: a mobile game utilizing data integration to enhance plant science education. Int J Comput Games Technol 2017 Mar 22:1-14. [doi: 10.1155/2017/3635061]
36. Ee RW, Yap KZ, Yap KY. Herbopolis - a mobile serious game to educate players on herbal medicines. Complement Ther Med 2018 Aug;39:68-79. [doi: 10.1016/j.ctim.2018.05.004] [Medline: 30012395]
37. Jouan J, de Graeuwe M, Carof M, Baccar R, Bareille N, Bastian S, et al. Learning interdisciplinarity and systems approaches in agroecology: experience with the serious game SEGAE. Sustainability 2020 May 26;12(11):4351. [doi: 10.3390/su12114351]
38. Yongyuth P, Prada R, Nakasone A, Kawtrakul A, Prendinger H. AgriVillage: 3D multi-language internet game for fostering agriculture environmental awareness. In: Proceedings of the International Conference on Management of Emergent Digital EcoSystems. 2010 Presented at: MEDES '10; October 26, 2010; Bangkok, Thailand. [doi: 10.1145/1936254.1936280]
39. Ali A, Paliktzoglou V, Gatzoulis C, Abedallat OA. Raising awareness on hydroponics via an educational video game using an indirect teaching method. In: Proceedings of the 9th IEEE-GCC Conference and Exhibition (GCCCE). 2017 Presented at: IEEE-GCC 2017; May 8-11, 2017; Manama, Bahrain. [doi: 10.1109/ieeegcc.2017.8448136]
40. Klit KJ, Pedersen KS, Stege H. A prospective cohort study of game-based learning by digital simulation of a pig farm to train agriculture students to reduce piglet mortality. Porcine Health Manag 2018;4:28. [doi: 10.1186/s40813-018-0105-6] [Medline: 30519479]
41. Tjernström E, Lybbert TJ, Hernández RF, Correa JS. Learning by (virtually) doing: experimentation and belief updating in smallholder agriculture. J Econ Behav Organ 2021 Sep;189(5):28-50. [doi: 10.1016/j.jebo.2021.03.001]
42. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009 Mar;24(1):485-500. [doi: 10.17705/1CAIS.02428]
43. Vanden Abeele V, de Schutter B, Geurts L, Desmet S, Wauters J, Husson J, et al. P-III: a player-centered, iterative, interdisciplinary and integrated framework for serious game design and development. In: Proceedings of the Joint Conference of the Interdisciplinary Research Group of Technology, Education, Communication, and the Scientific Network on Critical and Flexible Thinking. 2011 Presented at: ITEC/CIP/t 2011; October 19-21, 2011; Ghent, Belgium. [doi: 10.1007/978-3-642-33814-4_14]
44. Bandura A. Social Learning Theory. Upper Saddle River, NJ: Prentice Hall; 1977.
45. Schunk DH, Usher EL. Social cognitive theory and motivation. In: The Oxford Handbook of Human Motivation. New York, NY: Oxford University Press; 2012.
46. Fogg BJ. Persuasive Technology Using: Computers to Change What We Think and Do. Amsterdam, The Netherlands: Elsevier Science; 2003.
47. Bang M, Torstensson C, Katzeff C. The PowerHouse: a persuasive computer game designed to raise awareness of domestic energy consumption. In: Proceedings of the First International Conference on Persuasive Technology for Human Well-Being. 2006 Presented at: PERSUASIVE 2006; May 18-19, 2006; Eindhoven, The Netherlands. [doi: 10.1007/11755494_33]
48. Khaled R, Barr P, Noble J, Fischer R, Biddle R. Fine tuning the persuasion in persuasive games. In: Proceedings of the Second International Conference on Persuasive Technology. 2007 Presented at: PERSUASIVE 2007; April 26-27, 2007; Palo Alto, CA. [doi: 10.1007/978-3-540-77006-0_5]
49. Edad mediana por entidad federativa, serie de años censales de 1995 a 2020. Instituto Nacional de Estadística y Geografía (INEGI). 2020. URL: https://www.inegi.org.mx/app/tabulados/interactivos/?pxq=Poblacion_Poblacion_04_bb9a3db4-4c69-4231-aaaf-abef21dda472&idrt=123&opc=t [accessed 2024-01-25]
50. 2020 Essential facts about the video game industry. Entertainment Software Association (ESA). 2020 Jul 15. URL: https://www.theesa.com/wp-content/uploads/2024/03/Final-Edited-2020-ESA_Essential_facts.pdf [accessed 2024-02-25]

51. Shamah-Levy T, Vielma-Orozco E, Heredia-Hernández O, Romero-Martínez M, Mojica-Cuevas J, Cuevas-Nasu L, et al. Encuesta Nacional de Salud y Nutrición 2018-19: resultados nacionales. Instituto Nacional de Salud Pública. Cuernavaca: Instituto Nacional de Salud Pública; 2020. URL: https://ensanut.insp.mx/encuestas/ensanut2018/doctos/informes/ensanut_2018_informe_final.pdf [accessed 2024-01-12]
52. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Boston, MA: Addison-Wesley Publishing Company; 1975.
53. Bandura A. Self-efficacy. In: Ramachaudran VS, editor. Encyclopedia of Human Behavior. New York, NY: Academic Press; 1994.
54. Law E, Roto V, Vermereen AP, Kort J, Hassenzahl M. Towards a shared definition of user experience. In: Proceedings of the CHI '08 Extended Abstracts on Human Factors in Computing Systems. 2008 Presented at: CHI EA '08; April 5-10, 2008; Florence, Italy. [doi: 10.1145/1358628.1358693]
55. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
56. Fu FL, Su RC, Yu SC. EGameFlow: a scale to measure learners' enjoyment of e-learning games. Comput Educ 2009 Jan;52(1):101-112. [doi: 10.1016/j.compedu.2008.07.004]
57. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
58. Yoo HS, Kim SW. Virtual farmers training: realistic simulation with amusements using historic simulation and game storyline. Int J Multimed Ubiq Eng 2014;9(5):121-130. [doi: 10.14257/ijmue.2014.9.5.11]
59. Murray C, Grant MJ, Howarth ML, Leigh J. The use of simulation as a teaching and learning approach to support practice learning. Nurse Educ Pract 2008 Jan;8(1):5-8. [doi: 10.1016/j.nepr.2007.08.001] [Medline: 17951111]
60. Babin MJ, Rivière É, Chiniara G. Theory for practice: learning theories for simulation. In: Clinical Simulation: Education, Operations, and Engineering, Second Edition. Cambridge, MA: Academic Press; 2019:97-114. [doi: 10.1016/B978-0-12-815657-5.00008-5]
61. Rosenberg MJ. Cognitive structure and attitudinal affect. J Abnorm Psychol 1956 Nov;53(3):367-372. [doi: 10.1037/h0044579] [Medline: 13376244]
62. Cooper J, Croyle RT. Attitudes and attitude change. Annu Rev Psychol 1984;35:395-426. [doi: 10.1146/annurev.ps.35.020184.002143] [Medline: 19400622]
63. Agnisarman S, Madathil KC, Stanley L. A survey of empirical studies on persuasive technologies to promote sustainable living. Sustain Comput 2018 Sep;19:112-122. [doi: 10.1016/j.suscom.2018.08.001]
64. Pakarinen A, Parisod H, Smed J, Salanterä S. Health game interventions to enhance physical activity self-efficacy of children: a quantitative systematic review. J Adv Nurs 2017 Apr;73(4):794-811. [doi: 10.1111/jan.13160] [Medline: 27688056]
65. Jaccard J, Dodge T, Dittus P. Parent-adolescent communication about sex and birth control: a conceptual framework. New Dir Child Adolesc Dev 2002(97):9-41. [doi: 10.1002/cd.48] [Medline: 14964942]
66. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? a meta-analysis of the experimental evidence. Psychol Bull 2006 Mar;132(2):249-268. [doi: 10.1037/0033-2909.132.2.249] [Medline: 16536643]
67. Janakiraman S, Watson SL, Watson WR. Using game-based learning to facilitate attitude change for environmental sustainability. J Educ Sustain Dev 2018 Oct 17;12(2):176-185. [doi: 10.1177/0973408218783286]
68. Shliakhovchuk E. Video games as awareness raisers, attitude changers, and agents of social change. Int J Comput Games Technol 2024 Jan 30:1-10. [doi: 10.1155/2024/3274715]
69. Cohen EL. Enjoyment of a counter-hedonic serious digital game: determinants and effects on learning and self-efficacy. Psychol Pop Media Cult 2016 Apr;5(2):157-170. [doi: 10.1037/ppm0000052]
70. Hong T, Cabrera J, Beaudoin CE. The role of enjoyment in a serious game for binge drinking prevention: pretest-posttest quasi-experimental study. J Med Internet Res 2020 Nov 30;22(11):e21652. [doi: 10.2196/21652] [Medline: 33252348]
71. Espinosa-Curiel IE, Pozas-Bogarin EE, Martínez-Miranda J, Pérez-Espinosa H. Relationship between children's enjoyment, user experience satisfaction, and learning in a serious video game for nutrition education: empirical pilot study. JMIR Serious Games 2020 Sep 17;8(3):e21813. [doi: 10.2196/21813] [Medline: 32940609]
72. Giannakos MN. Enjoy and learn with educational games: examining factors affecting learning performance. Comput Educ 2013 Oct;68:429-439. [doi: 10.1016/j.compedu.2013.06.005]
73. Iten N, Petko D. Learning with serious games: is fun playing the game a predictor of learning success? Brit J Educ Technol 2014 Nov 11;47(1):151-163. [doi: 10.1111/bjet.12226]

Abbreviations

CG: control group
EG: experimental group
P-III: player-centered, iterative, interdisciplinary, and integrated
PSD: persuasive system design

RCT: randomized controlled trial
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 21.05.24; peer-reviewed by D Saxena, S Nur, L Malone; comments to author 13.06.24; revised version received 10.07.24; accepted 16.07.24; published 08.08.24.

Please cite as:
García de Alba-Chávez CA, Espinosa-Curiel IE, Michel-Nava RM
Exploring the Impact of a Persuasive Serious Video Game (Farmily) on Promoting Home Gardening Among Novices: Design and Randomized Controlled Trial
JMIR Serious Games 2024;12:e60771
URL: https://games.jmir.org/2024/1/e60771
doi:10.2196/60771
PMID:39116437

Exploring the Design of Upper Limb Strength Training Through High-Intensity Interval Training Combined With Exergaming: Usability Study

Shu-Cheng Lin[1], PhD; Jing-Yu Lee[1], BA; Yong Yang[2], PhD; Chu-Chun Fang[3], PhD; Hsiao-Lin Fang[4], PhD; Tien-Hung Hou[5,6], PhD

1

2

3

4

5

6

Corresponding Author:
Tien-Hung Hou, PhD

Abstract

Background: High-intensity interval training (HIIT) has become a popular exercise strategy in modern society, with the Tabata training method being the most popular. In the past, these training methods were mostly done without equipment, but incorporating exergaming into the training may provide a new option for muscle training.

Objectives: The aim of this study was to explore the differences in upper limb muscle activation using an HIIT program combined with exergaming.

Methods: A total of 15 healthy male participants were recruited for the study, and the differences in muscle activation were compared between push-ups and exergaming (Nintendo Switch Ring Fit Adventure with the Ring-Con accessory) during HIIT. Prior to the tests, participants underwent pretests, including maximal voluntary contractions of various muscle groups, maximal push-up tests, and maximal movement tests using the exergaming device. The push-up and exergaming tests were conducted on separate days to avoid interference, with a warm-up period of 5 minutes on a treadmill before testing. Muscle activation in the lateral and anterior portions of the deltoid muscle, the sternal and clavicular heads of the pectoralis major muscle, and the latissimus dorsi muscle were measured during the maximal voluntary contractions and single-round tests for each exercise mode. A repeated measures ANOVA was used to assess the variations in muscle activation observed across the 2 distinct modes of exercise, specifically push-ups and exergaming.

Results: In exergaming, the number of repetitions for push-ups was significantly fewer than for single-site exercises across both exhaustive (mean 23.13, SD 6.36 vs mean 55.67, SD 17.83; *P*=.001; effect size [ES]: 2.43) and single-round (mean 21.93, SD 7.67 vs mean 92.40, SD 20.47; *P*=.001; ES: 4.56) training. Heart rate differences were not significant (all *P*>.05), yet exergaming led to better muscle activation in specific muscle groups, particularly the right anterior deltoid (mean 48.00%, SD 7.66% vs mean 32.84%, SD 10.27%; *P*=.001; ES: 1.67) and right pectoralis major (sternal head: mean 38.99%, SD 9.98% vs mean 26.90%, SD 12.97%; *P*=.001; ES: 1.04; clavicular head: mean 43.54%, SD 9.59% vs mean 30.09%, SD 11.59%; *P*=.002; ES: 1.26) during exhaustive training. In single-round training, similar patterns were observed with the anterior deltoid (mean 51.37%, SD 11.76% vs mean 35.47%, SD 12.72%; *P*=.002; ES: 1.30) and pectoralis major (sternal head: mean 53.27%, SD 10.79% vs mean 31.56%, SD 16.92%; *P*=.001; ES: 1.53; clavicular head: mean 53.75%, SD 13.01% vs mean 37.95%, SD 14.67%; *P*=.006; ES: 1.14). These results suggest that exergaming may be more effective for targeted muscle activation.

Conclusions: In conclusion, HIIT can increase muscle activation in the upper extremities and can be incorporated into exergaming strategies to provide a fun and engaging way to exercise.

(JMIR Serious Games 2024;12:e51730) doi:10.2196/51730

KEYWORDS

muscle; electromyography; healthy; home training; exercise

Introduction

In recent years, motion-based video games have made substantial contributions to both medical education and sports training [1]. They have shown notable effects in the rehabilitation or training of upper limbs [2-4]. The enjoyment derived from gaming can enhance participants' motivation, and when combined with specific game design, it becomes one of the hot topics in research. Compared to longer-duration, moderate-intensity exercise, the strategy of high-intensity interval training (HIIT) [5] has become the mainstream exercise approach in modern society. The most popular approach is the Tabata training method, which involves performing 8 cycles of 20 seconds of all-out exercise, interspersed with 10 seconds of complete rest, for a total exercise time of 240 seconds [6]. Results have shown significant improvements in aerobic power [7], fat oxidation [8], and muscular endurance [9]. It can be observed that the HIIT strategy not only shortens exercise participation time but also has positive effects on the body. Tabata exercises, apart from running, also include various forms of bodyweight exercises, such as push-ups, squats, and burpees [10]. Among these, push-ups are the most used bodyweight exercise in Tabata training. In a study on muscle activation for strength training, Alizadeh et al [11] investigated the muscle activation patterns of push-ups and sit-ups, measuring the activation of the anterior and lateral portions of the deltoid muscle, as well as the sternal and clavicular heads of the pectoralis major muscle. The results showed variations in muscle activation levels despite the similarity in the exercises, highlighting differences in muscle engagement across different parts of the body. Another study by Putra et al [12] explored the muscle activation levels in the upper limbs during a boxing game while in standing and sitting positions in virtual reality gaming. The study found significant differences in the activation of the upper trapezius muscle during uppercut punches, whereas no differences were observed in straight and hook punches. Combining the findings of these 2 studies, it is evident that different exercises lead to varying levels of muscle activation in different muscle groups.

Push-ups, historically used to assess upper body strength, are frequently incorporated into HIIT sessions. This exercise primarily targets the deltoid, pectoralis major, and latissimus dorsi muscles. Similarly, virtual reality gaming offers training modes specifically designed to target these muscle groups. In summary, the research highlights the diverse muscle activation patterns associated with different exercises. Push-ups, a fundamental bodyweight exercise, have been traditionally used to assess upper body strength and are a common component of HIIT workouts, effectively engaging the deltoid, pectoralis major, and latissimus dorsi muscles. Additionally, virtual reality gaming provides tailored training modes focusing on these specific muscle groups. For individuals engaged in recreational physical activities, these conventional exercise methods might be perceived as monotonous due to their limited variation, potentially leading to reduced adherence to training. This lack of variety could negatively impact exercise adherence, as the "lack of enjoyment" is frequently cited as a common barrier to regular physical activity [13].

In studies on the application of HIIT in strength training, the focus has been on investigating the rest intervals between exercises [14] and movement speed [15]. Tomoo et al [14] examined the effects of long and short rest intervals between sets and found that shorter rest intervals led to higher muscle activation levels at the same exercise intensity. In contrast, Dora et al [15] found that faster movement speed resulted in higher muscle activation levels. Taken together, these studies suggest that shorter rest intervals and faster movement speed lead to greater muscle activation. Previous literature has also shown that shorter rest intervals can improve muscle adaptations during resistance training [16,17]. However, in upper body exercise design, push-ups are commonly included as one of the training movements and are frequently used to assess upper body muscle strength and endurance [18].

Exergaming have been used for exercise training for many years and have contributed to improving exercise participation [19]. Related exergaming devices includes Xbox 360, Nintendo Wii, Nintendo Switch, and Sony PlayStation 2. Through exergaming, aerobic capacity, agility, muscle strength, muscle endurance, and coordination can be improved [20]. Regarding muscle strength training, Willaert et al [21] showed that muscle activation can be improved by more than 40%. Although the level of activation is relatively low, this study aims to design a training program with the combination of HIIT and body-sensing video games to enhance the quality and effectiveness of the training. There are relatively few studies on the use of HIIT for muscle strength training, and the level of muscle activation using exergaming combined with HIIT has not been clarified. The aim of this study was to explore the differences in upper limb muscle activation using an HIIT program combined with exergaming.

Methods

Study Participants

In this study, we recruited 15 healthy male participants. They had an average age of 24.4 (SD 10.4) years, stood at an average height of 174 cm with a minimal variance of 0.05 cm, and weighed an average of 71.9 (SD 13) kg. Their BMI averaged at 23.5 (SD 3.57) kg/m^2. All participants maintained a regular exercise routine, engaging in physical activity 3 times per week over the past year, and had experience in performing push-ups correctly. They also completed the Physical Activity Readiness Questionnaire [22] and confirmed that they had no history of upper or lower limb skeletal muscle injury or major injury. Participants were instructed to avoid vigorous activity and the intake of caffeine or supplements that enhance muscle performance for 24 hours prior to the experiment. Before the study began, all participants provided their personal information and medical history and filled out the health questionnaires and informed consent form. Additionally, the data were proofread to ensure accuracy and readability.

Ethical Considerations

The human research ethics committee of the local university approved this study, which was also approved by the human research ethics committee of the National Cheng Kung

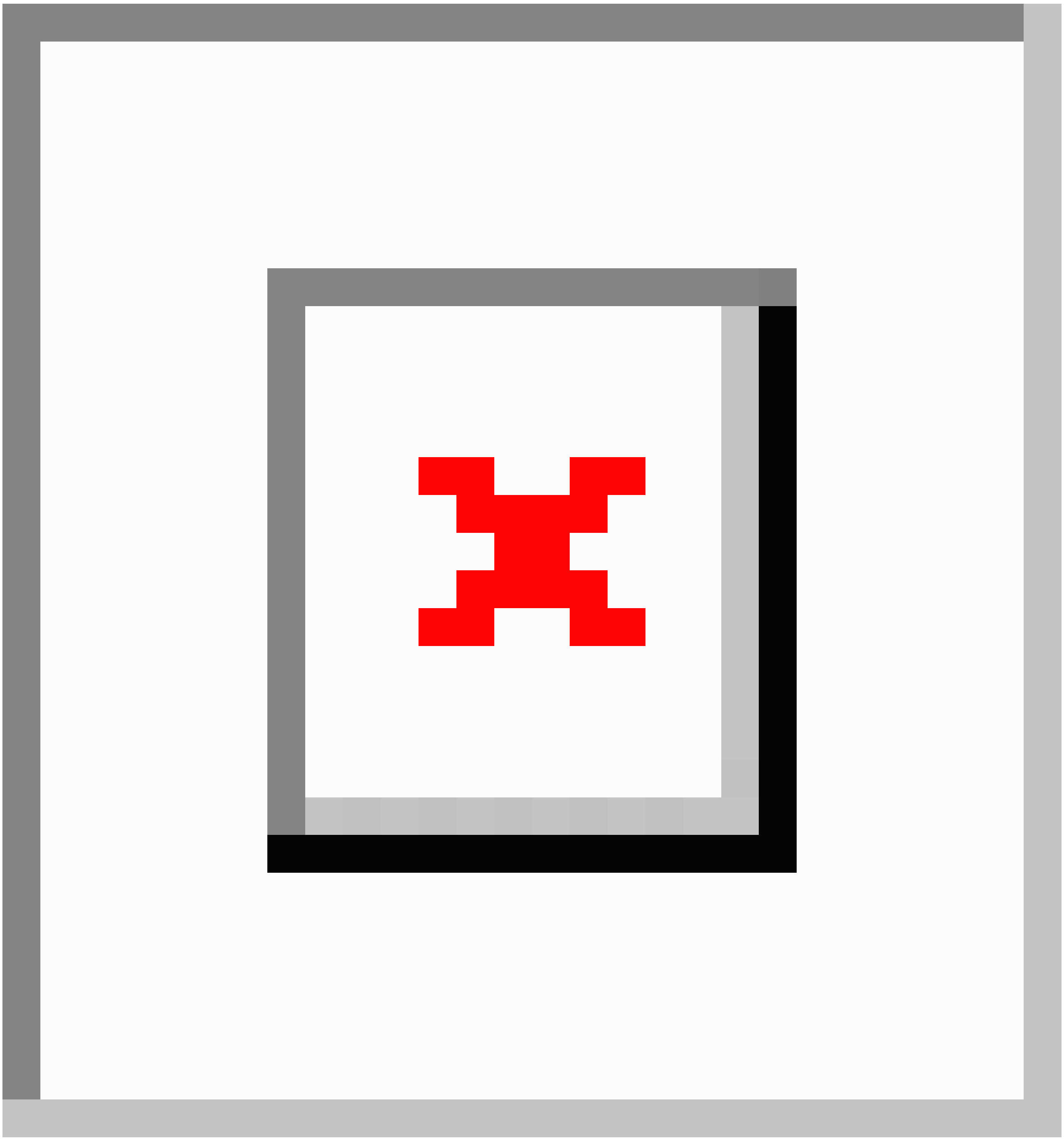
University, Taiwan (approval NCKU HREC-E-112-419-2). Users volunteered for this study and agreed to participate by signing an informed consent form. The research ensures the issues of privacy and confidentiality by assigning participants with numerical identifiers during the experiment to safeguard the confidentiality of their personal information. In terms of compensation, participants were volunteers and there was no remuneration involved.

Experimental Design

Overview

This study used a randomized, crossover, and repeated measures experimental design to compare the differences in muscle activation between push-up exercise and a Nintendo Switch Ring Fit Adventure exercise. Prior to the tests, participants underwent pretests, which included maximal voluntary contractions (MVCs) for each muscle group, maximal push-up tests, and maximal exercise tests for each part of the Nintendo Switch Ring Fit Adventure exercise. The 2 types of tests were conducted with a minimum interval of 3 days to avoid interference. Prior to each test, a 5-minute warm-up on the treadmill at a speed of 2 m/s was recommended. During the tests, muscle activation in various muscle groups was observed and measured, including the lateral and anterior portions of the deltoid muscle, the sternal and clavicular heads of the pectoralis major muscle, and the latissimus dorsi muscle, which were referenced and modified from previous studies by van den Tillaar [23], Alizadeh et al [11], and Maeo et al [24] (Figure 1). The aim was to explore the differences in upper limb muscle activation using an HIIT program combined with exergaming.

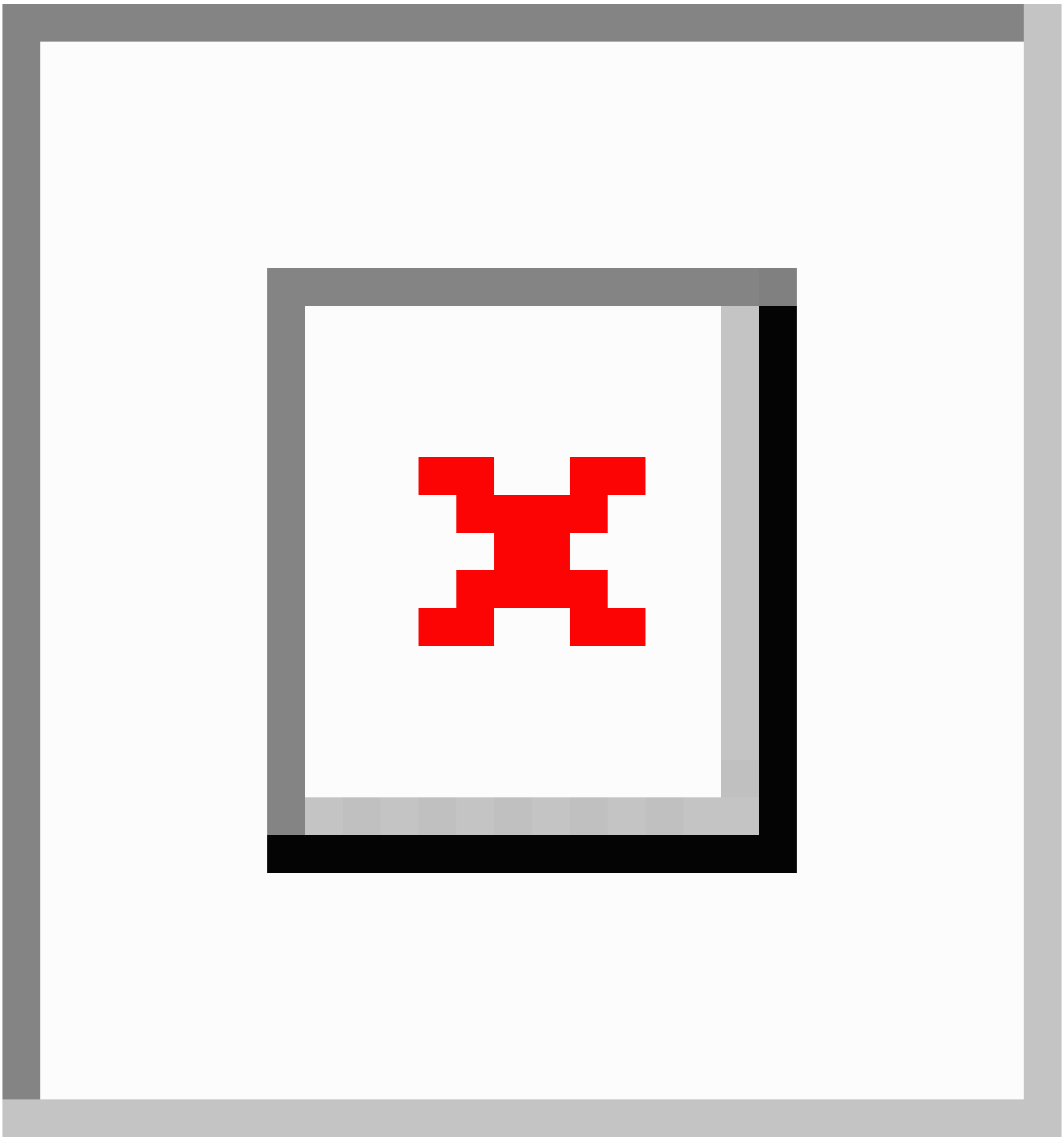
Figure 1. Experimental procedure diagram.



Exergaming

The exergaming selected for this study was Nintendo Switch Ring Fit Adventure, which combines exercise, adventure, and entertainment, allowing players to enjoy both physical workouts and gaming fun simultaneously. The game features an intuitive and user-friendly interface, suitable for players of all ages. It comes with a specialized fitness ring device (Ring-Con), an intelligent accessory that connects to the Nintendo Switch console. Through this ring, players can engage in various physical activities such as weightlifting, yoga, and aerobic exercises. The fitness ring sensor accurately captures players' movements and incorporates them into the gameplay. The game content involves unlocking levels and participating in fitness competitions through real-life physical movements. It offers a variety of fitness activities, each targeting different muscle groups, while also providing enjoyable gaming challenges. The trained muscle groups were the pectoralis major, deltoid, and latissimus dorsi muscles. The testing consisted of 2 modes: (1) maximum repetition test and (2) single-round test. The maximum repetition test was conducted using the extreme challenge mode. Participants were instructed to follow the game's pace of 60 beats per minute as closely as possible and maintain proper posture during each muscle group's testing to avoid compensation from other muscles. The single-round test was conducted using the challenge mode. Participants were instructed to perform repetitive actions as quickly as possible for 20 seconds (Figure 2).

Figure 2. Exercise training and exergaming model: (A) pectoralis major, (B) deltoid, and (C) latissimus dorsi.



Push-Up Tests

The push-up tests included 2 types of tests: (1) maximum repetition test and (2) single-round test. The maximum repetition test was based on the testing method described by Eckel et al [25]. Participants were guided to execute the test while in sync with a metronome set at 60 beats per minute, ensuring each movement, 1 second downward and 1 second upward, matched the rhythm precisely. This cadence was chosen to align with the pace of exercises conducted in the Nintendo Switch Ring Fit Adventure exercise, facilitating a consistent and controlled environment for comparison. During the test, the distance between the hands at the sternal notch level was measured and must be the same as the distance used in the single-round test. The elbow must be bent at 90° and the elbow must be fully extended when straightened. The single-round test was designed to match the duration of a single round of the Nintendo Switch Ring Fit Adventure exercise and involved performing the maximum number of squats possible within 20 seconds.

Heart Rate Tests

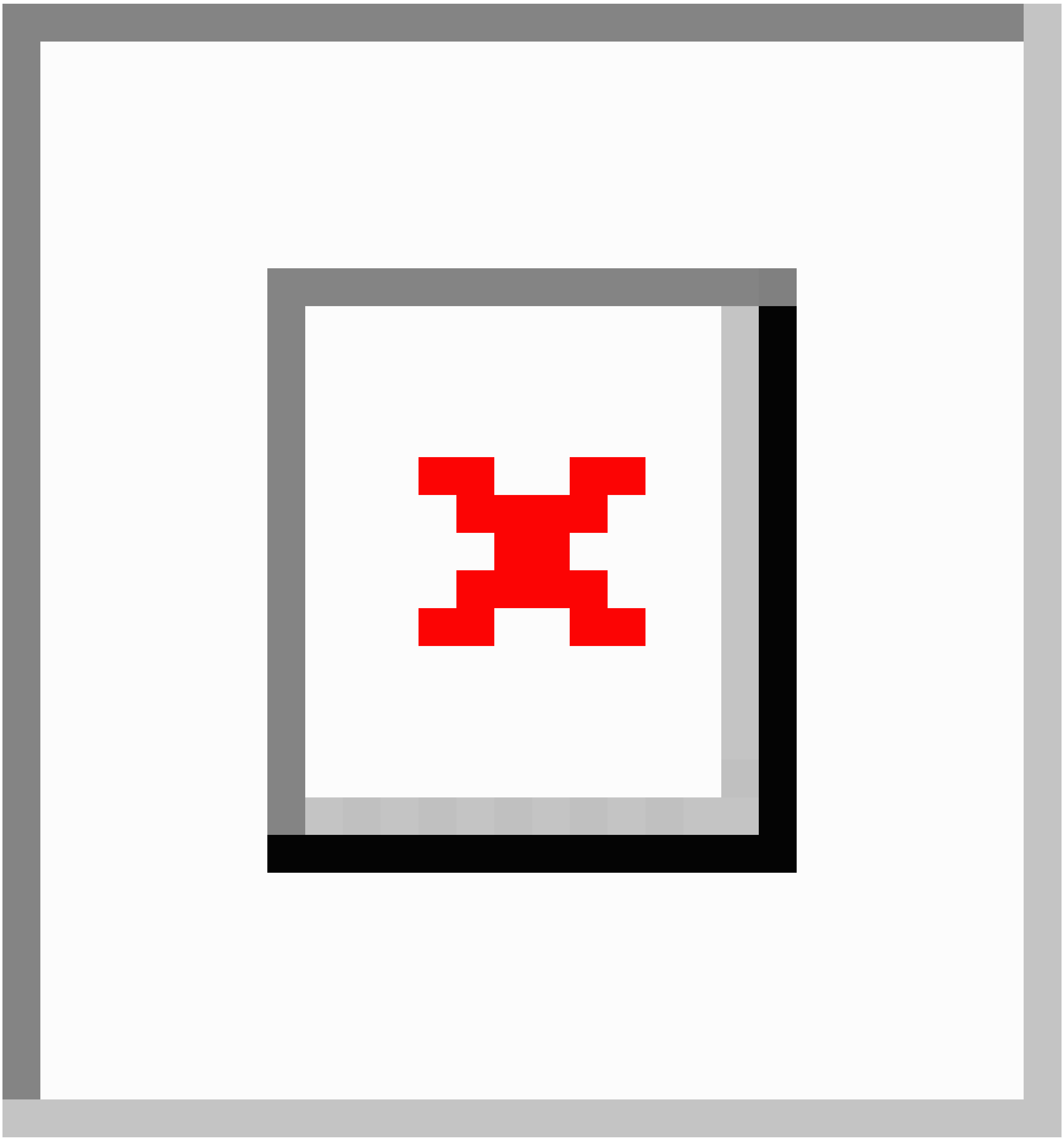
Heart rate measurements were taken using the iHeart heart rate sensor (Hexin) during push-up tests and exergaming sessions. The heart rate sensor was worn directly below the sternum and in direct contact with the skin, ensuring a comfortable fit that remained secure without slipping, even during exercise. After being fitted, the sensor was connected to the Polar Beat app (Polar Electro) for monitoring and recording purposes.

Electromyography

MVC Testing

In MVC testing, each muscle group underwent an isometric MVC prior to the test. After a running warm-up and a 3-minute rest period, the MVC test was performed according to the movements described by Konrad [26]. The test consisted of 3 attempts, each lasting 5 seconds with 1-minute rest intervals in between, and the participant was encouraged to exert maximum effort during each attempt. If the peak MVC values differed by more than 5%, additional testing was performed. The test procedures for each muscle group were as follows. (1) Deltoid: the participants were seated with their backs supported, and their arms were abducted to a position where they formed a 90° angle with the trunk, maintaining a horizontal plane. A rope, fixed to the arm, was then pulled upward to measure the force exerted. (2) Pectoralis major: the participants performed the test in a lying position with the elbows extended and flexed at 90°, holding onto a long bar with a fixed rope attached to each side of the bar and pulling upward as hard as possible to measure the force exerted. (3) Latissimus dorsi: the participants performed the test in a seated position, simulating the movement of a pull-up. The axis joint was flexed and abducted at 90°, and the rope was pulled downward the measure the force exerted, as shown in Figure 3. After the MVC testing, electromyography (EMG) electrodes were placed on the skin over the belly of the tested muscles for subsequent signal recording and analysis.

Figure 3. (A) Standardized illustrations for the MVC action of each muscle. (B) MVC sensor placement for each muscle. MVC: maximal voluntary contraction.



EMG Measurement and Analysis

The Trigno TM wireless foundation system (Delsys-EMGworks) was used for data collection in this study, measuring the anterior and lateral portions of the deltoid muscle, the sternal and clavicular heads of the pectoralis major muscle, and the latissimus dorsi muscle [11,23,24]. The system was configured with a sampling rate set at 2000 Hz per channel, tailored to the desired sample rate specifications. For the processing of EMG data, we used the EMGwork analysis software, which included steps of filtering and smoothing the EMG signals to ensure clarity and accuracy. The filtering process used a band-pass filter with a low-frequency cutoff at 20 Hz and a high-frequency cutoff at 500 Hz. Following this, the rectified EMG signals underwent further refinement using the root mean square method, which facilitated a detailed analysis of the signal's magnitude. During the MVC test, we determined the highest EMG amplitude recorded for each muscle group, referred to as EMGmax. The data processing method in exergaming was the same as the above, and the degree of muscle activation was calculated based on the values obtained from the standardized action test, expressed as a percentage of EMGmax.

Statistical Analysis

Data processing and analysis were performed using SPSS for Windows (version 20.0; IBM Corp). The data were presented as mean and SD. To investigate the variance in motion and heart rate across different movements derived from the 2 exercise models—strength endurance and single round—we used a repeated measures ANOVA. Furthermore, to assess the discrepancies in muscle activation elicited by the 2 distinct types of exercises, namely push-ups and exergaming, a paired-samples 2-tailed *t* test was used. Cohen *d* for effect size (ES) was calculated by the G*Power 3.1 software program (Heinrich-Heine-Universität), where the ESs of 0.2, 0.5, and 0.8 were considered small, medium, and large, respectively. Statistical significance was set as *P*<.05.

Results

Study Participants

For this study, a total of 15 male participants from the community were recruited. These participants were generally in good health. However, finding healthy female participants capable of performing push-ups was challenging due to their limited availability. Therefore, this study predominantly concentrated on male participants.

Comparing Strength, Performance, and Heart Rate: Push-Ups Versus Exergaming

The results showed that regardless of the exhaustive or single-round mode, the number of single-site repetitions in exergaming was significantly higher than that of push-ups (exhaustive: deltoid, mean 55.67, SD17.83; pectoralis major, mean 52.53, SD 13.61; and latissimus dorsi, mean 82.30, SD 20.82 vs push-up, 23.13, SD 6.36; *P*=.001; ES: 2.43, 2.77, and 3.84, respectively; single round: deltoid, mean 92.40, SD 20.47; pectoralis major, mean 104.27, SD 13.48; and latissimus dorsi, mean 97.33, SD 16.77 vs push-ups mean 21.93, SD 7.67; *P*=.001; ES: 4.56, 7.51, and 5.78, respectively). However, there was no difference in heart rate between the 2 modes (all *P*>.05). Taken together, these results suggest that both whole-body push-ups and single-site exergaming training can increase heart rate and can be used to train cardiorespiratory fitness (Table 1).

Table . Comparison of strength endurance, single-round performance, and heart rate between push-ups and exergaming.

Model and motion		Repetitions, mean (SD)	*F* test (*df*)	*P* value	Heart rate (bpm[a]), mean (SD)	*F* test (*df*)	*P* value
Strength endurance test			50.83	.001		3.92	.07
	Push-up	23.13 (06.36)			101.27 (14.79)		
	Exergaming, deltoid	55.67 (17.83)			96.80 (21.29)		
	Exergaming, pectoralis major	52.53 (13.61)			98.06 (18.76)		
	Exergaming, latissimus dorsi	82.30 (20.82)			91.40 (14.23)		
Single-round performance (20 s)			143.27	.001		1.56	.21
	Push-up	21.93 (07.67)			100.33 (16.32)		
	Exergaming, deltoid	92.40 (20.47)			104.80 (14.63)		
	Exergaming, pectoralis major	104.27 (13.48)			93.67 (17.11)		
	Exergaming, latissimus dorsi	97.33 (16.77)			104.13 (16.05)		

[a]bpm: beats per minute.

Strength Endurance: Push-Ups Versus Exergaming

According to the statistical results, in the exhaustive mode, the activation levels of the right anterior deltoid (mean 48.00%, SD 7.66% vs mean 32.84%, SD 10.27%; *P*=.001; ES: 1.67), right pectoralis major—sternal head (mean 38.99%, SD 9.98% vs mean 26.90%, SD 12.97%; *P*=.001; ES: 1.04), and right pectoralis major—clavicular head (mean 43.54%, SD 9.59% vs mean 30.09%, SD 11.59%; *P*=.002; ES:1.26) were significantly greater in the exergaming group than in the push-up group. Thus, these results suggest that exergaming have a better training effect on specific muscle groups (Table 2).

Table . Comparison of muscle activation between traditional push-ups and exergaming in the strength endurance test.

Muscle	Push-up activation (%), mean (SD)	Exergaming activation (%), mean (SD)	*t* test (*df*)	*P* value
Anterior deltoid	32.84 (10.27)	48.00 (7.66)	−4.096	.001
Lateral deltoid	37.36 (13.79)	44.56 (8.82)	−1.737	.10
Pectoralis major, sternal head	26.90 (12.97)	38.99 (9.98)	−4.358	.001
Pectoralis major, clavicular head	30.09 (11.59)	43.54 (9.59)	−3.784	.002
Latissimus dorsi	35.43 (10.39)	34.01 (18.05)	0.329	.75

Single-Round Test: Push-Ups Versus Exergaming

Based on the statistical analysis, in the single-round mode, the activation levels of the right anterior deltoid (mean 51.37%, SD 11.76% vs mean 35.47%, SD 12.72%; *P*=.002; ES: 1.30), right lateral deltoid (mean 52.08%, SD 10.79% vs mean 43.86%, SD 10.48%; *P*=.046; ES: 0.77), right pectoralis major–sternum head (mean 53.27%, SD 10.79% vs mean 31.56%, SD 16.92%; *P*=.001; ES: 1.53), and right pectoralis major–clavicle head (mean 53.75%, SD 13.01% vs mean 37.95%, SD 14.67%; *P*=.006; ES: 1.14) were significantly greater in the exergaming group than the push-up group. These results suggest that the exergaming designed as an HIIT exercise targeting a specific muscle group resulted in significantly higher muscle activation compared to push-ups (Table 3).

Table . Comparison of muscle activation between traditional push-ups and exergaming in the single-round test.

Muscle	Push-up activation (%), mean (SD)	Exergaming activation (%), mean (SD)	*t* test (*df*)	*P* value
Anterior deltoid	35.47 (12.72)	51.37 (11.76)	−3.705	.002
Lateral deltoid	43.86 (10.48)	52.08 (10.79)	−2.190	.046
Pectoralis major, sternal head	31.56 (16.92)	53.27 (10.79)	−4.266	.001
Pectoralis major, clavicular head	37.95 (14.67)	53.75 (13.01)	−3.236	.006
Latissimus dorsi	41.49 (11.00)	38.17 (16.87)	0.645	.529

Discussion

Principal Findings

The results indicated that, in terms of the number of repetitions performed, both the strength endurance and single-round tests showed significantly higher execution rates for the exergaming training mode compared to push-ups. However, there were no significant differences in heart rate between the 2 modes. Regarding muscle activation, in the strength endurance test, exergaming exhibited significantly higher activation levels than push-ups in the anterior deltoid, pectoralis major–sternal head, and pectoralis major–clavicular head muscles. In the single-round test, exergaming demonstrated significantly higher activation levels than push-ups in the anterior deltoid, lateral deltoid, pectoralis major–sternal head, and pectoralis major–clavicular head muscles.

Heart Rate Response in Exergaming

Interestingly, the study did not observe significant differences in heart rate between the exergaming and push-up groups. This result contrasts with previous studies indicating that exergaming can lead to higher heart rates due to the immersive and stimulating nature of video game–based exercises [27,28]. The lack of significant heart rate differences could be attributed to the individual variability in cardiovascular responses and the adaptability of participants to the exergaming interface. HIIT has been shown to be an effective way to improve cardiovascular fitness and overall health [7]. Heart rate is an important factor in both exergaming and HIIT. Monitoring heart rate can help individuals ensure that they are working at an appropriate intensity level to achieve their fitness goals. In HIIT, heart rate can be used to guide the high-intensity intervals and rest periods to optimize the workout's effectiveness [29]. Overall, the use of exergaming and HIIT can provide a fun and effective way to improve physical fitness and health, with heart rate monitoring serving as an important tool to help individuals achieve their goals.

Another potential explanation for these results is that the use of exergaming may increase motivation and engagement in physical activity, leading to greater adherence to exercise programs [30]. This is especially important given the high rates of sedentary behavior and physical inactivity in modern society. Exergaming may provide a fun and enjoyable way to engage in physical activity, potentially leading to increased frequency and duration of exercise sessions [31,32]. This is particularly relevant for individuals who may struggle to engage in more traditional forms of exercise due to boredom, lack of motivation, or physical limitations.

EMG Response in Exergaming

This study investigated the effects of HIIT on muscle activation in the upper extremities. The results suggest that exergaming may be a more effective training method for upper extremity muscle activation compared to push-up exercises. The reason for this difference may be due to the specific muscles activated during exergaming, as the Nintendo Switch Ring-Con requires movements that engage the lateral and anterior parts of the deltoid muscle, pectoralis major muscle, and latissimus dorsi muscle. These findings suggest that exergaming can be a viable option for those looking to improve upper extremity muscle activation.

In interpreting the results, the study found that the exergaming training mode exhibited superior performance in terms of the number of repetitions compared to traditional push-ups, both in the strength endurance test and the single-round test. This finding aligns with previous research indicating the effectiveness of exergaming in enhancing endurance and strength capacities [33,34]. The higher execution rates in the exergaming group suggest that this interactive gaming approach offers a more engaging and motivating environment, encouraging participants to perform better and prolong their workout sessions compared to conventional push-ups. EMG can reflect the response of muscles during strength training. Alizadeh et al [11] investigated 2 common strength training exercises, push-ups and sit-ups, and measured muscle activation in the major muscle groups involved, such as the lateral and anterior portions of the deltoid and the sternal and clavicular heads of the pectoralis major. The results showed that even with the same exercise, different muscle groups were activated to varying degrees, indicating the importance of focusing on specific muscle group activation for muscle training. Compared to traditional strength training, there has been relatively little research on exergaming, but Putra et al [12] investigated the activation levels of upper limb muscles during a punching game while in standing and sitting positions. The results showed significant differences in the activation of the upper trapezius muscle during the execution of an uppercut punch, but not for straight or hook punches, indicating that the fixedness of the movements also affects the activation levels of different muscle groups in exergaming. Taken together, these studies suggest that different movements can affect muscle activation levels in different muscle groups. Therefore, choosing appropriate exercises and training modes is important for muscle strength training. Push-ups are a common exercise that mainly trains muscle groups, such as the deltoid, pectoralis major, and latissimus dorsi muscles, and are commonly used in HIIT. Training modes in exergaming are also available for these muscle groups, providing more diverse options for muscle training. Overall, these research results indicate the importance of understanding the relationship between movements and muscle groups.

Regarding muscle activation patterns, the exergaming group demonstrated significantly higher activation levels in specific muscles compared to traditional push-ups. In the strength endurance test, the anterior deltoid, pectoralis major–sternal head, and pectoralis major–clavicular head muscles exhibited increased activation during exergaming sessions. These findings corroborate with prior studies highlighting the targeted muscle engagement achieved through exergaming interventions [21]. The single-round test further showed elevated muscle activation in the anterior deltoid, lateral deltoid, pectoralis major–sternal head, and pectoralis major–clavicular head muscles during exergaming activities. This specific muscle activation pattern emphasizes the comprehensive nature of exergaming exercises, engaging various upper body muscles simultaneously [35]. One potential explanation for these results is that the Ring-Con is a novel type of resistance training that provides a more targeted and isolated workout for specific muscle groups [36]. This may allow individuals to activate and recruit more muscle fibers, leading to increased muscle activation compared to more traditional exercises such as push-ups. Additionally, the Ring-Con provides a unique form of resistance that can be adjusted to individual fitness levels, potentially allowing for greater customization and variety in workout routines [12].

Conclusions

In conclusion, this study demonstrated that exergaming may be a more effective strategy for upper extremity muscle activation compared to push-up exercises during HIIT. The specific movements required by the Nintendo Switch Ring-Con may activate the lateral and anterior parts of the deltoid muscle, pectoralis major muscle, and latissimus dorsi muscle more effectively. Furthermore, the HIIT protocol used in this study provides a time-efficient method for strength training. Incorporating exergaming into an HIIT program may provide a more engaging and effective strategy for improving upper extremity muscle activation. Further research is needed to investigate the long-term effects of exergaming on upper extremity muscle activation and strength.

Acknowledgments

Financial support provided by the National Science and Technology Council (NSTC 112-2637-E-165-002) and the Ministry of Sciences and Technology (MOST 111-2813-C-165-026-H) of the Executive Yuan. It is gratefully acknowledged. In the writing process, we used ChatGPT (OpenAI) [37] for minor translation and editing.

Authors' Contributions

SCL carried out the laboratory experiments, analyzed the data, interpreted the results, prepared the figures and tables, and prepared the manuscript. JYL, HLF, YY and CCF assisted in the data collection and the discussion of the literature. THH designed the study, supervised the experimental procedure, and reviewed the entire preparation of the manuscript.

Conflicts of Interest

None declared.

References

1. Tolks D, Schmidt JJ, Kuhn S. The role of AI in serious games and gamification for health. JMIR Serious Games 2024 Jan 15;12(1):e48258. [doi: 10.2196/48258] [Medline: 38224472]
2. Koutsiana E, Ladakis I, Fotopoulos D, Chytas A, Kilintzis V, Chouvarda I. Serious gaming technology in upper extremity rehabilitation: scoping review. JMIR Serious Games 2020 Dec 11;8(4):e19071. [doi: 10.2196/19071] [Medline: 33306029]
3. Vieira C, Ferreira da Silva Pais-Vieira C, Novais J, Perrotta A. Serious game design and clinical improvement in physical rehabilitation: systematic review. JMIR Serious Games 2021 Sep 23;9(3):e20066. [doi: 10.2196/20066] [Medline: 34554102]
4. Villada Castillo JF, Montoya Vega MF, Muñoz Cardona JE, et al. Design of virtual reality exergames for upper limb stroke rehabilitation following iterative design methods: usability study. JMIR Serious Games 2024 Jan 11;12(1):e48900. [doi: 10.2196/48900] [Medline: 38206670]
5. Weston M, Taylor KL, Batterham AM, Hopkins WG. Effects of low-volume high-intensity interval training (HIT) on fitness in adults: a meta-analysis of controlled and non-controlled trials. Sports Med 2014 Jul;44(7):1005-1017. [doi: 10.1007/s40279-014-0180-z] [Medline: 24743927]
6. Viana RB, de Lira CAB, Naves JPA, Coswig VS, del Vecchio FB, Gentil P. Tabata protocol: a review of its application, variations and outcomes. Clin Physiol Funct Imaging 2019 Jan;39(1):1-8. [doi: 10.1111/cpf.12513] [Medline: 29608238]
7. Viana RB, Naves JPA, de Lira CAB, et al. Defining the number of bouts and oxygen uptake during the "Tabata protocol" performed at different intensities. Physiol Behav 2018 May 15;189:10-15. [doi: 10.1016/j.physbeh.2018.02.045] [Medline: 29486169]
8. Pearson RC, Olenick AA, Green ES, Jenkins NT. Tabata‐style functional exercise increases resting and postprandial fat oxidation but does not reduce triglyceride concentrations. Exp Physiol 2020 Mar;105(3):468-476. [doi: 10.1113/EP088330] [Medline: 31916294]
9. Menz V, Marterer N, Amin SB, Faulhaber M, Hansen AB, Lawley JS. Functional vs running low-volume high-intensity interval training: effects on VO2max and muscular endurance. J Sports Sci Med 2019 Aug 1;18(3):497-504. [Medline: 31427872]
10. Tabata I. Tabata training: one of the most energetically effective high-intensity intermittent training methods. J Physiol Sci 2019 Jul;69(4):559-572. [doi: 10.1007/s12576-019-00676-7] [Medline: 31004287]
11. Alizadeh S, Rayner M, Mahmoud MMI, Behm DG. Push-ups vs bench press differences in repetitions and muscle activation between sexes. J Sports Sci Med 2020 May 1;19(2):289-297. [Medline: 32390722]
12. Putra BAM, Masduchi RH, Tinduh D, Pawana I. Upper limb muscles activity during punches in virtual reality exergame on standing and sitting position. Surabaya Physical Medicine and Rehabilitation Journal 2021 Feb;3(1):14-22. [doi: 10.20473/spmrj.v3i1.21550]
13. Bartlett JD, Close GL, MacLaren DPM, Gregson W, Drust B, Morton JP. High-intensity interval running is perceived to be more enjoyable than moderate-intensity continuous exercise: implications for exercise adherence. J Sports Sci 2011 Mar;29(6):547-553. [doi: 10.1080/02640414.2010.545427] [Medline: 21360405]
14. Tomoo K, Suga T, Dora K, et al. Impact of inter-set short rest interval length on inhibitory control improvements following low-intensity resistance exercise in healthy young males. Front Physiol 2021 Nov 22;12:741966. [doi: 10.3389/fphys.2021.741966] [Medline: 34880772]
15. Dora K, Suga T, Tomoo K, et al. Effect of very low-intensity resistance exercise with slow movement and tonic force generation on post-exercise inhibitory control. Heliyon 2021 Feb 18;7(2):e06261. [doi: 10.1016/j.heliyon.2021.e06261] [Medline: 33665431]
16. Hughes L, Paton B, Rosenblatt B, Gissane C, Patterson SD. Blood flow restriction training in clinical musculoskeletal rehabilitation: a systematic review and meta-analysis. Br J Sports Med 2017 Jul;51(13):1003-1011. [doi: 10.1136/bjsports-2016-097071] [Medline: 28259850]
17. Goto K, Nagasawa M, Yanagisawa O, Kizuka T, Ishii N, Takamatsu K. Muscular adaptations to combinations of high- and low-intensity resistance exercises. J Strength Cond Res 2004 Nov;18(4):730-737. [doi: 10.1519/R-13603.1] [Medline: 15574075]
18. van den Tillaar R, Saeterbakken A. Effect of fatigue upon performance and electromyographic activity in 6-RM bench press. J Hum Kinet 2014 Apr 9;40:57-65. [doi: 10.2478/hukin-2014-0007] [Medline: 25031673]
19. Bock BC, Dunsiger SI, Ciccolo JT, et al. Mediators of physical activity between standard exercise and exercise video games. Health Psychol 2019 Dec;38(12):1107-1115. [doi: 10.1037/hea0000791] [Medline: 31512923]
20. Comeras-Chueca C, Marin-Puyalto J, Matute-Llorente A, Vicente-Rodriguez G, Casajus JA, Gonzalez-Aguero A. Effects of active video games on health-related physical fitness and motor competence in children and adolescents with overweight or obesity: systematic review and meta-analysis. JMIR Serious Games 2021 Oct 18;9(4):e29981. [doi: 10.2196/29981] [Medline: 34661549]

21. Willaert J, de Vries AW, Tavernier J, van Dieen JH, Jonkers I, Verschueren S. Does a novel exergame challenge balance and activate muscles more than existing off-the-shelf exergames? J Neuroeng Rehabil 2020 Jan 15;17(1):6. [doi: 10.1186/s12984-019-0628-3] [Medline: 31941518]
22. Adams R. Revised Physical Activity Readiness Questionnaire. Can Fam Physician 1999 Apr;45:992,995,1004-992,995,1005. [doi: 10.1111/sms.13964] [Medline: 10216799]
23. van den Tillaar R. Comparison of kinematics and muscle activation between push-up and bench press. Sports Med Int Open 2019 Sep 5;3(3):E74-E81. [doi: 10.1055/a-1001-2526] [Medline: 31508485]
24. Maeo S, Chou T, Yamamoto M, Kanehisa H. Muscular activities during sling- and ground-based push-up exercise. BMC Res Notes 2014 Mar 28;7(1):192. [doi: 10.1186/1756-0500-7-192] [Medline: 24678968]
25. Eckel TL, Watkins CM, Archer DC, et al. Bench press and pushup repetitions to failure with equated load. International Journal of Sports Science & Coaching 2017 Oct 12;12(5):647-652. [doi: 10.1177/1747954117733879]
26. Konrad P. The ABC of EMG: A Practical Introduction to Kinesiological Electromyography: Noraxon U.S.A. Inc; 2005:30-35 URL: https://www.noraxon.com/wp-content/uploads/2014/12/ABC-EMG-ISBN.pdf [accessed 2024-04-03]
27. Kircher E, Ketelhut S, Ketelhut K, et al. Acute effects of heart rate-controlled exergaming on vascular function in young adults. Games Health J 2022 Feb;11(1):58-66. [doi: 10.1089/g4h.2021.0196] [Medline: 34986027]
28. Kraft JA, Russell WD, Bowman TA, Selsor CWIII, Foster GD. Heart rate and perceived exertion during self-selected intensities for exergaming compared to traditional exercise in college-age participants. J Strength Cond Res 2011 Jun;25(6):1736-1742. [doi: 10.1519/JSC.0b013e3181e06f13] [Medline: 21386720]
29. Ito S. High-intensity interval training for health benefits and care of cardiac diseases - the key to an efficient exercise protocol. World J Cardiol 2019 Jul 26;11(7):171-188. [doi: 10.4330/wjc.v11.i7.171] [Medline: 31565193]
30. Sadeghi H, Jehu DA, Daneshjoo A, et al. Effects of 8 weeks of balance training, virtual reality training, and combined exercise on lower limb muscle strength, balance, and functional mobility among older men: a randomized controlled trial. Sports Health 2021;13(6):606-612. [doi: 10.1177/1941738120986803] [Medline: 33583253]
31. Sween J, Wallington SF, Sheppard V, Taylor T, Llanos AA, Adams-Campbell LL. The role of exergaming in improving physical activity: a review. J Phys Act Health 2014 May;11(4):864-870. [doi: 10.1123/jpah.2011-0425] [Medline: 25078529]
32. Vandoni M, Codella R, Pippi R, et al. Combatting sedentary behaviors by delivering remote physical exercise in children and adolescents with obesity in the COVID-19 era: a narrative review. Nutrients 2021 Dec 14;13(12):4459. [doi: 10.3390/nu13124459] [Medline: 34960011]
33. Viana RB, de Oliveira VN, Dankel SJ, et al. The effects of exergames on muscle strength: a systematic review and meta‐analysis. Scand Med Sci Sports 2021 Aug;31(8):1592-1611. [doi: 10.1111/sms.13964] [Medline: 33797115]
34. Esmaeilzadeh S, Kumpulainen S, Pesola AJ. Strength-cognitive training: a systemic review in adults and older adults, and guidelines to promote “strength exergaming” innovations. Front Psychol 2022 May 27;13:855703. [doi: 10.3389/fpsyg.2022.855703] [Medline: 35712202]
35. Novák D. Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice: IGI Global; 2015.
36. Owen PJ, Miller CT, Mundell NL, et al. Which specific modes of exercise training are most effective for treating low back pain? network meta-analysis. Br J Sports Med 2020 Nov;54(21):1279-1287. [doi: 10.1136/bjsports-2019-100886] [Medline: 31666220]
37. ChatGPT. OpenAI. URL: https://chat.openai.com/ [accessed 2024-04-04]

Abbreviations

EMG: electromyography
ES: effect size
HIIT: high-intensity interval training
MVC: maximal voluntary contraction

Edited by R Baranyi, T Leung; submitted 09.08.23; peer-reviewed by CH Chang, P Liu; revised version received 26.02.24; accepted 11.03.24; published 17.04.24.

Please cite as:
Lin SC, Lee JY, Yang Y, Fang CC, Fang HL, Hou TH
Exploring the Design of Upper Limb Strength Training Through High-Intensity Interval Training Combined With Exergaming: Usability Study
JMIR Serious Games 2024;12:e51730
URL: https://games.jmir.org/2024/1/e51730
doi:10.2196/51730

Home-Based, Low-Intensity, Gamification-Based, Interactive Physical-Cognitive Training for Older Adults Using the ADDIE Model: Design, Development, and Evaluation of User Experience

Teerawat Kamnardsiri[1], PhD; Sirintip Kumfu[2], MSc; Peeraya Munkhetvit[3], PhD; Sirinun Boripuntakul[2,*], PhD; Somporn Sungkarat[2,*], PhD

1

2

3

*these authors contributed equally

Corresponding Author:
Somporn Sungkarat, PhD

Abstract

Background: Declines in physical and cognitive function are natural biological processes, leading to an increased risk of falls. Promising evidence suggests that combined physical-cognitive exercise has beneficial effects in improving both physical and cognitive health. Although moderate-to-high exercise intensity is commonly recommended, it might be impractical for older adults facing physical limitations or contraindications. Thus, low-intensity exercise is a viable option. The main barriers to engaging in exercise in older adults include transportation, time, motivation, and enjoyment. To overcome these challenges, a home-based, gamification-based training system may provide an effective approach to enhance exercise adherence.

Objective: This study aimed to develop and evaluate the usability of a low-intensity, gamification-based, interactive physical-cognitive exercise for older adults in a home-based setting.

Methods: The prototype of a game-based physical-cognitive exercise was created following the ADDIE model (analysis, design, development, implementation, and evaluation) and assessed for user experience in older adults. A total of 15 older adults engaged in the game-based physical-cognitive exercise at home for 60 minutes per day, 3 days per week, for 4 weeks. The usability of the game-based training system was evaluated using the system usability scale (SUS) after completion of a 4-week training program. As for satisfaction, the 8-item Physical Activity Enjoyment Scale (PACES) questionnaire was used to assess participants' enjoyment level after 1 week and 4 weeks of training. Descriptive statistics were used to illustrate the SUS score. A Wilcoxon signed-rank test was used to compare the PACES scores between the first week and the end of the 4-week period, with significance set at $P<.05$.

Results: As for experts' consensus, the game-based training consisted of 3 games: Ocean Diver, Road Runner, and Moving and Memorizing. The games had 3 levels of difficulty: beginner, intermediate, and advanced. A computer vision–based system was selected as the delivery platform for a home setting. The total SUS score for all participants was mean 87.22 (SD 5.76), indicating the user's perception of the usability of a system ranging from good to excellent. At the end of the 4-week training, the total PACES score was significantly greater than the first week, suggesting an improvement in enjoyment (first week: mean 44.93, SD 3.99 vs fourth week: mean 50.53, SD 4.70; P=.001).

Conclusions: The prototype of low-intensity, gamification-based, interactive physical-cognitive training was designed and developed using the ADDIE model, which included both experts and end users in the process. The findings showed that the exergame prototype was a usable and practical approach for a home-based setting, enhancing older adults' enjoyment and motivation. Further research is warranted to determine the effectiveness of such gamification-based training in promoting physical and cognitive functions.

(JMIR Serious Games 2024;12:e59141) doi:10.2196/59141

KEYWORDS

exergame; physical-cognitive training; computer-based interventions; gamification; older adults; instructional design model; low-intensity

Introduction

Aging is a natural biological process leading to a gradual decline in both physical and cognitive abilities. Physical decline in older adults, which contributes to their susceptibility to falls, includes intraindividual factors such as muscle weakness, delayed reaction time, impaired vision and proprioception, and impaired balance [1]. In addition to a decline in physical function, older adults may also experience problems in cognitive function, primarily involving executive function, attention, and processing speed, which are significant risk factors for falls [2]. Given that physical and cognitive impairments coexist in aging and are indicators of the risk of falls, it is essential to identify effective strategies to enhance both physical and cognitive functions for fall prevention.

It is well established that exercise serves as an effective strategy for fall prevention among older adults by ameliorating physiological fall risk factors such as muscle weakness and poor balance [3,4]. Accumulating evidence has demonstrated that exercise benefits not only physical function but also cognitive function. Given that cognitive impairment is an independent risk factor for falls, improving cognitive function through exercise may further contribute to fall risk reduction [5-7]. Among various types of exercise, combined physical-cognitive training (simultaneous training) has a profound effect on cognitive and physical health, reducing the risk of falls in older adults, and it is superior to single-component training [8,9]. Although moderate-to-high intensity exercise is widely recommended for reducing the risk of falls, its applicability may be limited for certain older adults with physical restrictions or contraindications. Currently, a growing body of evidence has suggested that low-intensity exercise has a positive effect on cognitive and physical performance in adults and older adults with and without pathologic conditions [10-13]. With this, a low-intensity exercise may be a promising alternative approach for fall prevention in older adults who experience physical limitations that prevent them from engaging in moderate-to-high intensity exercise.

With current technological advancements, gamification, which involves using game-based mechanics to motivate individual action and learning for a specific target, is increasingly used in health care services, especially for promoting exercise [14,15]. Among gamification-based exercises, combined physical-cognitive training has emerged through exergames that require individuals to move their bodies to interact with the game for achieving training purposes [16-18]. Previous studies have shown that exergame-based interventions improve both physical aspects (eg, balance, gait, and physical fitness) and cognitive domains (eg, executive functions, memory, and processing speed) in older adults while also facilitating enjoyment and motivation, which are critical mediators of training adherence and goal-directed achievement [19-22]. Nowadays, many commercial interactive game-based training products have emerged, such as Kinect Xbox, Nintendo Wii, Sony PlayStation, virtual reality systems, and computer vision–based applications. Amid the exergames in the commercial market, human pose estimation, which is one of the computer vision–based platforms, has rapidly developed in recent years. Specifically, human pose estimation involves extracting an individual's joint positions from an image or video to create a skeletal shape. With this, an algorithm of machine learning creates a human pose estimation model based on dataset samples of human movements in space, which ultimately enables machines to correctly interact with individuals [23,24]. One interesting feature of computer vision–based systems using human pose estimation is that they are markerless full-body trackers that enable users to naturally interact with games in real time. Additionally, it is reliable for representing a person's movements, resulting in proper feedback and guidance to the users during exercise [25]. Its advantages include its low cost, user-friendliness, simpler operation, and the requirement of only a personal computer, thereby increasing accessibility. Therefore, a computer vision–based system appears to be a viable alternative to delivering home-based exercise programs for older adults.

The ADDIE model is an instructional systems design framework used to design and develop learning experiences that involve identifying the requirements and understanding the solutions that learners should achieve [26]. The ADDIE model comprises 5 phases: analysis, design, development, implementation, and evaluation [27]. There has been growing interest in using the ADDIE model to develop health services for various purposes [28-32]. However, previous works utilizing the ADDIE model to develop exergames aimed at enhancing physical and cognitive functions are scarce. Considering the usability challenges faced by older adults, developers should design exergame interfaces that are user-friendly for this population. Aside from the technologies used, commercial exergames available in the market may not be specific for older people in terms of training purposes, complexity, intensity, and safety. Thus, the ADDIE model may be a useful approach for designing and developing exergames for older adults, particularly in a home environment.

Collectively, a game-based, combined physical-cognitive training that is accessible and user-friendly, boosts motivation, and is tailored to suit older adults' capabilities would represent an efficient approach in terms of time and resources. Therefore, this study focused on developing the prototype of a home-based, low-intensity, gamification-based, interactive physical-cognitive training for older adults using the ADDIE model. Notably, the usability of the system and participants' satisfaction will be investigated to facilitate practical implementation for community-dwelling older adults.

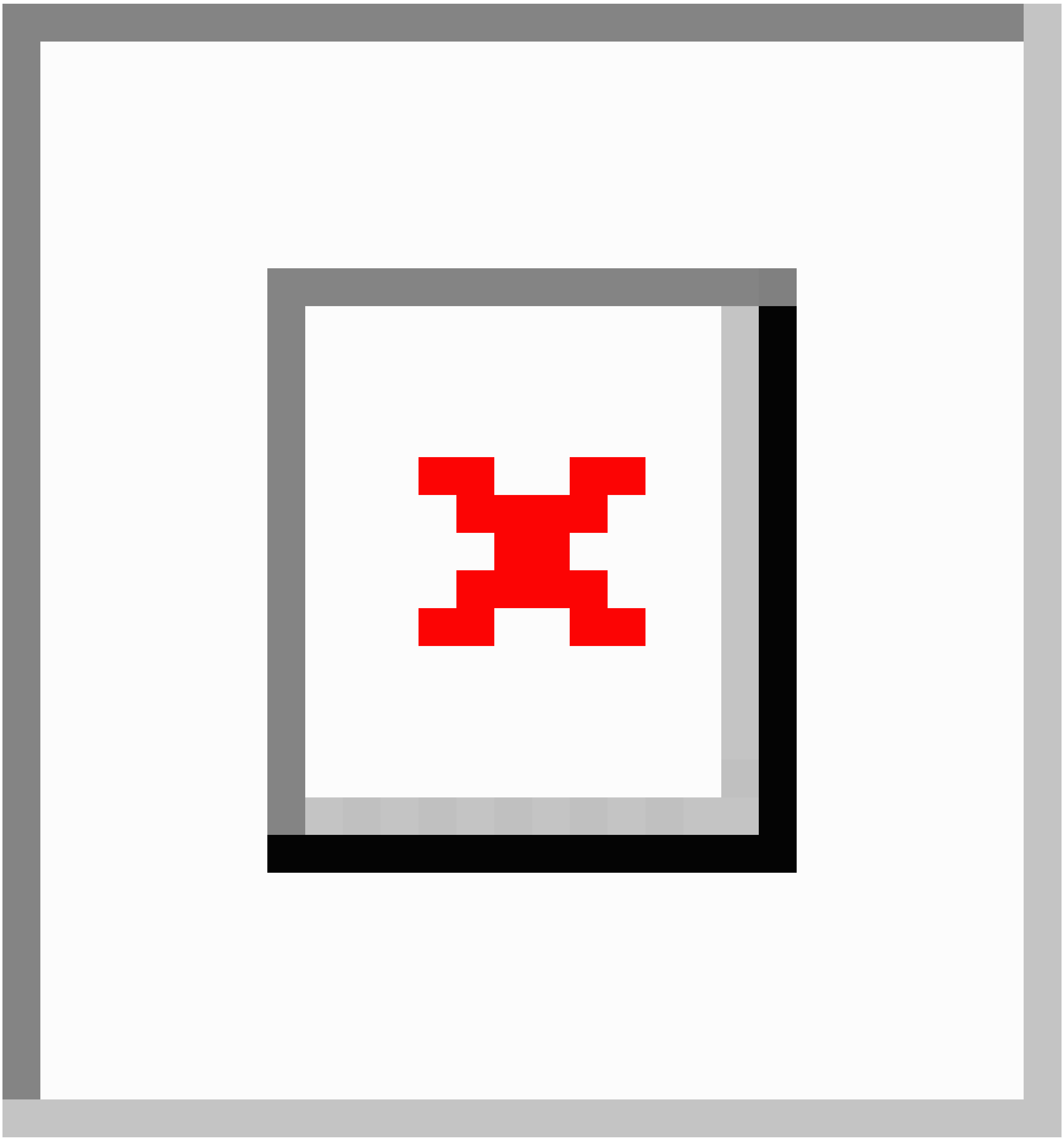
Methods

Study Design

Overview

This study utilized the ADDIE model to design and develop a low-intensity, combined physical-cognitive exercise in the form of an exergame delivered in a home-based format for older adults. The prototype of a gamification-based, interactive physical-cognitive training was built upon five phases: (1) analysis, (2) design, (3) development, (4) implementation, and (5) evaluation of target users' feedback, response, and satisfaction. The ADDIE model is depicted in Figure 1A, and the development milestones of this study are presented in Figure 1B.

Figure 1. (**A**) Development process of the game-based training system using the ADDIE (analysis, design, development, implementation, and evaluation) model. (**B**) Development milestone of the game-based training system prototype.



Phase I: Analysis Phase

During the analysis phase, a focus group expert interview was conducted to extract essential contents in terms of knowledge related to effective physical and cognitive training programs for older adults and potential core game ideas. A total of 7 experts participated in the brainstorming session, including 4 physical therapists and 1 occupational therapist (ages ranging from 38 to 60 years, with 10-25 years of experience in physical and cognitive rehabilitation for older adults, both with and without cognitive impairment), 1 game programmer (29 years old, with 4 years of experience in developing codebases for video games or related software), and 1 game artist (29 years old, with 5 years of experience in developing game mechanics and interfaces).

In this study, the core elements of the exergame consisted of two training components that were trained simultaneously: (1) a physical component at a low intensity level focusing on enhancing dynamic balance, coordination, and muscle strength and endurance of upper and lower limbs, and (2) a cognitive component focusing on memory, attention, and executive function (eg, anticipation, planning, switching, and inhibition) that closely related to balance and falls in older adults [2,33].

Regarding the delivery platform, a computer vision–based system was selected due to its low cost, accessibility, and simple operation, allowing older adults to independently administer self-training at home.

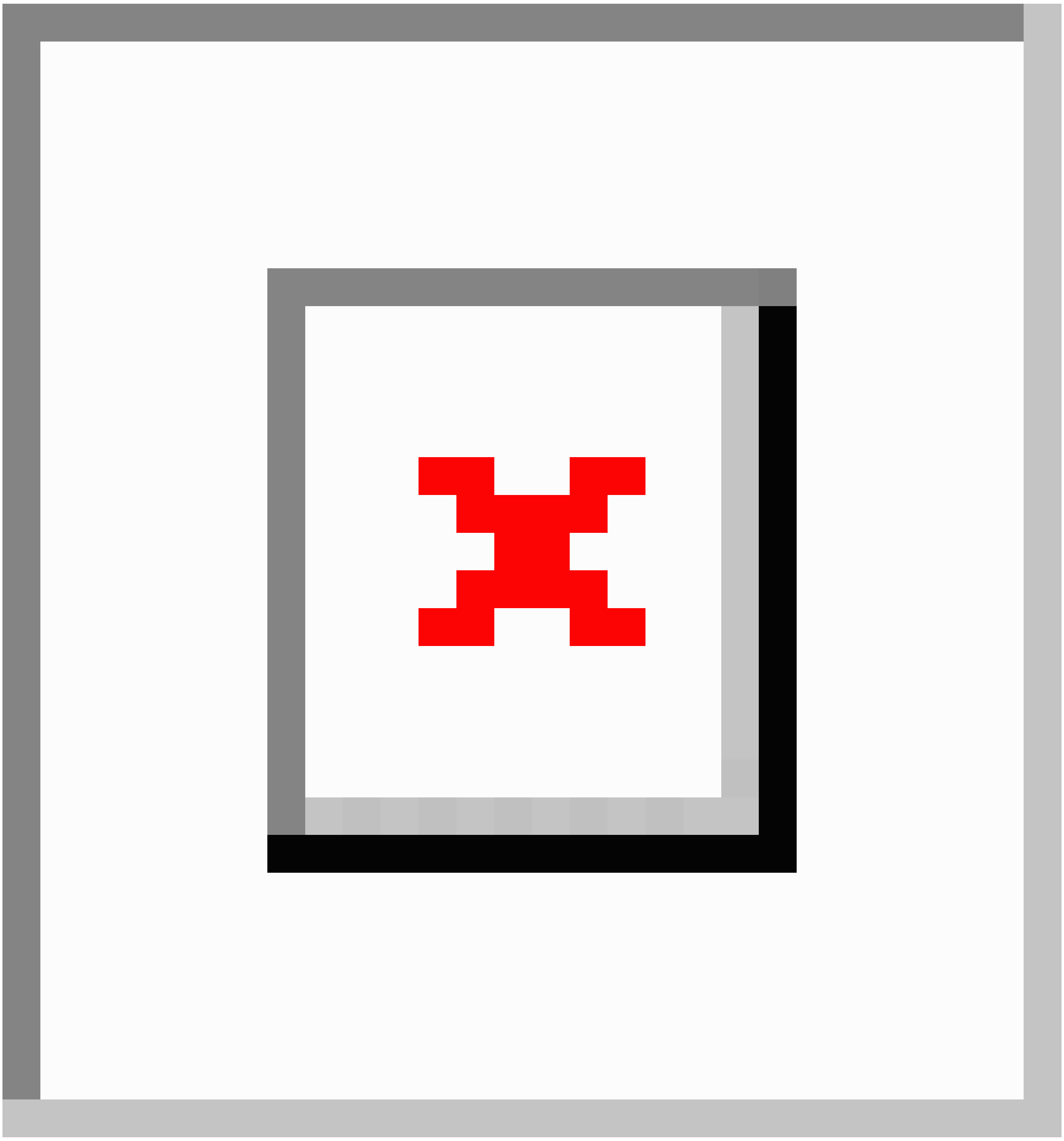
Phase II: Design Phase

In the design phase, the overall game structure and characteristics, including the game's goals, mechanics, and graphical user interface elements, were integrated and fine-tuned from a practical standpoint to form a design of the game-based training system prototype. Moreover, the progression of game difficulty was divided into 3 difficulty levels (ie, beginner, intermediate, and advanced) to ensure that the difficulty of the games was appropriate for each user. After that, the game blueprint underwent a critical review (using the assessment form found in Multimedia Appendix 1) by 5 experts in the field of cognitive and physical interventions for older adults, including 2 physical therapists, 1 occupational therapist, 1 neurologist, and 1 geriatrician, with ages ranging from 40 to 56 years and experience ranging from 18 to 34 years. The content validity of the exergames was analyzed using the index of item-objective congruence (IOC). The cutoff value is a flexible criterion, with a generally accepted minimum typically being 0.75 [34].

Phase III: Development Phase

Consensus on the game idea and design, along with comments and feedback from the experts during the design phase, was used to refine the game-based exercise prototype in the development phase. This prototype was created using the Unity 3D game engine software, incorporating computer vision techniques from BlazePose for 2D/3D pose estimation [35]. The landmarks of the BlazePose 2D/3D pose estimation, which include 33 key points, are provided in Multimedia Appendix 2. In this study, only 12 key points (of 33 points) were utilized to control the mechanics of the games. These key points included the following: (1) left_shoulder, (2) right_shoulder, (3) left_wrist, (4) right_wrist, (5) left_pinky, (6) right_pinky, (7) left_index, (8) right_index, (9) left_thumb, (10) right_thumb, (11) left_hip, and (12) right_hip. The various conditions for controlling the game's avatar are displayed in Figure 2.

Figure 2. Condition of each landmark for control of the avatar in the game. (A) and (B) Moving bilateral arms horizontally to swim in a frog stroke style. (C) Dropping bilateral arms to rest. (D) and (E) Moving bilateral arms as if running to control the forward movement. (F) Raising the left or right hand over the head to collect objects. (G) Raising the right hand over the head to select choice A. (H) Raising the left hand over the head to select choice B. (I) Raising both hands over the head to select choice C.



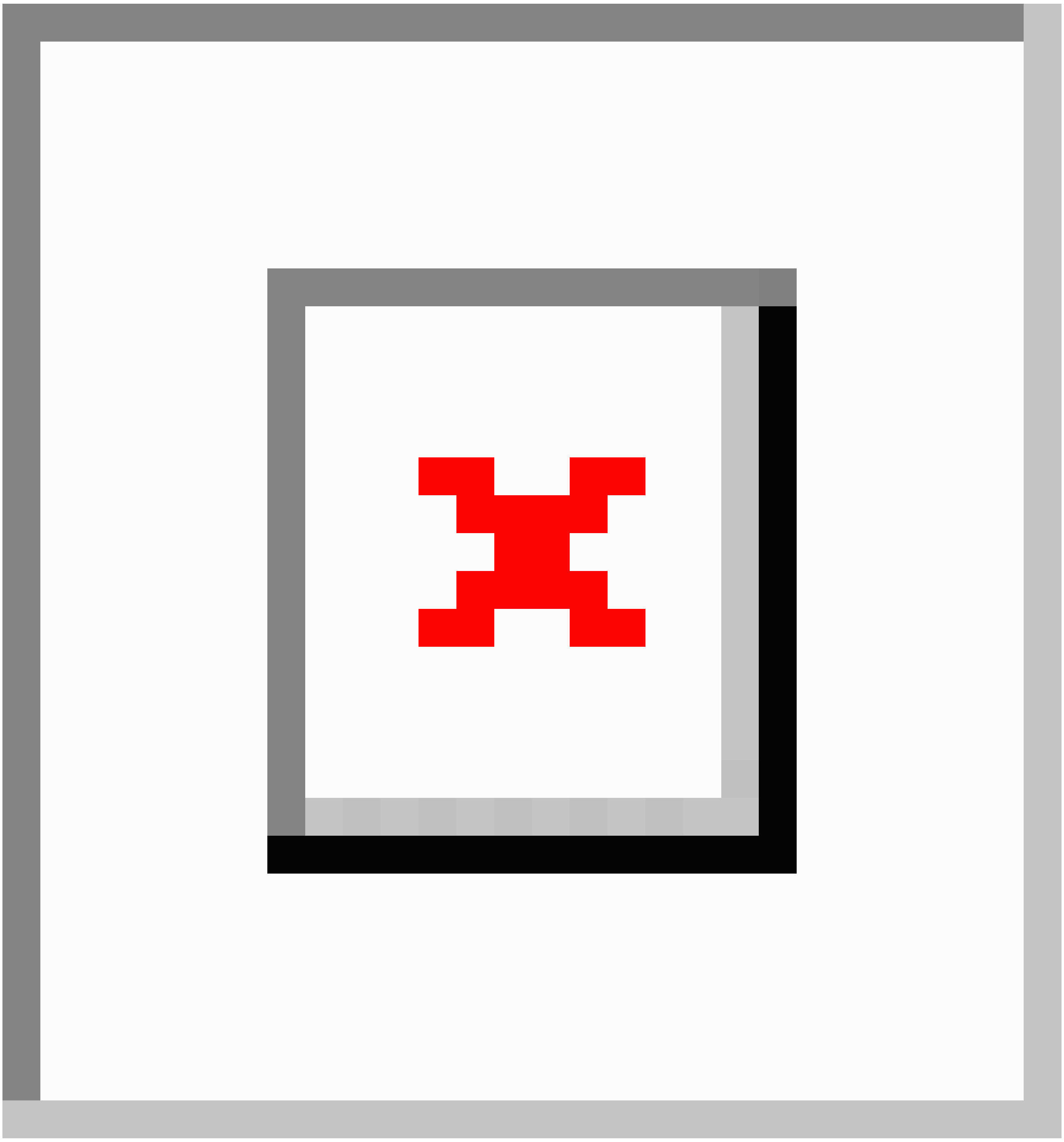
After completing phases I-III, the overall framework of the gamification-based, interactive physical-cognitive training system prototype, consisting of 5 components, was established (Figure 3):

1. User and the notebook with high-definition (HD) webcam sensor: The full-body movements were tracked in 3D coordinates (ie, the x-, y-, and z-axes) using 12 key points (Figure 2). The HD webcam sensor was used to track the full-body movements of the participants while they played the exergame (Figure 3A).
2. The gamification-based physical-cognitive training system: The digital game system consisted of 3 exergames (Game I: Ocean Diver; Game II: Road Runner; and Game III: Moving and Memorizing) with 3 levels of difficulty (beginner, intermediate, and advanced) and feedback (score, time, and error). Moreover, rules, goals, rewards, points, aesthetics, leaderboard, and game mechanics were used as elements of the game. The game was developed using the Unity game engine (Figure 3B).
3. Game programmer: The specialist who created the game using computer programming languages and game engine

software, including level design, coding, game development, and game testing (Figure 3C).

4. Game artist: The specialist who designed and created graphical content, including the graphical user interface, character, scenes, background, sound, and music (Figure 3D).
5. Domain knowledge: The experts with core knowledge in physical and cognitive training programs for older adults, both those with and without cognitive impairment (Figure 3E).

Figure 3. Gamification-based, interactive physical-cognitive training framework: (**A**) user interface and laptop with a camera, (**B**) system of gamification-based physical-cognitive training, (**C**) game programmer, (**D**) game artist, and (**E**) domain experts.



Phase IV: Implementation Phase

Overview

In this phase, the prototype of a low-intensity, gamification-based, interactive physical-cognitive exercise was implemented to representative target users to evaluate its usability and ability to promote enjoyment. The prototype of the exergame was delivered via the Windows platform and devices, including notebooks and personal computers.

Recruitment and Participants

A total of 15 community-dwelling older adults were recruited from community groups such as local senior schools and senior clubs, as well as through social media advertisements. Inclusion criteria were as follows: (1) age 60 years or older, (2) ability to comprehend instructions and willingness to participate, (3) adequate vision and hearing, or correction for impairments, and (4) sufficient physical performance, indicated by scoring at least 9/12 on the Short Physical Performance Battery, to ensure safety during exercise, particularly to prevent loss of balance and falls (Multimedia Appendix 3) [36]. Exclusion criteria were as follows: (1) being diagnosed with other neuromuscular conditions that affect cognitive function and physical ability (eg, stroke, Parkinson disease, dementia, and multiple sclerosis), and (2) having major health conditions that could not be controlled (eg, acute joint pain, asthma, hypertension, diabetes mellitus, and coronary artery disease).

Procedures

System Setup and Calibration

Prior to participants beginning the interactive exergame training at their homes, the researchers made a home visit for the system's setup, calibration, and operation. The capture volume was configured using a notebook or personal computer equipped with an HD webcam. The HD webcam was positioned on a table approximately 0.8 meters above the floor and about 2.5 meters away from the participant. The capture volume from left to right was set at approximately 3.0 meters, with the center being 1.5 meters from each side. The game was displayed on either notebook screens or computer monitors. The configuration of hardware for playing the game is provided in Multimedia Appendix 4. The game-based exercise was connected to the HD webcam sensor in the notebook or personal computer. During the calibration process, participants were instructed to move the avatar within the capture volume, with the boundaries marked by markers on the floor. Once the system had been fully set up, participants were able to use it independently.

User Experience Testing

The demographic information of eligible participants—including age, gender, BMI, and education level—was documented. The game goals, rules, mechanics, and controls were explained to the participants and they were provided with a demonstration. The participants interacted with the virtual game using their body and hand movements following each game's rules (ie, Ocean Diver, Road Runner, and Moving and Memorizing). The training duration was 35 minutes, which included warm-up and cool-down periods, with a rest interval of approximately 5 minutes, resulting in a total session time of around 60 minutes. Participants were allowed to progress to the next level if they achieved a score exceeding 50% in the current level. To assess the usability and enjoyment of the prototype, participants engaged in game-based training at their homes 3 times per week for 4 consecutive weeks (a total of 12 sessions).

Phase V: Evaluation Phase

Participants were asked to provide ratings for both the system's usability and their enjoyment during engagement with the game-based prototype after completing the 4-week testing period. The usability of the game-based physical-cognitive training system was assessed using the system usability scale (SUS) at the end of the training period [37]. The SUS rating scale was then converted from the original score (between 0 and 60) to an SUS score ranging from 0 to 100, where an SUS score above 68 is considered an above average ("good") or marginal acceptance level, and scores above 85 are considered an "excellent" acceptance level [38]. Additionally, the level of enjoyment during exercise engagement was assessed after the first and fourth week of training using the 8-item Physical Activity Enjoyment Scale (PACES) questionnaire, a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) [38]. The sum of PACES scores was calculated, with a higher score indicating a greater level of enjoyment during exercise engagement. Moreover, participants were interviewed about their perceptions of various aspects of the exergame during gameplay using probing questions (Multimedia Appendix 5). In the interview, assessors asked core questions sequentially and documented the responses. Upon completing the questioning, assessors summarized the key insights from the interview to ensure a comprehensive understanding of the participant's responses. Afterward, assessors extracted key themes from the responses provided by all participants to gather both positive and negative feedback, as well as information for improvements and refinements.

Ethical Considerations

The research protocol received approval from the Human Ethical Review Board of the Chiang Mai University (approval number AMSEC-66EX-036). All participants were informed about the study's purpose and procedure before providing written informed consent to enroll in the study. All data used for this study were anonymized. No compensation was provided to the participants.

Statistical Analysis

The content validity of the exergames, as assessed by experts, was determined using the IOC, with an IOC value greater than 0.75 indicating an acceptable level of validity. Descriptive statistics were used to illustrate the demographic profile of the participants and the SUS score after 4 weeks. The Wilcoxon signed-rank test was used to address differences in the PACES scores between the first week and the end of the 4-week training period. The significance level was set at $P<.05$. Analysis was conducted using SPSS (version 21.0; IBM Corp).

Results

Phase I-III: Analysis, Design, and Development Phases

The game-based training system underwent a critical review by 5 experts in the field of cognitive and physical interventions for older adults. The results indicated that the content validity was highly acceptable, with IOC values ranging from 0.80 to 1.00. The summary of the consensus features of the low-intensity, gamification-based, interactive physical-cognitive exercise is presented in Table 1 and Textbox 1. This system consisted of 3 exergames (Ocean Diver, Road Runner, and Moving and Memorizing), all aimed at enhancing simultaneous physical and cognitive function in older adults. Each exergame had 3 levels of difficulty (beginner, intermediate, and advanced), in which game complexity progressed by increasing the complexity of physical and cognitive demands. The total game play time was 35 minutes. In accordance with exercise principles, warm-up and cool-down sessions were included as part of the training. The exergame prototype was delivered via a computer vision–based system that participants could operate independently in a home-based setting.

Table . Summary of the features of the developed gamification-based, interactive physical-cognitive training.

	Game I: Ocean Diver	Game II: Road Runner	Game III: Moving and Memorizing
Actions	Swimming in a frog stroke style underwater to collect predetermined objects in the correct order and collect coins as a bonus	Running along the road to collect as many predetermined objects as possible and collect coins as a bonus	Walking along the road to collect coins while listening to a story, then answering questions related to the story at the end of the game
Rules and game mechanics	• Moving bilateral arms in a frog stroke style to control the avatar's upward or downward direction • Stepping to the left or right to control the avatar's left or right movement direction	• Moving bilateral arms as if running to control the avatar's forward movement • Stepping to the left or right to control the avatar's left or right movement direction • Raising the left or right arm to control the avatar to collect predetermined objects	• Stepping to the left or right to control the avatar's left or right movement direction • Listening to a story and collecting coins as a bonus • Answering questions related to the story at the end of the game
Virtual environment	An underwater world with aquatic animals and treasures	A suburban village with fruits, vegetables, and animals	A suburban village with coins
Goal of the game: physical components	• Improve dynamic balance and coordination of upper and lower limbs • Improve upper and lower limb muscle strength and endurance	• Improve dynamic balance and coordination of upper and lower limbs • Improve reaction and response time • Improve upper and lower limb muscle strength and endurance	• Improve dynamic balance • Improve lower limb muscle strength and endurance
Goal of the game: cognitive components	Improve memory, attention, and executive function (ie, planning, sequencing, and inhibiting)	Improve memory, attention, and visuospatial ability	Improve memory (delayed recall) and attention

Textbox 1. Framework of the gamification-based physical-cognitive training.

- Physical component training progression: the difficulty levels progressed by increasing the subject's movement speed of upper and lower limbs (eg, adding the number of objects).
- Cognitive component training progression: the difficulty levels progressed by increasing the subject's cognitive demands to play the game (eg, adding memory and attention requirements, increasing the speed of processing, and adding the number of objects).
- Rewards: points and coins
- Estimated playtime: 35 minutes
- Health topic: gamification-based physical-cognitive training prototype
- Targeted age group: older adults (age ≥60 years)
- Short description of the gamification-based training: gamification-based physical-cognitive training is an interactive game-based training system for older adults using high-definition webcam sensor technology. It comprises 3 games involving physical training (ie, dynamic balance; coordination; reaction and response time; and upper and lower limb muscle strength and endurance) and cognitive training (ie, memory or delayed recall, attention, and executive function including planning, sequencing, inhibiting, visuospatial ability.
- End user or target player: older adults (individual self-training)
- Clinical support needed: physical therapists and related geriatrics health care professionals
- Data shared with clinician: data are saved and stored on the hard disk; however, points and scores are given as feedback on the laptop monitor at the end of each game.
- Type of game: physical, cognitive, action, real-time strategy
- Behavior change procedure used: gamification-based physical-cognitive training enhances extrinsic motivation and engagement in older adults.
- Setting: gamification-based physical-cognitive training can be set in a room environment.
- Device requirements: PC/notebook/laptop with HD webcam
- Sensors used: HD webcam

Phase IV-V: Implementation and Evaluation Phases

Participant Characteristics

In total, 15 community-dwelling older adults aged ≥60 years took part in the study; of the participants, 8 were female and 7 were male. None of the participants had prior experience with game-based training. The demographic characteristics of the participants are presented in Table 2. All participants completed the training program (mean number of exercise sessions 11.07, SD 1.95, with a total of 12 sessions). No falls or other adverse events were reported.

Table . Characteristics of the study participants.

Characteristics	Mean (SD)	Range (minimum-maximum)
Age (years)	65.27 (4.40)	64.00 - 77.00
BMI (kg/m^2)	24.09 (2.59)	20.03 - 28.73
Education level (years)	15.67 (2.58)	9.00 - 21.00

User Experience in Using the Gamification-Based Exercise System

The user experience in using the game-based exercise was determined using the SUS questionnaire and the PACES. The overall results of the SUS score are illustrated in Figure 4. The mean SUS score was 87.22 (SD 5.76), which corresponds to an A or "excellent to highest possible" acceptance level (based on the grade scale and adjective rating; Figure 4A). Further data exploration showed that, of the 15 participants, 6 rated this gamification-based exercise system prototype as "good," 5 rated it as "excellent," and 4 rated it as "highest possible." The mean SUS score for each question ranged from 72.22 (question item 6) to 98.89 (question item 3), signifying ratings from a marginal "good" to the "highest possible" acceptance level (Figure 4B).

As for the 8-item PACES score, the total PACES score of all items at the end of the first week of training was 44.93 (SD 3.99); the average on each item ranged from 5.27 to 5.93. After 4 weeks, it was 50.53 (SD 4.70); the average on each item ranged from 5.93 to 6.67 (Table 3). A Wilcoxon signed-rank test demonstrated that both the average on each item and total score after the 4-week training were significantly greater than those after 1 week of training (P=.001), indicating an improvement in enjoyment and satisfaction after exercise engagement for 4 consecutive weeks.

Figure 4. Overview of usability results of the game-based exercise prototype using the SUS questionnaire at the end of the 4-week testing period. (A) The SUS score and its relationship with the grade scale (F, D, C, B, and A), adjective ratings (lowest possible, poor, okay, good, excellent, and highest possible), and acceptability scores (with marginal acceptability at 68.0). (B) The SUS score for each question. SUS: system usability scale.

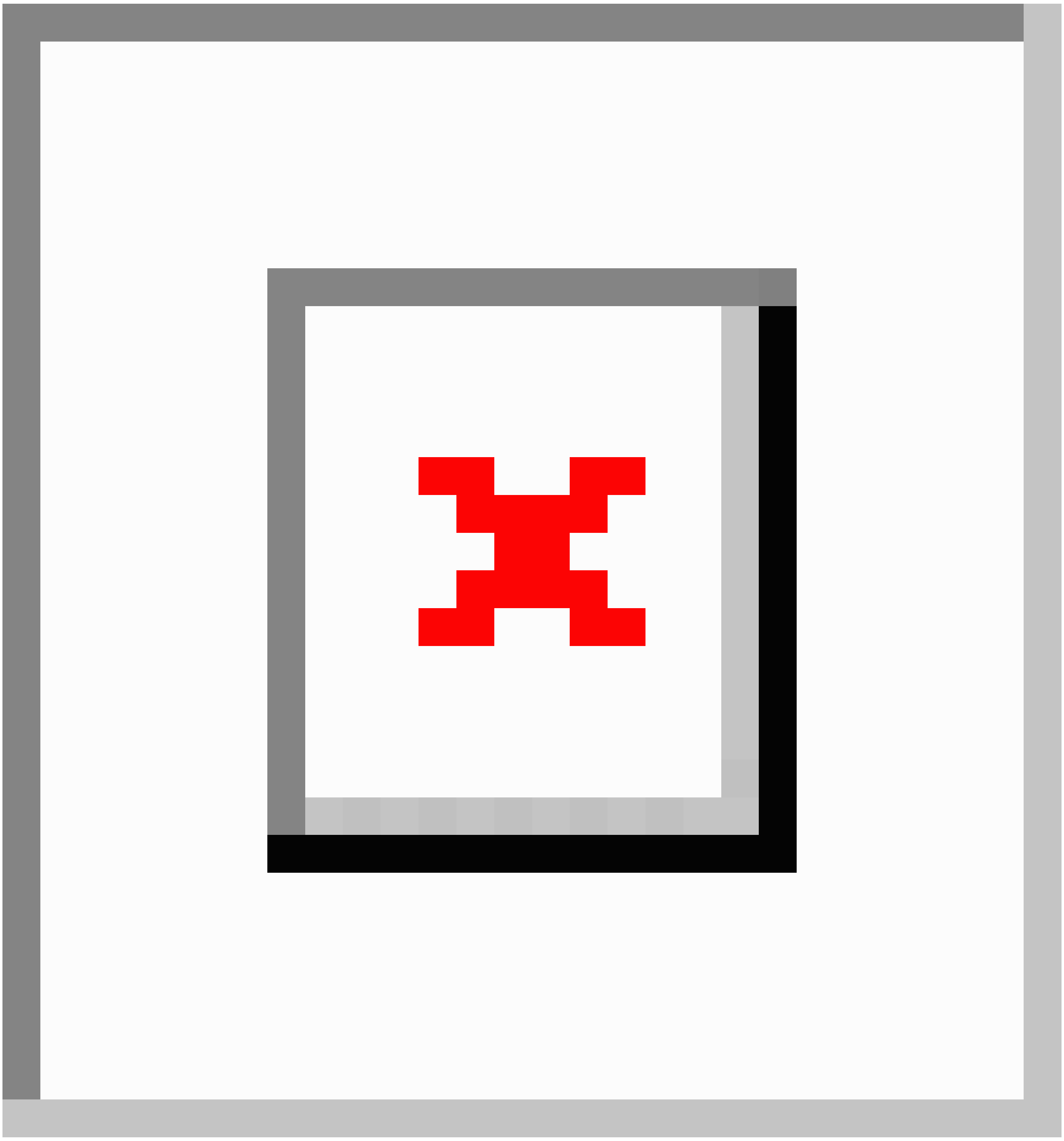


Table . The PACES score after 1 week of training and after completion of 4 weeks of training (n=15). All items were rated on a 7-point scale from 1 (strongly disagree) to 7 (strongly agree).

Question items	Week 1, mean (SD)	Week 4, mean (SD)	*P* value[a]
1. I find it pleasurable	5.53 (0.83)	6.33 (0.82)	.003
2. It's a lot of fun	5.27 (0.88)	5.93 (0.88)	.02
3. It's very pleasant	5.27 (0.88)	6.27 (0.88)	.002
4. It's very invigorating	5.60 (0.63)	6.13 (0.83)	.01
5. It's very gratifying	5.67 (0.62)	6.27 (0.70)	.01
6. It's very exhilarating	5.73 (0.59)	6.40 (0.74)	.004
7. It's very stimulating	5.93 (0.59)	6.53 (0.52)	.01
8. It's very refreshing	5.93 (0.59)	6.67 (0.62)	.005
Total scale of all items (56 points)	44.93 (3.99)	50.53 (4.70)	.001

[a]Wilcoxon signed-rank test with significant difference at $P<.05$.

User Feedback of the Gamification-Based Exercise System

Comments and feedback from the participants on the developed gamification-based exercise system prototype through probing interviews are summarized for each question domain in Table 4.

Table . The key responses of participants regarding the game characteristics and their perceptions during engagement with the game-based prototype.

Question domains	Positive feedback	Negative feedback
Game mechanics, rules, and interface	• The storytelling game (Moving and Memorizing) was enjoyable, with its content, purpose, and display perfectly suited to personal preferences • The Ocean Diver game was very enjoyable and well-designed, providing a refreshing experience	• It is hard to judge how far to jump to grab animals or objects on time • Sometimes in the game, the movements did not align with the actual actions • The characters and objects displayed in the game were relatively small • Sometimes, the game system froze
Game instructions	Instructions for the game were clear and easy to understand	No negative feedback
Gameplay experience	• The exergame was very helpful as it promoted physical activity and memory • The exergame could be played at any time, which perfectly fits the schedule • There was a noticeable improvement in sleep quality	• In the Ocean Diver game, some objects that needed to be collected were placed in tight spots, making it stressful when participants were unable to pick them up
Exercise dosage	The exercise was not too difficult, as it allowed me to keep up with the training	In the first week of training, the legs were quite fatigued; however, after the first week, it went away
Feedback for improvements	Overall, the games were satisfying, enhancing both enjoyment and physical activity	Making the movements in the game better match the user's actions, increasing the size of the characters and objects, and reducing the game system's freezing issues

Discussion

Principal Findings

This study aimed to design, develop, and determine the user experience of a prototype of low-intensity, gamification-based training targeted at improving physical and cognitive performance in older adults within a home environment using the ADDIE model. The idea of the core concept and platform of the exergame were extracted by integrating the knowledge with well-recognized physical and cognitive outcomes in older adults, considering practical standpoints, and then soliciting critical appraisal by experts. The content validity of the exergames was highly acceptable, demonstrating that the game's features, rules, and exercise components were well-aligned with both the intended therapeutic goals and practical considerations. The experiences of older adults in terms of usability and enjoyment of the exergame prototype were assessed and their feedback was collected.

The output of this study was the prototype of a gamification-based, interactive physical-cognitive training program in a home-based setting, which consisted of 3 exergames: Ocean Diver, Road Runner, and Moving and Memorizing. Each game had 3 levels of difficulty: beginner, intermediate, and advanced. The user-perceived usability of the game-based exercise prototype was rated as good to excellent, and there were improvements in enjoyment and satisfaction after exercise engagement over 4 consecutive weeks. Additionally, the overall end user feedback indicated that the motivation for engaging in an exergame is its inherent interest and enjoyability.

Comparison to Prior Work

In recent years, an increasing number of studies have developed and evaluated the effects of exergames on both physical and cognitive functions in older adults. However, the majority of these studies have used moderate-to-high physical intensity levels [19,39,40], which are generally suitable for relatively healthy individuals without significant physical restrictions. Only a limited number of studies focused on developing low-intensity exergames, which may be more practical for older adults who are physically limited or less active, as they require low exertion and low-impact exercise modes. To the best of our knowledge, our game-based exercise is the first prototype of low-intensity, gamification-based, interactive physical-cognitive training specifically designed to enhance both physical and cognitive abilities in older adults. We found that the exergame prototype was usable in a home environment and effectively enhanced the enjoyment and motivation of older adults. The feedback also indicated potential future applications for self-training at home. These findings are in line with previous studies suggesting that computer-based interventions, a delivery platform used in this study, offer a practical, safe, and efficient approach to encouraging adherence to exercise [41,42].

Currently, various approaches to product development may involve end users at different stages of the production process, depending on the specific product requirements and goals. Among various product development approaches, the ADDIE model was selected because it aligns well with our exergame product, which is relatively new and unfamiliar to older adults. This model involves experts in the analysis and design phases (phases I and II) and end users in the implementation and evaluation phases (phases IV and V). Given that experts have knowledge and experience in the fields of physical and cognitive rehabilitation for older adults, as well as game development, they played a key role in the analysis and design phases, which are crucial for achieving both theoretical accuracy and practical feasibility to ensure high-quality outcomes. A representative group of the end users (older adults) tested the exergame prototype and provided feedback during the implementation and evaluation phases. The feedback and suggestions from end users were essential in shaping the final product, guaranteeing that the prototype was closely tailored to their needs and preferences. A previous study suggested that individualization of the exercise in terms of needs and interests is necessary to enhance individual engagement, thereby contributing to positive intervention outcomes [43]. In addition, this study utilized a computer vision–based system as a technology platform to deliver an exercise program (in the form of an exergame) in a home-based environment, aiming to provide accessibility and ease of integration into users' lives. Previous findings have suggested that home-based exercise exerts greater exercise engagement than center-based exercise due to its accessibility and flexibility in the schedule, allowing individuals to integrate it conveniently into their daily life [44,45]. Taken together, the benefits of using the ADDIE model approach and delivering the platform as a computer vision–based system for the development of an exergame prototype may represent a potential strategy for enhancing the enjoyment, motivation, and engagement of older adults in exercise in a home-based setting.

Regarding the user experience with the game-based training system, participants reported increased enjoyment, as indicated by their PACES scores, and provided positive feedback on the enhancement of both their enjoyment and physical activity. The increase in the enjoyment level of participants may, at least in part, be attributed to several fundamental aspects of the game. These include game elements that provide real-time feedback (ie, scores and rewards), game mechanics (ie, real-time interface and sound effects), and game rules (optimal difficulty level, grading from simple to advanced), which in turn results in high enjoyment and self-efficacy empowerment. Our findings are supported by prior research, which suggests that intrasession feedback, positive reward, and the use of graded exercise may enhance intrinsic motivation and self-confidence in the capability to exercise, potentially reinforcing repetitive desirable behavior [46,47]. Some negative feedback on the game's characteristics was received, including issues with asynchronization between user actions and the display, system freezing, and the need to enlarge characters and objects. Thus, additional adjustments to the game mechanics are necessary prior to its deployment for end users. Apart from the game characteristics, the usability of a game-based system is a key aspect in enhancing acceptability and exercise adherence. As we found, the users' perception of the usability of the developed game-based training system, as determined by the SUS score, ranged from good to highest possible. The positive response from usability testing may potentially be due to the game interface and operations being specifically designed to be used independently by older adults. Consistently, a previous study has suggested that the feature design of exergames, particularly an interface that is friendly to older users, is a crucial factor contributing to the positive acceptance of older users who are unfamiliar with new technology [48]. Nonetheless, the lowest scores on the SUS questionnaire were found in items 5 and 6 (scoring 74.44 and 72.22, respectively), which are related to game mechanics and consistent with feedback for improvements. Therefore, further refinement of the exergame system is warranted before implementation for end users, particularly concerning technical errors and game mechanics issues encountered by participants.

Strengths and Limitations

To the best of our knowledge, this game-based home exercise is the first prototype customized for older adults with low physical capacity, aiming to simultaneously stimulate specific cognitive domains and physical components. Despite the positive results, this study has certain limitations that need to be addressed. First, participants in this study had relatively high educational backgrounds and were socially active, which may introduce potential bias, as they may be more familiar with technology. Further, they did not exhibit obvious physical restrictions as the study inclusion/exclusion criteria were set to minimize potential adverse events during participation. Second, the sample size was relatively small. Together, the findings may not be generalizable to a broader older adult population and should be considered preliminary, warranting cautious interpretation. Third, exercise adherence was observed over a 4-week training period; hence, long-term adherence remains unknown. Last, since this study focused on developing and assessing the usability of an exergame prototype for home-based settings, hard outcomes, such as physical and cognitive parameters in response to exergame engagement, have not yet been established. Further research that includes larger sample sizes, diverse demographic backgrounds, extended training periods, and examination of effectiveness through comprehensive outcomes is needed.

Future Directions

The outcome of this study is the prototype of a home-based, low-intensity exercise program that is practical and has the potential to enhance the enjoyment and motivation of older adults. The next step will be to refine the game-based prototype based on end user critical feedback and examine its effectiveness in promoting the physical and cognitive functions of community-dwelling older adults, thereby giving this new type of exergame a promising future.

Conclusions

In this study, a low-intensity, gamification-based, interactive physical-cognitive training system was developed for older adults with limited capacity to engage in moderate-to-high intensity exercise. The exergame prototype, delivered via a computer vision–based platform for a home-based, self-training exercise, was well-received by the end users for its usability and enjoyment. Although these findings hold promise for implementing the exergame in the target population, further research is warranted to determine its effectiveness in promoting physical and cognitive functions.

Acknowledgments

The authors would like to thank Nathapong Metawatcharekul and Thanet Galong for their contribution to the development of the exergames, and Poosri T for her assistance with data collection. The authors would also like to thank participants for their engagement and feedback. Special thanks to all the experts for their comprehensive comments in the development of the exergame prototype. This study has been funded by the National Research Council of Thailand (NRCT; N42A660640) and the Faculty of Associated Medical Sciences, Chiang Mai University (author SS).

Data Availability

The datasets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

KT, KS, MP, BS, and SS contributed to the study conception and idea, experimental design, data collection, and approval of the final manuscript for publication. KT was responsible for intellectual contributions related to gamification for health, research planning, data analysis and interpretation of the results, and drafting the manuscript. SS and BS were responsible for intellectual contributions related to exercise sciences for older adults, data analysis and interpretation of the results, drafting the manuscript, and reviewing the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Content validity assessment form for experts.
[PDF File, 265 KB - games_v12i1e59141_app1.pdf]

Multimedia Appendix 2
Landmark of BlazePose 2D/3D pose estimation.
[PNG File, 132 KB - games_v12i1e59141_app2.png]

Multimedia Appendix 3
Short Physical Performance Battery scoring sheet.
[PDF File, 349 KB - games_v12i1e59141_app3.pdf]

Multimedia Appendix 4

Environmental configuration of the gamification-based interactive physical-cognitive training system.
[PNG File, 611 KB - games_v12i1e59141_app4.png]

Multimedia Appendix 5
The probing questions on perceptions of the game characteristics and user experiences during engagement with the game-based training system prototype.
[PDF File, 64 KB - games_v12i1e59141_app5.pdf]

References

1. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]
2. Zhang W, Low LF, Schwenk M, Mills N, Gwynn JD, Clemson L. Review of gait, cognition, and fall risks with implications for fall prevention in older adults with dementia. Dement Geriatr Cogn Disord 2019;48(1-2):17-29. [doi: 10.1159/000504340] [Medline: 31743907]
3. Sherrington C, Michaleff ZA, Fairhall N, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 2017 Dec;51(24):1750-1758. [doi: 10.1136/bjsports-2016-096547] [Medline: 27707740]
4. Sherrington C, Fairhall NJ, Wallbank GK. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424. [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
5. Öhman H, Savikko N, Strandberg TE, Pitkälä KH. Effect of physical exercise on cognitive performance in older adults with mild cognitive impairment or dementia: a systematic review. Dement Geriatr Cogn Disord 2014;38(5-6):347-365. [doi: 10.1159/000365388] [Medline: 25171577]
6. American College of Sports Medicine, Chodzko-Zajko WJ, Proctor DN, et al. American College of Sports Medicine position stand. Exercise and physical activity for older adults. Med Sci Sports Exerc 2009 Jul;41(7):1510-1530. [doi: 10.1249/MSS.0b013e3181a0c95c] [Medline: 19516148]
7. Bamidis PD, Vivas AB, Styliadis C, et al. A review of physical and cognitive interventions in aging. Neurosci Biobehav Rev 2014 Jul;44:206-220. [doi: 10.1016/j.neubiorev.2014.03.019] [Medline: 24705268]
8. Gavelin HM, Dong C, Minkov R, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
9. Theill N, Schumacher V, Adelsberger R, Martin M, Jäncke L. Effects of simultaneously performed cognitive and physical training in older adults. BMC Neurosci 2013 Sep 23;14:103. [doi: 10.1186/1471-2202-14-103] [Medline: 24053148]
10. Suwabe K, Byun K, Hyodo K, et al. Rapid stimulation of human dentate gyrus function with acute mild exercise. Proc Natl Acad Sci USA 2018 Oct 9;115(41):10487-10492. [doi: 10.1073/pnas.1805668115] [Medline: 30249651]
11. Wang R, Zhang H, Li H, et al. The influence of exercise interventions on cognitive functions in patients with amnestic mild cognitive impairment: a systematic review and meta-analysis. Front Public Health 2022;10:1046841. [doi: 10.3389/fpubh.2022.1046841] [Medline: 36457329]
12. Gothe NP. Examining the effects of light versus moderate to vigorous physical activity on cognitive function in African American adults. Aging Ment Health 2021 Sep;25(9):1659-1665. [doi: 10.1080/13607863.2020.1768216] [Medline: 32426993]
13. Cai H, Li G, Jiang S, Yin H, Liu P, Chen L. Effect of low-intensity, Kinect-based Kaimai-style Qigong exercise in older adults with type 2 diabetes. J Gerontol Nurs 2019 Feb 1;45(2):42-52. [doi: 10.3928/00989134-20190111-05] [Medline: 30690653]
14. Robson K, Plangger K, Kietzmann JH, McCarthy I, Pitt L. Is it all a game? Understanding the principles of gamification. Bus Horiz 2015 Jul;58(4):411-420. [doi: 10.1016/j.bushor.2015.03.006]
15. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]
16. Sun TL, Lee CH. An impact study of the design of exergaming parameters on body intensity from objective and gameplay-based player experience perspectives, based on balance training exergame. PLoS One 2013;8(7):e69471. [doi: 10.1371/journal.pone.0069471] [Medline: 23922716]
17. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
18. Stanmore E, Stubbs B, Vancampfort D, de Bruin ED, Firth J. The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials. Neurosci Biobehav Rev 2017 Jul;78:34-43. [doi: 10.1016/j.neubiorev.2017.04.011] [Medline: 28442405]

19. Phirom K, Kamnardsiri T, Sungkarat S. Beneficial effects of interactive physical-cognitive game-based training on fall risk and cognitive performance of older adults. Int J Environ Res Public Health 2020 Aug 21;17(17):6079. [doi: 10.3390/ijerph17176079] [Medline: 32825555]
20. Zhao Y, Feng H, Wu X, et al. Effectiveness of exergaming in improving cognitive and physical function in people with mild cognitive impairment or dementia: systematic review. JMIR Serious Games 2020 Jun 30;8(2):e16841. [doi: 10.2196/16841] [Medline: 32602841]
21. Zhao C, Zhao C, Zhao M, et al. Effect of exergame training on working memory and executive function in older adults. Sustainability 2022;14(17):10631. [doi: 10.3390/su141710631]
22. Mellecker R, Lyons EJ, Baranowski T. Disentangling fun and enjoyment in exergames using an expanded design, play, experience framework: a narrative review. Games Health J 2013 Jun;2(3):142-149. [doi: 10.1089/g4h.2013.0022] [Medline: 24761322]
23. Gong W, Zhang X, Gonzàlez J, et al. Human pose estimation from monocular images: a comprehensive survey. Sensors (Basel) 2016 Nov 25;16(12):1966. [doi: 10.3390/s16121966] [Medline: 27898003]
24. Samkari E, Arif M, Alghamdi M, Al Ghamdi MA. Human pose estimation using deep learning: a systematic literature review. MAKE 2023;5(4):1612-1659. [doi: 10.3390/make5040081]
25. Vonstad EK, Su X, Vereijken B, Bach K, Nilsen JH. Comparison of a deep learning-based pose estimation system to marker-based and Kinect systems in exergaming for balance training. Sensors (Basel) 2020 Dec 4;20(23):6940. [doi: 10.3390/s20236940] [Medline: 33291687]
26. Szabo DA. Adapting the ADDIE instructional design model in online education. Stud UBB Psychol-Paed 2022 Jun 30;67(1):125-140 [FREE Full text] [doi: 10.24193/subbpsyped.2022.1.08]
27. Aldoobie N. ADDIE model. Am Int J Contemp Res 2015;5(6):68-72 [FREE Full text] [doi: 10.30845/aijcr]
28. Makmee P, Wongupparaj P. Virtual reality-based cognitive intervention for enhancing executive functions in community-dwelling older adults. Psychosoc Interv 2022 Jul;31(3):133-144. [doi: 10.5093/pi2022a10] [Medline: 37361011]
29. Kim D. Development and effect of virtual reality practice program for improving practical competency of caregivers specializing in dementia. Healthcare (Basel) 2021 Oct 18;9(10):1390. [doi: 10.3390/healthcare9101390] [Medline: 34683070]
30. Choi HK, Lee K, Lee SH. Developmental study on “Smart Silver Care”: a mobile application to alleviate loneliness in older adults within the community. Healthcare (Basel) 2023 Aug 23;11(17):2376. [doi: 10.3390/healthcare11172376] [Medline: 37685410]
31. Ae-Ri J, Kowoon L, Eun-A P. Development and evaluation of the information and communication technology-based Loneliness Alleviation Program for community-dwelling older adults: a pilot study and randomized controlled trial. Geriatr Nurs 2023;53:204-211. [doi: 10.1016/j.gerinurse.2023.07.011] [Medline: 37544264]
32. Bicen A, Uzunboylu H, Burgul NS. Evaluation of participants’ opinions on online physical fitness training. J Sport Psychol 2020;29(4):25-32 [FREE Full text]
33. Amboni M, Barone P, Hausdorff JM. Cognitive contributions to gait and falls: evidence and implications. Mov Disord 2013 Sep 15;28(11):1520-1533. [doi: 10.1002/mds.25674] [Medline: 24132840]
34. Turner RC, Carlson L. Indexes of item-objective congruence for multidimensional items. Int J Test 2003 Jun;3(2):163-171. [doi: 10.1207/S15327574IJT0302_5]
35. Bazarevsky V, Grishchenko I, Raveendran K, Zhu TL, Zhang F, Grundmann M. BlazePose: on-device real-time body pose tracking. ArXiv. Preprint posted online on Jun 17, 2020. [doi: 10.48550/arXiv.2006.10204]
36. de Fátima Ribeiro Silva C, Ohara DG, Matos AP, Pinto A, Pegorari MS. Short physical performance battery as a measure of physical performance and mortality predictor in older adults: a comprehensive literature review. Int J Environ Res Public Health 2021 Oct 10;18(20):10612. [doi: 10.3390/ijerph182010612] [Medline: 34682359]
37. Brooke J. SUS-A quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry, 1st edition: CRC Press; 1996:189-194. [doi: 10.1201/9781498710411]
38. Mullen SP, Olson EA, Phillips SM, et al. Measuring enjoyment of physical activity in older adults: invariance of the Physical Activity Enjoyment Scale (PACES) across groups and time. Int J Behav Nutr Phys Act 2011 Sep 27;8:103. [doi: 10.1186/1479-5868-8-103] [Medline: 21951520]
39. Béraud-Peigné N, Maillot P, Perrot A. The effects of a new immersive multidomain training on cognitive, dual-task and physical functions in older adults. Geroscience 2024 Apr;46(2):1825-1841. [doi: 10.1007/s11357-023-00952-w] [Medline: 37743415]
40. Béraud-Peigné N, Maillot P, Perrot A. The user experience of an immersive and interactive wall exergame in older adults. Games Health J 2023 Jun;12(3):220-227. [doi: 10.1089/g4h.2022.0075] [Medline: 36251866]
41. Anggraini IT, Xiao Y, Funabiki N, Shih CL, Huang WC, Fan CP. A hand gesture exergame for dementia development suppression and its applications to elders. Presented at: 2024 IEEE International Conference on Consumer Electronics (ICCE); Jan 5-8, 2024; Las Vegas, NV, USA p. 1-3. [doi: 10.1109/ICCE59016.2024.10444306]
42. Goumopoulos C, Drakakis E, Gklavakis D. Feasibility and acceptance of augmented and virtual reality exergames to train motor and cognitive skills of elderly. Comp 2023;12(3):52. [doi: 10.3390/computers12030052]

43. Vseteckova J, Deepak-Gopinath M, Borgstrom E, et al. Barriers and facilitators to adherence to group exercise in institutionalized older people living with dementia: a systematic review. Eur Rev Aging Phys Act 2018;15:11. [doi: 10.1186/s11556-018-0200-3] [Medline: 30455778]
44. Allen NE, Sherrington C, Suriyarachchi GD, Paul SS, Song J, Canning CG. Exercise and motor training in people with Parkinson's disease: a systematic review of participant characteristics, intervention delivery, retention rates, adherence, and adverse events in clinical trials. Parkinsons Dis 2012;2012:854328. [doi: 10.1155/2012/854328] [Medline: 22191077]
45. Ashworth NL, Chad KE, Harrison EL, Reeder BA, Marshall SC. Home versus center based physical activity programs in older adults. Cochrane Database Syst Rev 2005 Jan 25;2005(1):CD004017. [doi: 10.1002/14651858.CD004017.pub2] [Medline: 15674925]
46. Nicolson PJA, Bennell KL, Dobson FL, Van Ginckel A, Holden MA, Hinman RS. Interventions to increase adherence to therapeutic exercise in older adults with low back pain and/or hip/knee osteoarthritis: a systematic review and meta-analysis. Br J Sports Med 2017 May;51(10):791-799. [doi: 10.1136/bjsports-2016-096458] [Medline: 28087567]
47. Dobson F, Bennell KL, French SD, et al. Barriers and facilitators to exercise participation in people with hip and/or knee osteoarthritis: synthesis of the literature using behavior change theory. Am J Phys Med Rehabil 2016 May;95(5):372-389. [doi: 10.1097/PHM.0000000000000448] [Medline: 26945211]
48. Lu MH, Lin W, Yueh HP. Development and evaluation of a cognitive training game for older people: a design-based approach. Front Psychol 2017;8:1837. [doi: 10.3389/fpsyg.2017.01837] [Medline: 29089914]

Abbreviations

ADDIE: analysis, design, development, implementation, and evaluation
HD: high definition
IOC: index of item-objective congruence
PACES: Physical Activity Enjoyment Scale
SUS: system usability scale

Edited by A Coristine; submitted 03.04.24; peer-reviewed by E Giannouli, S Moza; revised version received 22.09.24; accepted 23.09.24; published 29.10.24.

Please cite as:
Kamnardsiri T, Kumfu S, Munkhetvit P, Boripuntakul S, Sungkarat S
Home-Based, Low-Intensity, Gamification-Based, Interactive Physical-Cognitive Training for Older Adults Using the ADDIE Model: Design, Development, and Evaluation of User Experience
JMIR Serious Games 2024;12:e59141
URL: https://games.jmir.org/2024/1/e59141
doi:10.2196/59141

Examining the Feasibility, Acceptability, and Preliminary Efficacy of an Immersive Virtual Reality–Assisted Lower Limb Strength Training for Knee Osteoarthritis: Mixed Methods Pilot Randomized Controlled Trial

Hermione Hin Man Lo[1], MPH, MScN; Marques Ng[1], PhD; Pak Yiu Hugo Fong[2], MBBS, MSc; Harmony Hoi Ki Lai[2], MAGC; Bo Wang[2], EdD; Samuel Yeung-shan Wong[2], MPH, MD, CCFP; Regina Wing Shan Sit[2], MBBS, MD, DCH, DPD, PDip, DipMed

1

2

Corresponding Author:
Regina Wing Shan Sit, MBBS, MD, DCH, DPD, PDip, DipMed

Abstract

Background: Knee osteoarthritis prevalently causes significant pain, activity limitations, psychological distress, and reduced quality of life. Despite lower limb strength training being a core treatment for knee osteoarthritis, adherence remains a challenge, prompting the exploration of virtual reality (VR) to improve exercise compliance. Recent research suggests the potential of VR in providing enhanced pain management and functional outcomes for knee osteoarthritis, necessitating further exploration of immersive VR technology.

Objective: We aimed to study the feasibility, acceptability, and preliminary efficacy of an immersive VR-assisted lower limb strength training for knee osteoarthritis (VRiKnee).

Methods: A convergent, parallel, mixed methods study was conducted in 30 participants with knee osteoarthritis. After 1:1 randomization, the VRiKnee group (n=15) was assigned to perform repetitive concentric quadriceps and isometric vastus medialis oblique exercise in an immersive environment using a head-mounted display for 12 weeks. The control group (n=15) completed the same exercises without VRiKnee. VRiKnee participants were interviewed at week 12 to study VRiKnee acceptability and user experience. Quantitative data included feasibility outcomes such as recruitment, dropout, and exercise adherence rates, and effectiveness outcomes such as the numeric rating scale, the Western Ontario and McMaster Universities Osteoarthritis Index (100 points) pain and function subscales, and objective physical activity measured by metabolic equivalents of task using an ActivPAL accelerometer. Qualitative data were analyzed by thematic analysis, followed by integration with quantitative data using joint displays.

Results: The recruitment rate was 100% (N=30), with enrollment of 30 participants in 7.57 weeks. The median age was 63.5 (IQR 61.8 - 66.3) years, with 76% (n=23) being female. The response rates were 80% and 93% for the VRiKnee and control groups. Dropout rates were 13% for VRiKnee and 7% for the control group. Median exercise adherence was 77% (IQR 37-104%) for VRiKnee and 62% (IQR 40-166%) for the control group, respectively, with adherence reduction over this study's period and no significant intergroup differences (P=.82). No statistically significant differences were observed in primary and secondary outcomes, though positive trends were observed in pain and stiffness. Cybersickness was reported by 5 (33%) participants in the VRiKnee group. In the qualitative analysis, 4 themes, 11 subthemes, and 16 quotes were generated, identifying facilitators and barriers with practical suggestions to enhance the usability of VRiKnee.

Conclusions: VRiKnee demonstrated feasibility, acceptability, and potential efficacy in managing knee osteoarthritis. Future trials of larger sample sizes and better VR designs will confirm its role in clinical practice.

Trial Registration: Chinese Clinical Trial Registry CHiCTR2100046313; https://www.chictr.org.cn/showprojEN.html?proj=125404

(JMIR Serious Games 2024;12:e52563) doi:10.2196/52563

KEYWORDS

virtual reality; VR; immersive; knee; joint; arthritis; arthritic; osteoarthritis; knee osteoarthritis; gerontology; geriatric; older adult; elder; elderly; older person; older people; aging; pain; function; acceptability; user experience; RCT; randomized; controlled trial; limb; strength; muscle; muscular; physiotherapy

Introduction

Knee osteoarthritis is a common degenerative joint disease with a global prevalence of 22.9% in participants aged 40 years and older [1]. It is a major contributor to disability worldwide, causing considerable pain, activity limitations, psychological distress, and reduced quality of life for those affected [2-4]. According to Osteoarthritis Research Society International in 2019, land-based exercise is a core treatment for knee osteoarthritis, for which lower limbs strength training is often recommended [5-7]. In clinical practice, exercise programs often involve initial supervision by a clinician, followed by unsupervised home exercise. Ideally, regular participation in exercise should be one of the long-term goals for knee osteoarthritis; unfortunately, adherence to home exercise is often poor [8]. Therefore, strategies are needed to improve adherence to home exercise.

Recently, studies have suggested that technology advancements may increase attractiveness of exercise programs, thus further improving its compliance, adherence, and clinical outcomes [9,10]. In this context, innovative approaches, such as virtual reality (VR), have gained attention as potential interventions to enhance exercise adherence and outcomes for participants with knee osteoarthritis. VR is a digital technology that incorporates the use of interactive simulations created with computer hardware and software to present users with opportunities to engage in environments that appear and feel similar to real-world objects and events [11,12]. Within VR applications, an important distinction can be made between immersive and nonimmersive media, which differs in spatial presences [13]. With immersive technology, participants view the full panorama and are essentially inside the created environment. In a nonimmersive environment, virtual content is based on how the device (PC, smartphone, or tablet) is moved or rotated, and participants are only external observers.

An increasing number of trials have demonstrated the positive role of VR-assisted interventions in chronic pain management [14]. VR distracts users from their noxious pain perceptions by shifting their focus into video games, thus increasing their pain tolerance [15,16]. The gaming elements further enhance users' performances and motivation to exercise [17,18]. VR also facilitates skill-building for regulating painful stimuli through stimulating visual, auditory, and proprioception senses [16]. Numerous studies have supported VR-assisted physical therapy in reducing pain and improving function in low back pain and neck pain [19-21]. However, little is known about the effectiveness of VR-assisted physical therapy for knee osteoarthritis. To the best of our knowledge, only 3 trials evaluated nonimmersive VR-assisted balance and proprioception training, and findings suggested its benefits in reducing pain and improving function in knee osteoarthritis [22-24]. Since evidence has highlighted that the exposure to an immersive VR is able to elicit a better sense of presence and potentially impact the effectiveness of VR treatments [25], it is worth to explore immersive VR-assisted physical therapy in the rehabilitation of knee osteoarthritis.

This study aimed to pilot-test the feasibility, acceptability, and preliminary efficacy of an immersive VR-assisted lower limb strength training for knee osteoarthritis (VRiKnee) through a mixed methods approach. We hypothesized that VRiKnee was feasible as a home-based exercise for knee osteoarthritis. The quantitative measures would provide insights into the trend of clinical effectiveness, guiding the design of larger-scale trials, whereas the qualitative measures would uncover potential barriers, enabling the enhancement of future VR interventions.

Methods

Study Design

A convergent, parallel, mixed method study was used to gain an in-depth understanding of feasibility and acceptability of applying immersive VR on our selected population [26]. Quantitative and qualitative data were conducted concurrently in a 12-week, 2-arm, pilot randomized controlled trial design (CHiCTR2100046313). This study was conducted in Hong Kong from June 1, 2021, to March 18, 2022, in 3 separate batches. All collected quantitative and qualitative data were equally weighted, independently analyzed, and then integrated to generate results [26].

Ethical Considerations

This study complied with the Declaration of Helsinki, and ethical approval was obtained from the Joint Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee (2021.052). Written informed consent was obtained from all participants. All data were deidentified and kept confidential and were disposed on completion of this study.

Settings and Participants

Participants were recruited from an existing community project titled "CUHK-Jockey Club Pain Relief Project for Seniors," a charity program that offered nonpharmacological pain management to older adults with chronic musculoskeletal pain [27]. The eligibility criteria were screened by a trained research assistant and confirmed by a primary care physician. The inclusion criteria were a diagnosis of knee osteoarthritis based on clinical and radiographic criteria, as defined by the American Rheumatology College; moderate to severe knee pain for at least 3 months; pain intensity score ≥4 on a numeric rating scale (NRS) of 10; and use of a smartphone [28]. The exclusion criteria were decompensated organic and psychiatric disease; contraindications to VR therapy due to history of epilepsy or severe myopia (>−3.5 diopters); and comorbidities that may impede active participation in this study. Participants were recruited in 3 separate batches based on the availability of lower limb sensors.

Interventions

All participants attended a 30-minute health talk led by a registered physiotherapist to explain and demonstrate the home-based exercises. Further, 2 sets of lower limbs strengthening exercise were selected from the Ottawa panel clinical practice guideline and validated by a physiotherapist and a primary care physician [29]. This included repeated knee-extension exercise for concentric quadriceps training and squeezing a fitness ring between the thighs for vastus medialis oblique isometric training, both to be performed in a sitting position.

Participants in the VRiKnee group were instructed to perform lower limb exercises for 12 weeks using an immersive VR platform developed by our team. A smartphone app (iOS or Android) which captured outdoor garden scenes was used to create an immersive environment and participants were immersed using a head-mounted display (HMD) device (VR Shinecon 5.0). The app provided a virtual environment with ambient audio, an amateur as a coach, and interactive animations with visual feedback that encourage practice, that is, virtual flowers that would blossom with successful exercise moves and a timer for unlocking the 3 levels of difficulty. The app also recorded participants' accumulated moves, total exercise time, and game level in a virtual scoreboard (Multimedia Appendix 1).

Participants in the control group were instructed to perform the same set of exercises as guided by paper-based education pamphlets. Both groups were advised to perform the exercise 5 days per week, with an expected duration of 30 min/d for 12 weeks. An exercise diary was given to the control group to record exercise participation.

Sample Size Calculation

We applied the stepped rule of thumb in this pilot sample size calculation [30]. With a proposed future main trial design of 90% power and 2-sided 5% significance, a total of 30 participants (15 at each arm) would be able to detect an assumed effect size of 0.5.

Randomization, Allocation, Concealment, and Blinding

An off-site statistician performed 1:1 randomization using Random Allocation Software (version 1.0) to allocate and control groups. The allocation sequence was concealed from investigators and participants using sequentially numbered, opaque, and sealed envelopes. The corresponding envelopes were opened at the time of intervention assignment after all the enrolled participants had undergone all baseline assessments. It was not possible to blind participants and research assistants implementing interventions in this open-label study; however, data collectors and outcome adjudicators were blinded to the allocation status.

Data Collection

The outcome measures were collected through face-to-face interviews at baseline, 6 weeks, and 12 weeks. Demographic data such as age, gender, the number, and type of comorbid illnesses were recorded.

Quantitative Outcomes

Primary Outcomes

The primary outcome was to assess the feasibility of VRiKnee for knee osteoarthritis. Assessments included recruitment rate, dropout rate, response rate, intervention adherence rate, and adverse events. Recruitment rate was defined by the proportion of eligible participants who successfully enrolled into this study. Response rate was the percentage of usable responses obtained from our quantitative data questionnaires, calculated over the number of eligible participants. Dropout rate was the proportion of participants who dropped out of this study before its completion. Intervention adherence rate was measured by the Timeline Followback for Exercise, which is a validated retrospective self-reported calendar that documents exercise participation [31]. It was calculated by the total self-reported exercise time over the expected time spent on lower limb strengthening exercises in this study's period, that is, total exercise time (min) in wk/(30 min per d × 5 d per wk × number of wk). Adverse events or side effects were collected by participants' exercise records at each visit. Success criteria for this pilot study were set with prespecified thresholds of a 60% recruitment rate, a 70% response rate, and a <20% dropout rate.

Secondary Outcomes

Secondary outcomes evaluated the treatment effect at 12 weeks. These included pain intensity, physical function, and health-related quality of life. Pain intensity was measured using the NRS (0 - 10) [32] and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale [33,34]. Physical function was measured subjectively by the WOMAC function subscale [33,34] and objectively by the number of steps captured by ActivPAL monitor (PAL Technologies Ltd). ActivPAL is a thigh-worn accelerometer that distinguishes body posture and movement and provides objective exercise participation measurements [35,36]. Participants were assigned to wear the ActivPAL at their anterior upper thigh for 7 consecutive days at baseline and at 12 weeks, respectively. All ActivPAL data that recorded at least 4 days of activities over 20 hours were considered as valid [37,38]. The metabolic equivalents of task (MET) equation was used to measure the level of physical activities [39]. Health-related quality of life was measured by the EuroQol-5D [40,41].

Qualitative Outcomes

A concurrent approach was used to understand the acceptability toward VRiKnee. All participants were recruited to focus group interviews at 12-week using convenience sampling and qualitative data was collected until theoretical saturation [42]. A licensed counseling assistant (HHKL) with experience in qualitative interviews and a registered nurse (HHML) held the semistructured interviews with open-ended questions (Multimedia Appendix 2) to assess user experience and technology acceptance [43,44]. To engage participants in a dialogue, follow-up questions were asked according to individual responses. Since this study was conducted during the COVID-19 pandemic, both face-to-face and Zoom (Zoom Video Communications, Qumu Corporation) videoconferencing were used for focus group interviews [45].

Statistical Methods and Analysis

Quantitative data were statistically analyzed with IBM SPSS Statistics software (version 28.0.1.1, IBM Corp). Between-group differences were compared with the Mann-Whitney test for continuous variables and the Fisher exact test for categorical variables [46]. Prior to the intention-to-treat (ITT) analysis, missing values were imputed using a Markov chain Monte Carlo model, in which 10 completed data sets were imputed under the assumption that data were missing at random [47-49]. Imputed dependent variables included the NRS, WOMAC, EQ-5D, and MET; imputed independent variables or covariates included gender, number of chronic diseases, and baseline MET. The parameters were combined according to the Rubin rule, and a linear mixed model was used to analyze results. Analysis of covariance was used to assess the intervention effects on secondary outcomes at week 6 and 12 following the ITT analysis, adjusting baseline measurements at randomization [50].

Qualitative data were collected and analyzed with our quantitative data. Each semistructured interview was transcribed verbatim and analyzed using thematic analysis [51]. Further, 2 researchers (HHML and HHKL) read the transcripts line by line to familiarize with the data and initial codes were then formulated. The coding structure was then validated by 2 authors (MN and RWSS) independently. Codes sharing similar meanings were consolidated and organized into potential themes. Thematic maps were generated for review of themes in relation to the coded extract [52]. Similar codes were grouped into categories and themes and were discussed within the research team until consensus was met. We initiated coding and analysis of data early after batch one to identify possible patterns and themes, and qualitative data was collected until theoretical saturation was reached [53]. The researcher reflexivity of this study was enhanced with several strategies [54]. First, the interviews were led by HHKL, a counseling assistant with experience in qualitative interviews; HHML is a registered nurse with experience in older adult care, who participated in observation note taking during the interviews. Second, a reflexive journal was kept throughout the research process to increase transparency. To minimize individual biases, reflexive team discussions were carried out among the research team throughout the interpretation phase. The work was supported by RWSS, a family medicine physician with clinical and research experience, and MN who is an expert on mixed method studies.

Finally, data-mixing occurred when both quantitative and qualitative data were analyzed using side-by-side comparisons and joint displays [55].

Results

Overview

In total, 30 participants with a median age of 63.5 (IQR 61.8 - 66.3) years, 76% (n=23) of whom were female, were recruited. No statistical differences were found between the groups in terms of baseline measurements of variables (all *P* values were >.05). Participants' baseline characteristics are summarized in Table 1.

Table . Participants' baseline characteristics.

Demographic data	Total (N=30)	VR[a] group (n=15)	Control group (n=15)	*P* value[b]
Age (years), median (IQR)	63.5 (61.8 - 66.3)	63 (60 - 67)	64 (62 - 65)	.97
Gender (female), n (%)	23 (76.7)	10 (66.7)	13 (86.7)	.39
Retired, n (%)	23 (76.7)	12 (80)	11 (70.3)	>.99
Number of comorbidities[c], median (IQR)	2 (0 - 2.3)	2 (1.5 - 2.5)	1 (0 - 2)	.22
BMI (kg/m^2), median (IQR)	22.8 (21 - 25.7)	23.8 (21.3 - 26.6)	21.6 (20.9 - 24.4)	.17
NRS[d] (score: range 0 - 10), median (IQR)	5.5 (4-7)	6 (5 - 7)	5 (3 - 6)	.11
WOMAC[e] (score: range 0 - 2400), median (IQR)	703 (424.3 - 1231.8)	889 (464 - 1150)	464.5 (284 - 1252)	.33
WOMAC pain (score: range 0 - 500), median (IQR)	142.5 (95.8 - 250)	150 (112 - 240)	120 (67 - 271)	.35
WOMAC stiffness (score: range 0 - 200), median (IQR)	71 (35.3 - 117.8)	99 (54 - 113)	47 (17 - 132)	.12
WOMAC function (score: range 0 - 1700), median (IQR)	478 (273.5 - 829.5)	527 (320 - 828)	377 (170 - 834)	.33
EQ-VAS[f], median (IQR)	70 (57.5 - 76.3)	65 (50 - 80)	70 (60 - 70)	.57
Metabolic equivalents of task[g], median (IQR)	34.7 (33.5 - 35.4)	34.1 (33.2 - 35.2)	34.9 (33.9 - 35.6)	.20

[a]VR: virtual reality.

[b]Fisher exact test was used for categorical variables . Mann-Whitney test was used for nonnormal continuous variables.

[c]Participants reported comorbidities including hypertension, diabetes mellitus, dyslipidemia, cerebrovascular disease, indigestion, cataract, hyperthyroidism or hypothyroidism, tinnitus, fatty liver, gall stones, indigestion, constipation, gastritis, cataract, cerebrovascular disease, and cancer.

[d]NRS: numeric rating scale.

[e]WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

[f]EQ-VAS: Euro-Qol-visual analogue scale.

[g]One invalid ActivPAL record due to device failure from control group.

Primary Outcomes

Recruitment was conducted between May 27, 2022, to November 3, 2022, in 3 separated batches; the total recruitment time was 7.57 weeks. A total of 100 registrations were received; 29 did not have knee osteoarthritis, 30 did not have smartphones that operate on Android or iOS, and 11 had comorbidities that precluded their participation. The remaining 30 eligible participants were all enrolled to this study and successfully randomized, with a recruitment rate of 100%. The response rates were 80% and 93% for the VR group and control group, respectively. The median exercise adherence for VR and control groups was 78.89% (IQR 47.75%-142%) and 68.75% (IQR 38.57%-188%) during the initial 6 weeks, respectively; both groups had reduced exercise adherence, with rates decreasing to 77.22% (IQR 36.78%-104%) for the VR group and 62.08% (IQR 40.43%-166%) for the control group over the 12 weeks. The adherence rates in terms of bout time are shown in Table 2.

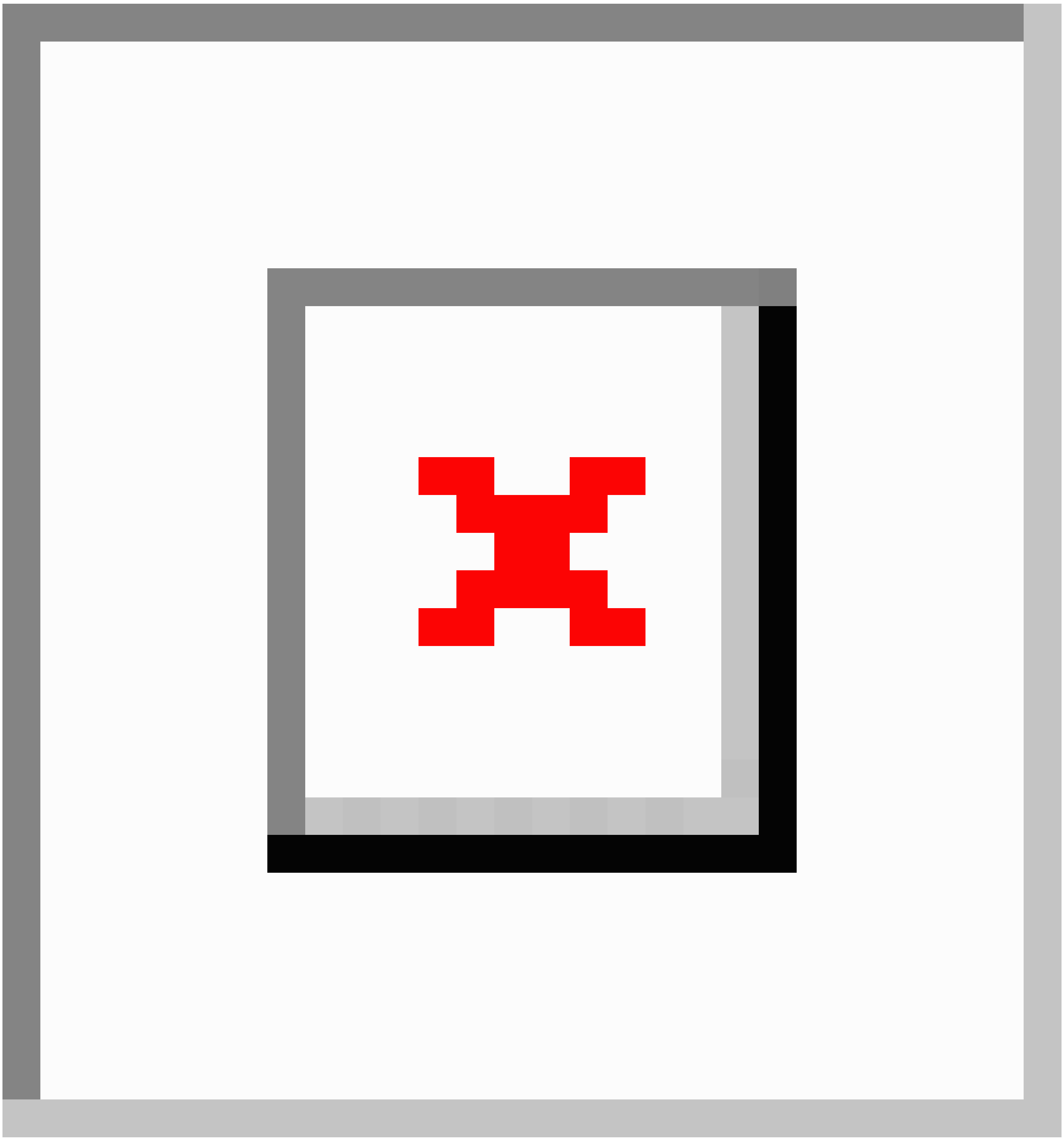
In total, 2 participants dropped out in the VRiKnee group and 1 in the control group, with a dropout rate of 13.3% and 6.7%, respectively. Reasons for dropout included hearing impairment (n=1) and worries for personal health (n=1) in the VR group and emotional distress (n=1) in the control group. Adverse events included 5 (33.3%) participants reporting cybersickness in the VR group. This study workflow is summarized in Figure 1.

Table . Primary outcomes between the intervention and control groups at week 12.

Primary outcomes	VRiKnee[a]	Control	*P* value[b]
Recruitment rate (%)	100	100	N/A[c]
Dropouts, n (%)	2 (13.3)	1 (6.67)	>.99
Response rate (%)	80	93	N/A
Intervention adherence in terms of bout time measured by the TLFB-E [d], median (IQR)			
Weeks 1 - 6	710 (429.75 - 1275.75)	618.75 (347.13 - 1692.75)	.78
Weeks 1 - 12	1390 (662 - 1877.8)	1117.5 (727.8 - 2988)	.82
Cybersickness, n (%)	5 (33.3)	N/A	N/A

[a]VRiKnee: Immersive virtual reality–assisted lower limb strength training for knee osteoarthritis.
[b]Mann-Whitney test and intention-to-treat analyses were used for nonnormal continuous variables. Fisher exact was used for categorical variables .
[c]N/A: not applicable.
[d]TLFB-E: Timeline Followback for Exercise.

Figure 1. CONSORT diagram for reporting pilot and feasibility trials. CONSORT: Consolidated Standards of Reporting Trials; VR: virtual reality.



Secondary Outcomes

At 12 weeks, no statistically significant between-group differences were detected for all secondary outcomes in this small pilot (Table 3). Yet, we detected a trend favoring the use of VRiKnee versus control in reducing the NRS (Cohen *d*=−0.084, *P*=.64) and WOMAC pain subscale (Cohen *d*=−0.089, *P*=.62) and improving the WOMAC stiffness subscale (Cohen *d*=−0.190, *P*=.29).

Table . Between-group differences of secondary outcomes at baseline, week 6, and week 12.

Variables	VRiKnee[a] (n=15), mean (SD)	Control (n=15), mean (SD)	Mean difference between groups (95% CI, VRiKnee – control)[b]	*P* value	Cohen *d*
NRS [c]					
Week 0	5.93 (1.86)	4.80 (1.49)	N/A[d]	N/A	N/A
Week 6	5.48 (1.74)	5.12 (1.22)	−0.04 (−1.31 to 1.22)	.95	−0.012
Week 12	4.78 (1.77)	4.56 (2.39)	−0.42 (−2.20 to 1.36)	.64	−0.084
Total WOMAC [e]					
Week 0	902.93 (454.75)	751.97 (549.29)	N/A	N/A	N/A
Week 6	926.73 (544.51)	782.95 (444.68)	82.60 (−281.71 to 446.91)	.66	0.080
Week 12	791.30 (425.06)	742.36 (510.53)	−61.31 (−358.88 to 236.26)	.69	−0.073
WOMAC pain					
Week 0	189.53 (96.45)	163.87 (115.27)	N/A	N/A	N/A
Week 6	184.60 (117.51)	187.51 (102.71)	−14.20 (−99.41 to 71.02)	.74	−0.059
Week 12	160.16 (82.37)	160.86 (112.29)	−17.77 (−87.78 to 52.24)	.62	−0.089
WOMAC stiffness					
Week 0	91.27 (46.39)	65.60 (54.48)	N/A	N/A	N/A
Week 6	87.79 (51.21)	64.06 (40.28)	10.53 (−25.05 to 46.10)	.56	0.104
Week 12	68.79 (42.30)	69.04 (47.40)	−17.31 (−49.52 to 14.91)	.29	−0.190
WOMAC function					
Week 0	622.13 (339.36)	522.50 (393.99)	N/A	N/A	N/A
Week 6	653.83 (384.54)	530.50 (326.02)	89.66 (−165.68 to 345.00)	.49	0.124
Week 12	560.63 (312.72)	513.69 (364.69)	−22.55 (−232.23 to 187.14)	.83	−0.038
EQ-VAS [f]					
Week 0	62.00 (19.62)	67.33 (17.20)	N/A	N/A	N/A
Week 6	67.58 (18.62)	70.83 (11.98)	0.45 (−10.35 to 11.25)	.94	0.014
Week 12	71.56 (12.79)	67.69 (17.42)	5.15 (−4.75 to 15.05)	.31	0.184
ActivPAL – MET[g] per day					
Week 0	34.18 (1.24)	34.98 (1.35)	N/A	N/A	N/A
Week 12	34.35 (1.26)	34.91 (1.96)	−0.11 (−1.50 to 1.30)	.88	−0.146

[a]VRiKnee: immersive virtual reality–assisted lower limb strength training for knee osteoarthritis.

[b]Adjusting for baseline scores: analysis of covariance and intention-to-treat analyses were used. Covariates included intervention group, gender, age, comorbidities, metabolic equivalents of task, and variable at baseline.

[c]NRS: numeric rating scale.

[d]N/A: not applicable.

[e]WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

[f]EQ-VAS: Euro-Qol-visual analogue scale.

[g]MET: Metabolic equivalents of task.

Qualitative Outcomes

Overview

In total, 10 out of the 15 VR participants joined the qualitative interviews. Their background characteristics are summarized in Table 4. The duration of interviews lasted from 41 to 67 minutes. Further, 4 themes, 11 subthemes, and 16 quotes were generated upon analysis. The four themes included (1) embracing the use of VRiKnee, (2) facilitators of VRiKnee implementation, (3) barriers to VRiKnee adherence, and (4) suggestions for VR intervention development. The subthemes within each theme are presented in Tables 5-8.

Table . VRiKnee[a] focus group participants characteristics at baseline, week 6, and week 12.

Participant	Age (years)	Gender	Chronic diseases (n)	MET[b]	NRS[c] (score: range 0 - 10)		WOMAC-pain[d] (score: range 0 - 500)		
					Baseline	Net change	Baseline	Week 6	Week 12
1	60	Female	5	32.69	8	−5	223	115	159
2	60	Male	2	33.42	3	+4	116	168	241
3	60	Female	2	35.38	7	−4	353	296	119
4	68	Male	2	33.86	5	−3	150	66	76
5	62	Male	3	33.24	5	−1	240	118	88
6	63	Female	0	32.20	8	−1	232	275	324
7	61	Female	0	34.66	6	−3	112	32	43
8	62	Female	2	34.15	6	−1	247	198	214
9	67	Female	1	36.49	6	−1	96	125.5	158.5
10	66	Male	2	34.04	5	0	97	68	N/A[e]

[a]VRiKnee: immersive virtual reality–assisted lower limb strength training for knee osteoarthritis.

[b]MET: metabolic equivalents of task.

[c]NRS: numeric rating scale.

[d]WOMAC-pain: Western Ontario and McMaster Universities Osteoarthritis Index pain subscale.

[e]N/A: not applicable (denotes missing data).

Table . Comparison between the quantitative and qualitative findings for embracing the use of VRiKnee[a].

Quantitative findings	Qualitative findings (subthemes)	Conclusion
Our pilot study had a 100% recruitment rate. The recruitment periods of the 3 separate batches were completed in 2 days, 6 weeks, and 1.3 weeks, respectively.	*Craving technology despite challenges*: Participants expressed their interest in VR[b] and all interviewees were positive to future use of advanced technology as a treatment modality.	Findings were complementary; VRiKnee was well-received by participants, with better exercise adherence and positive user experiences.
The participant retention rate was 87% and 93% for VRiKnee and control, respectively. The median exercise adherence for VRiKnee (77%, IQR 37%-104%) was higher than that for the control group (62%, IQR 40%-166%).	*Overcoming exercise inertia*: VRiKnee promoted a sense of achievement and satisfaction in users with exercise inertia by using gaming elements.	These findings support the above conclusions.

[a]VRiKnee: virtual reality–assisted lower limb strength training for knee osteoarthritis.

[b]VR: virtual reality.

Table . Comparison between the quantitative and qualitative findings for facilitators of VRiKnee[a] implementation.

Quantitative findings	Qualitative findings (subthemes)	Conclusion
Although this small pilot RCT[b] did not yield statistically significant results for both primary and secondary outcomes, we did observe a preliminary trend on the primary outcome of pain reduction on the NRS[c] (*P*=.64; Cohen *d*=−0.084) and WOMAC-pain[d] (*P*=.62, Cohen *d*=−0.089) at week 12.	*Improved knee osteoarthritis symptoms*: Participants who adhered to VRiKnee reported improvements in knee pain and function.	Findings were supplementary. Qualitative findings suggest that VRiKnee may offer potential benefits in managing knee osteoarthritis symptoms and enhancing exercise adherence, which might explain the observed positive, though statistically nonsignificant trends in some outcomes.
The median exercise adherence for VRiKnee (77%) was higher than control group (62%).	*Digital records enhancing exercise adherence*: Participants reported improved exercise compliance when using VRiKnee, with 5 participants noting that its record-keeping feature served as a reminder for them to exercise. Furthermore, 2 participants expressed their enhanced adherence, attributing to the competitive element of VRiKnee's virtual scoreboard.	These findings support the above conclusions.

[a]VRiKnee: immersive virtual reality–assisted lower limb strength training for knee osteoarthritis.

[b]RCT: randomized controlled trial.

[c]NRS: numeric rating scale.

[d]WOMAC-pain: Western Ontario and McMaster Universities Osteoarthritis Index pain subscale.

Table . Comparison between the quantitative and qualitative findings for potential barriers to VRiKnee[a] adherence.

Quantitative findings	Qualitative findings (subthemes)	Conclusion
Both groups had reduced exercise adherence over the 12 week period: VRiKnee group reduced from 78% to 56%, while control reduced from 69% to 48%.	*Boredom from repetition*: Our participants conveyed a sense of boredom resulting from stagnant audiovisual elements and repetitive exercise moves, which were perceived as too easy.	Findings were complementary. Quantitative results showed reduced exercise adherence with time, whereas the qualitative findings shed light on several challenges faced by participants during the VR[b] intervention. Addressing these issues is essential to enhance the overall user experience and promote better adherence to VR-based interventions in the future.
N/A[c]	*Technological challenges*: Participants suggested that the use of electronic appliances required additional time and effort. They also expressed a growing sense of frustration during prolonged use, especially when persistent technological errors were encountered.	These findings support the above conclusions.
No relevant quantitative data were available.	*Inconvenience of HMD[d] during exercise*: Participants reported that the use of HMD during exercise was inconvenient, attributing to its bulkiness when their mobile phones were installed and the sensation of feeling confined while wearing it over their eyes.	These findings support the above conclusions.
2 (13%) VR group participants reported cybersickness throughout the intervention.	*HMD-induced cybersickness*: 2 additional participants reported experiencing dizziness during the use of HMD in the focus group interviews; 1 participant reported experiencing blurred vision after prolonged use of HMD.	These findings support the above conclusions.

[a]VRiKnee: immersive virtual reality–assisted lower limb strength training for knee osteoarthritis.

[b]VR: virtual reality.

[c]N/A: not available.

[d]HMD: head-mounted display.

Table . Comparison between the quantitative and qualitative findings for suggestions for VR[a] intervention development.

Quantitative findings	Qualitative findings (subthemes)	Conclusion
No relevant quantitative data were available.	*Enriching VR app's audiovisual features*: Our participants suggested to increase the variety of scenes, interactive animations, and background music in different levels, which may contribute to a more engaging and enjoyable VR experience for users.	N/A[b]
No relevant quantitative data were available.	*Enhancing VR haptic interfaces*: Our participants suggested to improve the interfaces between the VR app and the motion sensor to ensure a more accurate digital record of the exercise moves.	N/A
No relevant quantitative data were available.	*Delivering VR experiences on 2D large screens*: Half of the participants expressed a preference in performing VR-assisted exercise using large "television screens" for an enhanced VR experience.	N/A

[a]VR: virtual reality.

[b]N/A: not available.

Theme 1: Embracing the Use of VRiKnee

Overview

Participants expressed a positive reception toward the implementation of VRiKnee in their treatment, acknowledging its potential as a valuable intervention.

Craving Technology Despite Challenges

Participants expressed a strong desire to incorporate advanced technology in managing their health conditions, despite facing certain challenges with the technology. The enthusiasm toward using VR for their health needs was unanimous among all participants.

> *I really like it, actually, I am very fond of new technologies.* [P8]

Overcoming Exercise Inertia

Participants highlighted the hedonic aspects of VR, such as visual stimulation and sound effects, had played a crucial role in overcoming exercise inertia. Even novice users, who initially hesitated to engage in physical activity, reported a newfound enjoyment in the VRiKnee experience, leading to a sense of accomplishment.

> *It feels like a competition when there are consecutive exercise records in the app, having it done yesterday and today. It gives me a sense of fulfillment and a little desire to challenge myself.* [P1]

Theme 2: Facilitators of VRiKnee Implementation

Overview

Physiological improvements and enhanced exercise compliance were potential facilitators for the use of VRiKnee.

Improved Knee Osteoarthritis Symptoms

Participants reported experiencing tangible improvements in their knee osteoarthritis symptoms. The perceived benefits fostered a sense of hope and optimism, and encouraged engagement.

> *After using it (VRiKnee), I'm not sure if it's my perception, but I felt I had more strength in my legs.* [P4]

Digital Records Enhancing Exercise Adherence

Participants highlighted the role of digital exercise records provided by VRiKnee in boosting their exercise adherence. The availability of these records served as helpful reminders, encouraging them to maintain regular exercise habits and thereby improving their compliance.

> *I believe that using VRiKnee allows me to keep an exercise record. This helps me to keep a regular schedule and a more disciplined exercise habit.* [P5]

Theme 3: Potential Barriers to VRiKnee Adherence

Overview

Some barriers to VRiKnee adherence were also recognized.

Boredom From Repetition

Participants experienced declining interest and excitement in VRiKnee due to the repetitive nature of visual images and music.

> *At first, the activities seemed interesting because they provided visual motivation. After a few months, it became a bit monotonous looking at the same screen.* [P1]

> *If you do it more frequently, the first level is no longer challenging.… If I could choose, I would start from level two.* [P4]

> *VR was exciting and fresh for the first and second day, but it becomes boring after a while. So I stopped using it after a week.* [P10]

Technological Challenges

Participants faced challenges in navigating VRiKnee and operating the VR system, leading to potential frustration and reduced engagement.

> *During these past few months, most of the time was spent to deal with technology, whether the sensors were working properly or not.* [P1]

> *I wasted a lot of time in dealing with this (VRiKnee). I realized that the battery power was always not enough; I also need to deal with the machine (HMD). Sometimes my phone screen froze, and my records were all gone, so I have to start all over again from level one and two.* [P6]

> *The screen turned upside down suddenly after you succeeded a few moves. Not only that, the images also kept spinning around! Why does it have to be like this? Can't I just lift my legs on my own! I'm not interested in doing (VRiKnee) anymore.* [P7]

Inconvenience of HMD During Exercise

Participants found the HMD to be heavy and uncomfortable to wear, and eventually abandoned the use of the HMD.

> *Well, I felt that wearing the HMD on my head creates a lot of obstacles for exercising. The longer you use it, the more uncomfortable you feel.* [P10]

HMD-Induced Cybersickness

Participants generally avoided using the HMD due to experiencing cybersickness, including dizziness and blurred vision. This discomfort negatively impacted their engagement with VRiKnee exercises, leading to reduced adherence to the intervention.

> *Looking backwards with the HMD causes dizziness. You must turn around and focus on the focal point to proceed to the next step.* [P5]

> *If I use the HMD for long time and focus too much, I can't read small fonts and even the large fonts!* [P2]

Theme 4: Suggestions for VR Intervention Development

Overview

Participants offered valuable insights and suggestions for the development of VR interventions.

Enriching VR App's Audiovisual Features

Most participants expressed the need to increase the variation of scenes, interactive animations, and background music across different levels of exercise difficulty. They believe these would make VRiKnee more appealing and engaging, thereby motivating them to continue its use.

> *It would be more appealing if the images changed after each level instead of being the same all the time. Additionally, the music variety could be improved, as it was monotonous! As we progress to higher levels, incorporating more lively and upbeat music would make it more engaging.* [P10]

Enhancing VR Haptic Interfaces

Enhancing haptic interfaces would offer a more immersive and engaging experience for users, potentially increasing their motivation and adherence to the VRiKnee

> *The interface (between the app and lower limb sensor) can be less complicated! And the sensor should be able to respond accurately!* [P7]

Delivering VR Experiences on 2D Large Screens

Participants preferred to use 2D large screens which would provide a more comfortable and less intrusive experience, allowing them to engage with VRiKnee without the discomfort of wearing a HMD.

> *For us, it's hard to see with a small screen, so it's better to use a larger screen.* [P1]

Integration of Quantitative and Qualitative Findings

Quantitative and qualitative findings are integrated and displayed in Tables 5-8.

Discussion

Principal Findings

In summary, VRiKnee was well-received by participants and provided positive user experience. The qualitative analysis revealed that VRiKnee showed promising potential for managing knee osteoarthritis symptoms and improving exercise adherence, which could explain the positive trends observed in some of the quantitative outcomes. However, the quantitative results indicated a decline in exercise adherence over time. The qualitative findings highlighted various challenges faced by the participants during the VR intervention, emphasizing the need to address these issues in order to enhance the overall user experience and foster better adherence to VR-based interventions in future applications.

Comparison With Previous Studies and Implications for Research and Practice

Recently, a study conducted by Özlü et al [24] investigated the impact of disease-specific gamification through HMD on pain, physical function, and balance in participants with knee osteoarthritis. While this study population consisted of younger participants, with a mean age of 53 (SD 10.19) years, the researchers documented similar adverse effects such as cybersickness, nausea, and headache. Consequently, if immersive VR technology is to be employed in the management of knee osteoarthritis or other chronic diseases, the level of immersion must be carefully considered [56,57]. Interestingly, some participants suggested that using 2D screens may offer a more comfortable alternative, thereby prompting the exploration of the cave automatic virtual environment. This technology entails a cube-shaped room where screens project computer-generated images onto the walls, floor, and ceiling, allowing users to interact with the virtual environment via handheld controllers or trackers. Previous application of the cave automatic virtual environment system in training patients with Parkinson disease has yielded promising preliminary results [58].

Furthermore, Özlü et al [24] also observed that the beneficial effects of VR intervention diminished over time. Therefore, to sustain the positive effects of VR interventions, ensuring compliance is of utmost importance. Overcoming several obstacles during the design phase can facilitate this process. For instance, regular content updates can maintain player engagement and interest by introducing new levels, characters, game modes, or features that expand the gameplay experience. Additionally, community engagement is essential in fostering a strong player community. This can be achieved through encouraging player interaction, sharing of experiences, and providing feedback via online platforms such as forums and social media, as well as through in-game features such as leaderboards and multiplayer modes. Finally, continuous improvements based on player feedback will play a critical role in enhancing the intervention.

Strengths and Limitations

This study's strengths lie in its use of a mixed methods design, which contributes to a comprehensive understanding of VRiKnee, enhances validity through triangulation, provides valuable contextual insights, and generates practical implications on the design of future VR interventions. Further, 1 limitation of this study is the small sample size, which may restrict the trial's ability to detect meaningful effect sizes.

Conclusion

The feasibility and acceptability of VRiKnee in managing knee osteoarthritis have been demonstrated, suggesting its potential clinical efficacy. However, further confirmation through larger scale trials is necessary. VRiKnee has shown promise in enhancing exercise adherence, although a decline was observed over time. Participants faced challenges during the VR intervention, underscoring the need to address these barriers to improve adherence in future implementations.

Acknowledgments

This study was funded by the Hong Kong Jockey Club Charities Trust, under the project "CUHK-HKJC Pain Relief Projects for Seniors."

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

RWSS, the corresponding author, led the conceptualization, methodology, and writing of the original paper with HHML and MN, who served as the first and second authors, respectively. HHML, HHKL, and PYHF managed the project administration, while HHKL conducted the research investigation. BW, the statistician, curated the data set and performed formal analysis. Validation of results was carried out by HHML and HHKL. RWSS was responsible for the VRiKnee software development and funding acquisition or resources. The paper submission was supervised by RWSS and SYSW. The paper was then reviewed by HHML.

Conflicts of Interest

None declared.

Multimedia Appendix 1
This file displays the illustrations for immersive virtual reality–assisted lower limb strength training for knee osteoarthritis (VRiKnee).
[DOCX File, 1066 KB - games_v12i1e52563_app1.docx]

Multimedia Appendix 2
Semistructured interview guide
[DOCX File, 21 KB - games_v12i1e52563_app2.docx]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 8135 KB - games_v12i1e52563_app3.pdf]

References

1. Cui A, Li H, Wang D, Zhong J, Chen Y, Lu H. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine 2020 Dec;29-30:100587. [doi: 10.1016/j.eclinm.2020.100587] [Medline: 34505846]
2. Ackerman IN, Bucknill A, Page RS, et al. The substantial personal burden experienced by younger people with hip or knee osteoarthritis. Osteoarthr Cartil 2015 Aug;23(8):1276-1284. [doi: 10.1016/j.joca.2015.04.008]

3. Zambon S, Siviero P, Denkinger M, et al. Role of osteoarthritis, comorbidity, and pain in determining functional limitations in older populations: European Project on Osteoarthritis. Arthritis Care Research 2016 Jun;68(6):801-810. [doi: 10.1002/acr.22755]
4. Safiri S, Kolahi AA, Smith E, et al. Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. Ann Rheum Dis 2020 Jun;79(6):819-828. [doi: 10.1136/annrheumdis-2019-216515]
5. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. Br J Sports Med 2015 Dec;49(24):1554-1557. [doi: 10.1136/bjsports-2015-095424] [Medline: 26405113]
6. Li Y, Su Y, Chen S, et al. The effects of resistance exercise in patients with knee osteoarthritis: a systematic review and meta-analysis. Clin Rehabil 2016 Oct;30(10):947-959. [doi: 10.1177/0269215515610039] [Medline: 26471972]
7. Bannuru RR, Osani MC, Vaysbrot EE, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthr Cartil 2019 Nov;27(11):1578-1589. [doi: 10.1016/j.joca.2019.06.011]
8. Nicolson PJA, Hinman RS, Kasza J, Bennell KL. Trajectories of adherence to home-based exercise programs among people with knee osteoarthritis. Osteoarthr Cartil 2018 Apr;26(4):513-521. [doi: 10.1016/j.joca.2018.01.009]
9. Palazzo C, Klinger E, Dorner V, et al. Barriers to home-based exercise program adherence with chronic low back pain: patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 Apr;59(2):107-113. [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]
10. Rousseaux F, Bicego A, Ledoux D, et al. Hypnosis associated with 3D immersive virtual reality technology in the management of pain: a review of the literature. J Pain Res 2020;13:1129-1138. [doi: 10.2147/JPR.S231737] [Medline: 32547176]
11. Weiss PL, Kizony R, Feintuch U, Katz N. Virtual reality in neurorehabilitation. In: Micheal ES, editor. Textbook of Neural Repair and Rehabilitation: Cambridge University Press; 2006:182-197. [doi: 10.1017/CBO9780511545078]
12. Goudman L, Jansen J, Billot M, et al. Virtual reality applications in chronic pain management: systematic review and meta-analysis. JMIR Serious Games 2022 May 10;10(2):e34402. [doi: 10.2196/34402] [Medline: 35536641]
13. Ventura S, Brivio E, Riva G, Baños RM. Immersive versus non-immersive experience: exploring the feasibility of memory assessment through 360° technology. Front Psychol 2019;10:2509. [doi: 10.3389/fpsyg.2019.02509]
14. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. J Pain Res 2019;12:2053-2085. [doi: 10.2147/JPR.S200498] [Medline: 31308733]
15. Gupta A, Scott K, Dukewich M. Innovative technology using virtual reality in the treatment of pain: does it reduce pain via distraction, or is there more to it? Pain Med 2018 Jan 1;19(1):151-159. [doi: 10.1093/pm/pnx109] [Medline: 29025113]
16. Ahmadpour N, Randall H, Choksi H, Gao A, Vaughan C, Poronnik P. Virtual reality interventions for acute and chronic pain management. Int J Biochem Cell Biol 2019 Sep;114:105568. [doi: 10.1016/j.biocel.2019.105568]
17. Piech J, Czernicki K. Virtual reality rehabilitation and exergames—physical and psychological impact on fall prevention among the elderly—a literature review. Appl Sci 2021;11(9):4098. [doi: 10.3390/app11094098]
18. Colloca L, Raghuraman N, Wang Y, et al. Virtual reality: physiological and behavioral mechanisms to increase individual pain tolerance limits. Pain 2020 Sep 1;161(9):2010-2021. [doi: 10.1097/j.pain.0000000000001900] [Medline: 32345915]
19. Bordeleau M, Stamenkovic A, Tardif PA, Thomas J. The use of virtual reality in back pain rehabilitation: a systematic review and meta-analysis. J Pain 2022 Feb;23(2):175-195. [doi: 10.1016/j.jpain.2021.08.001] [Medline: 34425250]
20. Ahern MM, Dean LV, Stoddard CC, et al. The effectiveness of virtual reality in patients with spinal pain: a systematic review and meta-analysis. Pain Pract 2020 Jul;20(6):656-675. [doi: 10.1111/papr.12885] [Medline: 32196892]
21. Brea-Gómez B, Torres-Sánchez I, Ortiz-Rubio A, et al. Virtual reality in the treatment of adults with chronic low back pain: a systematic review and meta-analysis of randomized clinical trials. Int J Environ Res Public Health 2021 Nov 11;18(22):11806. [doi: 10.3390/ijerph182211806] [Medline: 34831562]
22. Elshazly FAA, Nambi SG, Elnegamy TE. Comparative study on virtual reality training (VRT) over sensory motor training (SMT) in unilateral chronic osteoarthritis—a randomized control trial. Int J Med Res Health Sci 2016;5(8):7-16 [FREE Full text]
23. Lin YT, Lee WC, Hsieh RL. Active video games for knee osteoarthritis improve mobility but not WOMAC score: a randomized controlled trial. Ann Phys Rehabil Med 2020 Nov;63(6):458-465. [doi: 10.1016/j.rehab.2019.11.008] [Medline: 31981832]
24. Özlü A, Ünver G, Tuna H, Menekşeoğlu AK. The effect of a virtual reality-mediated gamified rehabilitation program on pain, disability, function, and balance in knee osteoarthritis: a prospective randomized controlled study. Games Health J 2023 Apr;12(2):118-124. [doi: 10.1089/g4h.2022.0130] [Medline: 36603100]
25. Parsons TD, Rizzo AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. J Behav Ther Exp Psychiatry 2008 Sep;39(3):250-261. [doi: 10.1016/j.jbtep.2007.07.007] [Medline: 17720136]
26. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research: SAGE Publications; 2007.
27. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004 Jan 10;363(9403):157-163. [doi: 10.1016/S0140-6736(03)15268-3] [Medline: 14726171]
28. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957 Dec;16(4):494-502. [doi: 10.1136/ard.16.4.494] [Medline: 13498604]

29. Brosseau L, Taki J, Desjardins B, et al. The Ottawa panel clinical practice guidelines for the management of knee osteoarthritis. part two: strengthening exercise programs. Clin Rehabil 2017 May;31(5):596-611. [doi: 10.1177/0269215517691084] [Medline: 28183213]
30. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073. [doi: 10.1177/0962280215588241] [Medline: 26092476]
31. Panza GA, Weinstock J, Ash GI, Pescatello LS. Psychometric evaluation of the Timeline Followback for Exercise among college students. Psychol Sport Exerc 2012 Nov;13(6):779-788. [doi: 10.1016/j.psychsport.2012.06.002] [Medline: 22844226]
32. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S240-S252. [doi: 10.1002/acr.20543] [Medline: 22588748]
33. Bellamy N. Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. Semin Arthritis Rheum 1989 May;18(4 Suppl 2):14-17. [doi: 10.1016/0049-0172(89)90010-3] [Medline: 2786253]
34. Xie F, Li SC, Goeree R, et al. Validation of Chinese Western Ontario and Mcmaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement. Qual Life Res 2008 May;17(4):595-601. [doi: 10.1007/s11136-008-9340-7]
35. Edwardson CL, Winkler EAH, Bodicoat DH, et al. Considerations when using the activPAL monitor in field-based research with adult populations. J Sport Health Sci 2017 Jun;6(2):162-178. [doi: 10.1016/j.jshs.2016.02.002] [Medline: 30356601]
36. Carlson JA, Tuz-Zahra F, Bellettiere J, et al. Validity of two awake wear-time classification algorithms for activPAL in youth, adults, and older adults. J Meas Phys Behav 2021;4(2):151-162. [doi: 10.1123/jmpb.2020-0045] [Medline: 34447927]
37. Ekegren CL, Edwards ER, Kimmel L, Gabbe BJ. Do levels of sedentary behaviour and physical activity differ according to weight-bearing status after lower limb fracture? A prospective cohort study. J Orthop Trauma Rehabil 2021 Jan 1;28:221049172110204. [doi: 10.1177/22104917211020436]
38. Jawad BN, Petersen J, Andersen O, Pedersen MM. Variations in physical activity and sedentary behavior during and after hospitalization in acutely admitted older medical patients: a longitudinal study. BMC Geriatr 2022;22:209. [doi: 10.1186/s12877-022-02917-8]
39. Harrington DM, Welk GJ, Donnelly AE. Validation of MET estimates and step measurement using the ActivPAL physical activity logger. J Sports Sci 2011 Mar;29(6):627-633. [doi: 10.1080/02640414.2010.549499] [Medline: 21360402]
40. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997 Nov;35(11):1095-1108. [doi: 10.1097/00005650-199711000-00002] [Medline: 9366889]
41. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology (Oxford) 1999 Sep;38(9):807-813. [doi: 10.1093/rheumatology/38.9.807] [Medline: 10515639]
42. Morse JM, Niehaus L. Mixed Method Design: Principles and Procedures: Left Coast Press; 2009.
43. Sagnier C, Loup-Escande E, Lourdeaux D, Thouvenin I, Valléry G. User acceptance of virtual reality: an extended technology acceptance model. Int J Hum Comput Interact 2020 Jul 2;36(11):993-1007. [doi: 10.1080/10447318.2019.1708612]
44. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006 Mar;25(2):91-97. [doi: 10.1080/01449290500330331]
45. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. Int J Qual Methods 2019 Jan 1;18:160940691987459. [doi: 10.1177/1609406919874596]
46. Chan YH. Biostatistics 102: quantitative data--parametric & non-parametric tests. Singapore Med J 2003 Aug;44(8):391-396. [Medline: 14700417]
47. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. Yale J Biol Med 2013 Sep;86(3):343-358. [Medline: 24058309]
48. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 Sep;8(3):206-213. [doi: 10.1007/s11121-007-0070-9] [Medline: 17549635]
49. Zhang P. Multiple imputation: theory and method. Int Statistical Rev 2003 Dec;71(3):581-592. [doi: 10.1111/j.1751-5823.2003.tb00213.x]
50. Vickers AJ. Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data. BMC Med Res Methodol 2005 Nov 3;5:35. [doi: 10.1186/1471-2288-5-35] [Medline: 16269081]
51. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
52. Braun V, Clarke V. Thematic Analysis: American Psychological Association; 2012.
53. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2021 Mar 4;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
54. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qual Psychol 2022;9(1):3-26. [doi: 10.1037/qup0000196]

55. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
56. Stanney K, Lawson BD, Rokers B, et al. Identifying causes of and solutions for cybersickness in immersive technology: reformulation of a research and development agenda. Int J Hum Comput Interact 2020 Nov 25;36(19):1783-1803. [doi: 10.1080/10447318.2020.1828535]
57. Martirosov S, Bureš M, Zítka T. Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality. Virtual Real 2022;26(1):15-32. [doi: 10.1007/s10055-021-00507-4] [Medline: 34025203]
58. Foreman KB, Wilson C, Dibble LE, Merryweather AS. Training persons with Parkinson disease using an advanced CAVE virtual reality system. FASEB J 2019 Apr;33(S1). [doi: 10.1096/fasebj.2019.33.1_supplement.335.4]

Abbreviations

HMD: head-mounted display
ITT: intention-to-treat
MET: metabolic equivalents of task
NRS: numeric rating scale
VR: virtual reality
VRiKnee: immersive virtual reality–assisted lower limb strength training for knee osteoarthritis
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

Edited by A Mavragani, T Leung; submitted 08.09.23; peer-reviewed by W Li; revised version received 23.06.24; accepted 24.06.24; published 27.09.24.

Please cite as:
Lo HHM, Ng M, Fong PYH, Lai HHK, Wang B, Wong SYS, Sit RWS
Examining the Feasibility, Acceptability, and Preliminary Efficacy of an Immersive Virtual Reality–Assisted Lower Limb Strength Training for Knee Osteoarthritis: Mixed Methods Pilot Randomized Controlled Trial
JMIR Serious Games 2024;12:e52563
URL: https://games.jmir.org/2024/1/e52563
doi:10.2196/52563

Effects of a Serious Game for Adolescent Mental Health on Cognitive Vulnerability: Pilot Usability Study

Eva De Jaegere[1], PhD; Kees van Heeringen[1], MD, PhD; Peter Emmery[2], MD; Gijs Mommerency[3], MSc; Gwendolyn Portzky[1], PhD

1

2

3

Corresponding Author:
Eva De Jaegere, PhD

Abstract

Background: Adolescent mental health is of utmost importance. E-mental health interventions, and serious games in particular, are appealing to adolescents and can have beneficial effects on their mental health. A serious game aimed at improving cognitive vulnerability (ie, beliefs or attitudes), which can predispose an individual to mental health problems, can contribute to the prevention of these problems in adolescents.

Objective: This study aimed to assess the feasibility of the prototype of a serious game called "Silver."

Methods: The prototype of the serious game was developed using a user-centered participatory design. The prototype of Silver focused on 1 aspect of a serious game for improving cognitive vulnerability in adolescents, that is, the recognition and identification of cognitive distortions. Through the game, players were required to identify and classify the character's thoughts as helpful or unhelpful. Upon successful advancement to the next level, the task becomes more challenging, as players must also identify specific types of cognitive distortions. A pre- and posttest uncontrolled design was used to evaluate the game, with a 1-week intervention phase in which participants were asked to play the game. Participants aged 12-16 years were recruited in schools. The outcomes of interest were the recognition of cognitive distortions and presence of participants' cognitive distortions. The game was also evaluated on its effects, content, and usefulness.

Results: A total of 630 adolescents played Silver and completed the assessments. Adolescents were significantly better at recognizing cognitive distortions at the pretest (mean 13.09, SD 4.08) compared to the posttest (mean 13.82, SD 5.09; t_{629}=−4.00, P<.001). Furthermore, their cognitive distortions decreased significantly at the posttest (mean 38.73, SD 12.79) compared to the pretest (mean 41.43, SD 10.90; t_{629}=7.98, P<.001). Participants also indicated that the game helped them recognize cognitive distortions. Many participants considered the game appealing (294/610, 48.2%) but boring (317/610, 52%) and preferred a more comprehensive game (299/610, 49%).

Conclusions: Findings from this study suggest that a serious game may be an effective tool for improving cognitive vulnerability in adolescents. The development of such a serious game, based on the prototype, is recommended. It may be an important and innovative tool for the universal prevention of mental health problems in adolescents. Future research on the effects of the game is warranted.

(JMIR Serious Games 2024;12:e47513) doi:10.2196/47513

KEYWORDS

e-health; cognitive psychology; cognitive distortion; cognitive vulnerability; digital health; serious games; adolescent mental health; prototype; adolescent; prevention; eHealth

Introduction

Mental illness is one of the biggest health burdens worldwide, and adolescent mental health is a particular global concern [1]. According to the United Nations' recently published report on children's mental health, approximately 1 in 7 adolescents aged between 10 and 19 years experience a diagnosed mental disorder globally [2]. Approximately 40% of these disorders are attributable to anxiety and depression [1]. The psychological distress and economic costs due to adolescent mental health problems are enormous. In 2021, the invisible economic cost due to mental health problems in adolescents was estimated at US $387 billion per year globally [2]. Moreover, these mental health problems are important risk factors for self-harm and suicide among adolescents [3,4].

Early access to treatment can improve outcomes, but limited treatment resources result in long waiting lists and undertreatment [5]. Society, and health systems in particular, need to cope with the increasingly high demands to reduce costs and prevent self-harm and suicide among adolescents. E-mental health interventions are already being used across health care and may be more accessible, engaging, and acceptable options [6-8]. These can take various forms, such as text-based programs, multimedia and interactive programs, virtual reality–based programs, and serious games [9-12]. Evidence of their beneficial effects is increasingly provided [6,13-15]. Regarding serious games, a recent review concluded that the limited evidence indicates a beneficial effect on reducing mental health problems [11,16-19]. Thus, there is a massive potential for serious games to be a new, emerging treatment modality that is more acceptable and engaging, as it uses game mechanics, especially for adolescents who are major users of computerized games in the present-day digital world. However, more research is needed [11,18,20].

This study therefore aimed at providing further evidence of the beneficial effects of serious games by piloting a prototype ("Silver") developed to counter cognitive characteristics, which are known to increase the vulnerability to common mental health problems such as depression and anxiety [21]. More specifically, the prototype targets adolescents (aged 12-16 years), aiming to gain insight into cognitive distortions (ie, negative, biased thoughts that influence people's interpretation of themselves or the world [22]) and their effect on feelings and behaviors. Cognitive distortions, including all-or-nothing thinking, overgeneralization, or mind reading, can contribute to the development of mental health issues such as depression or anxiety during adolescence [23-25]. Identifying and modifying these distortions and replacing them with more balanced, helpful thoughts—that is, cognitive restructuring—is a common technique used in cognitive behavioral therapy (CBT) [22], which has been demonstrated to prevent mental health issues during this vulnerable development phase [26,27] and improve mental well-being [28]. Given the relationship between cognitive distortions and adolescent mental health, as well as the potential of CBT-based techniques to address these distortions, this study aimed to investigate the specific impact of the prototype on adolescents' cognitive processing. Therefore, the primary hypothesis was that adolescents would improve at recognizing and categorizing helpful and unhelpful thoughts (ie, cognitive distortions) after playing Silver. Furthermore, playing the game was expected to lead to a decline in cognitive distortions. Finally, the usability of Silver was assessed, focusing on the appeal of the game.

Methods

Participants and Recruitment

Participants were aged between 12 and 16 years and had a smartphone or tablet. Participants were excluded if they were not proficient in Dutch. Recruitment took place from August 2017 to October 2017 via schools.

Ethical Considerations

School directors of 8 secondary schools with different curricula across Flanders (ie, the Dutch-speaking region in Belgium) consented to participate in the study. Parents or guardians were informed about the study and given the opportunity to decline their child's participation (opt out). Web-based assent was obtained from the adolescents before the start of the study. Participants did not receive any form of compensation for their involvement in this study. The data were deidentified prior to analysis to safeguard participants' privacy. The study was approved by the Commission for Medical Ethics of the University Hospital Ghent (Belgian registration B670201731975).

Design and Procedure

The prototype was evaluated using a pre- and posttest uncontrolled design with a 1-week intervention phase. Before participants received access to the prototype, they were asked to fill in a web-based questionnaire (pretest). Immediately after completing the questionnaire, they received access to Silver, which they were asked to play daily for 1 week. After 1 week, they were asked to fill in the second web-based questionnaire (posttest).

Intervention

The prototype of the serious game Silver aims to reduce cognitive distortions in adolescents. The prototype is based on a cognitive behavioral framework and focuses on 1 element of mental health improvement, that is, gaining insight into cognitive distortions and their effects on feelings and behaviors. The prototype was designed and developed in a cocreative manner, in which the target users themselves (ie, adolescents aged 12-16 years) were involved, as well as a clinical child psychologist, a child and adolescent psychiatrist, and professional game designers. The design, therefore, was user centered and participatory. The cocreation process was managed by the company that developed the game's prototype.

The game is set up in 3 different worlds inhabited by anthropomorphic animals. Each world has different chapters that can be played. The more a player progresses in the game, the more difficult it becomes. The game always starts with an animal that is stuck in his or her mind and therefore is in a "glitch." The incident preceding the "glitch" is explained through a flashback where the player is shown how this came about (see Figure 1). The events represent difficult situations that are very relatable for adolescents (eg, not getting likes on a social media post). Afterward, the player is shown the character's thoughts. Each time after reading a thought, the player must decide whether it is a helpful or unhelpful thought (ie, cognitive distortion). If the thoughts are correctly recognized, the unhelpful thoughts are fired upon by little robots and the helpful thoughts return to the character's head. When enough helpful thoughts have been collected, the character is released from his or her "glitch" and the chapter is completed. At higher levels, the player is also asked to indicate which type of unhelpful thought it is. At first, 2 types of cognitive distortions are introduced, that is, future thinking and all-or-nothing thinking. Afterward, 1 more type of cognitive distortion is added, that is, mind reading (see Figure 2). Thus, the further you progress through the game, the more difficult it becomes as the thoughts can be categorized into more types of cognitive distortions. In this way, the player learns to gain insight into the different types of cognitive distortions.

Figure 1. Screenshot of Silver, showing a flashback to the incident before the glitch. “Maar vanochtend stak zijn moeder nog een kakelvers woestijnvosje in zijn lunchbox.” means “But this morning, his mother stuck another brand-new desert fox in his lunch box.”

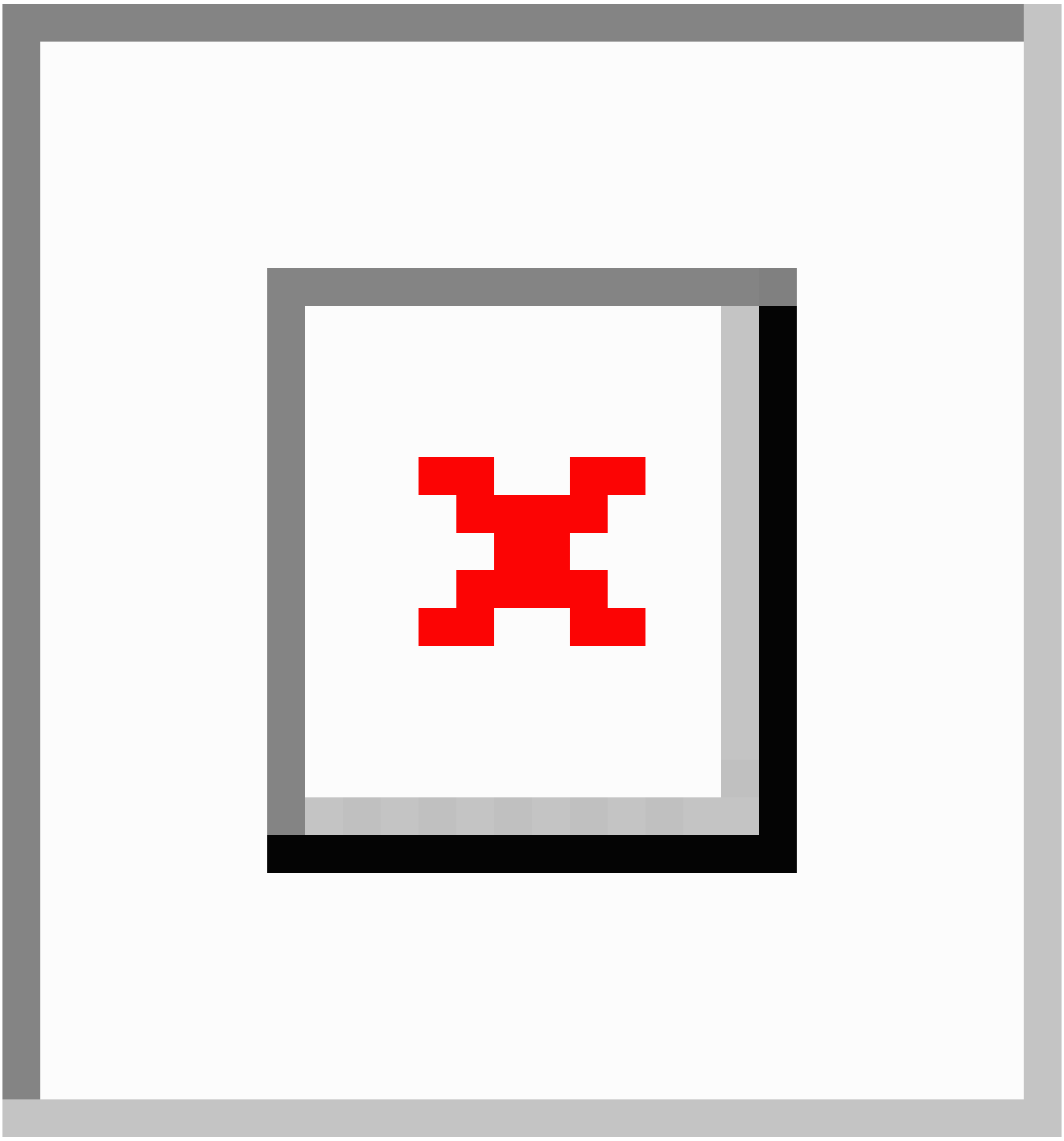
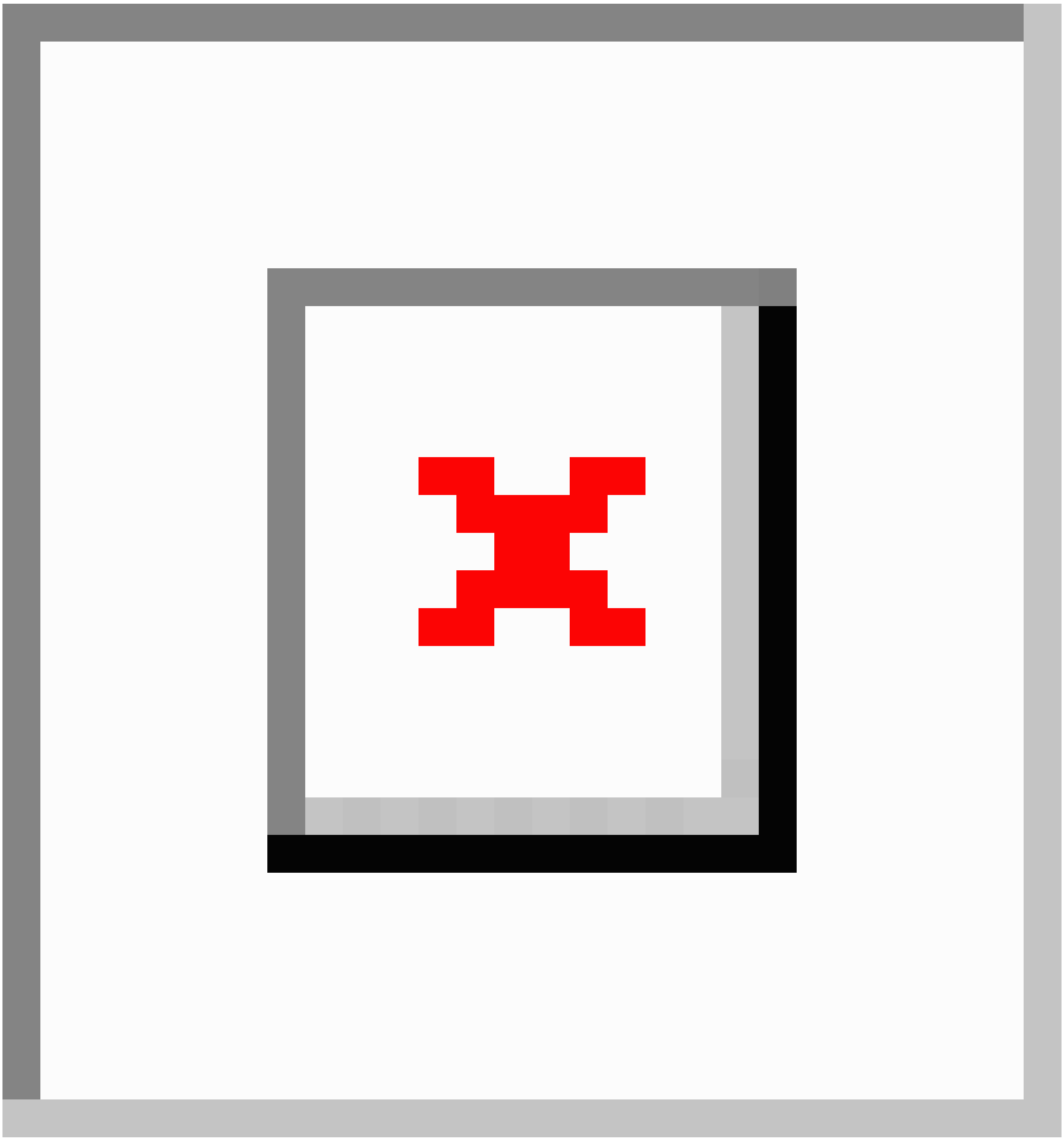


Figure 2. Screenshot of Silver, showing the process of identifying types of cognitive distortions. “Aan tafel zijn ze al zeker met mij aan het lachen.” means “At the table, they are already definitely laughing at me.“; “Helpend” means “Helpful”; “Toekomst” means ”Future”; “Gedachten lezen” means “Read minds”; and “Zwart-wit” means “Black and white.”



Measures

All outcome measures were collected via self-report questionnaires, which were administered on the web.

Demographics

Sociodemographic information (ie, gender, age, and education) was assessed at the pretest. Data on sex (male, female, or other), age (in years), and education (first or second grade and type of curriculum, ie, general secondary education, technical secondary education, secondary education in the arts, and vocational education) were obtained. In addition, participants were asked if they had ever been in therapy for psychological problems.

Media and Game Use

At baseline, data were collected regarding participants’ use of media by asking participants what type of media they used and how often they used it on a 5-point scale (1=“never”; 5=“daily”). Items regarding game use assessed whether they ever played computer or video games and whether they still played them. Furthermore, participants were asked on which devices they played the games and how often they did this on a 5-point scale (1=“never”; 5=“daily”). They also gave an estimate about their knowledge about games, ranging from 1= “no knowledge” to 4=“expert.”

Primary Outcome Measure: Recognition of Negative and Positive Automatic Thoughts

The primary outcome measure focused on recognizing helpful and unhelpful thoughts. A questionnaire was developed, in which 20 items of the Dysfunctional Attitude Scale, Form A, Dutch translation [29,30] were used. Participants were asked to classify each item as helpful, unhelpful, or "I do not know." Items were scored as true or false. Scores ranged from 0 to 20, with higher scores indicating a better identification of negative and positive automatic thoughts. The internal consistency of the scale in this study was α=.82.

Secondary Outcome Measure: Presence of Cognitive Distortions

The Children's Negative Cognitive Error Questionnaire–Revised [31] is a 16-item self-report questionnaire that assesses cognitive distortions in those aged 9-17 years. The questionnaire consists of 5 subscales that measure 5 categories of cognitive distortions: "underestimation of the ability to cope," "personalizing without mind reading," "mind reading," "selective abstraction," and "overgeneralizing." In each item, a situation is described, followed by a possible thought about the situation. Participants are asked to rate on a 5-point Likert scale how much the thought corresponds to what they would think in that situation, ranging from "almost exactly like I would think" (5 points) to "not at all like I would think" (1 point). Total scores range from 16 to 80, with a higher score reflecting more distorted cognitive errors. The total scale has a good level of internal consistency and good test-retest reliability [31,32].

Game Evaluation

At the posttest, participants were asked to rate various statements regarding the effects, content, and usefulness of the game. Items were rated on a 5-point Likert scale (1="completely disagree"; 5="completely agree"). An example of an item is "By playing the game I will be able to recognize my own cognitive distortions." Participants were also asked how they would rate the game overall on a scale from 0 to 10.

Statistical Analysis

Power and sample size could not be based on previous studies due to a lack of comparable studies. An effect size of 0.3 was assumed. To detect such an effect size with α=.05 and β=.80, a total sample of 500 participants was calculated. However, since a possible high dropout of 70% to 75% was expected [33], the total required sample size was estimated at 1753.

Differences between the participants of the study and those who dropped out during the study, as well as differences between those who played the game (ie, gamers) and those who did not play it (ie, nongamers), were examined with χ^2 tests (for categorical variables) and 2-tailed independent-sample *t* tests (for continuous variables). Mean changes between the pre- and posttests were carried out using 2-tailed paired-samples *t* tests. The corresponding effect sizes were assessed using Cohen *d*. A significance level of .05 was used for all outcome analyses. All data were analyzed using SPSS software (version 27; IBM Corp).

Results

Sociodemographic Characteristics and Baseline Outcome Measure

A total of 1654 adolescents signed up to take part in the study. Among these, 1140 took part in the pre- and posttests. Table 1 presents the differences between the participants who completed both the pre- and posttests and those who only completed the pretest.

Table . Sociodemographic characteristics and baseline outcome measures of participants who completed both the pre- and posttests and participants who only completed the pretest.

Characteristics		Pre- and posttests (n=1140)	Pretest only (n=514)	*P* value
Sex, n (%)				
	Female	700 (61.4)	303 (58.9)	.34
	Male	429 (37.6)	206 (40.1)	.34
	Other	11 (1)	5 (1)	.59
Age (y), mean (SD)		13.40 (1.32)	13.86 (1.36)	.001
Education curriculum, n (%)				
	First-grade GSE[a]	306 (26.8)	84 (16.4)	<.001
	First-grade VE[b]	55 (4.8)	79 (15.4)	<.001
	Second-grade GSE	513 (45)	158 (30.8)	<.001
	Second-grade TSE[c]	173 (15.2)	56 (10.9)	.02
	Second-grade SEA[d]	20 (1.8)	51 (9.9)	<.001
	Second-grade VE	73 (6.4)	85 (16.6)	<.001
Treatment for psychological problems, n (%)				
	Never been to therapy	984 (86.3)	419 (81.7)	.02
	More than a year ago	68 (6)	32 (6.2)	.83
	Less than a year ago	50 (4.4)	31 (6)	.15
	In therapy	38 (3.3)	31 (6)	.01
Baseline outcome measures, mean (SD)				
	Recognizing automatic thoughts	12.64 (4.37)	11.51 (4.82)	<.001
	CNCEQ-R[e]	40.81 (10.97)	40.66 (12.13)	.82

[a]GSE: general secondary education.

[b]VE: vocational education.

[c]TSE: technical secondary education.

[d]SEA: secondary education in the arts.

[e]CNCEQ-R: Children's Negative Cognitive Error Questionnaire–Revised.

Of the 1140 adolescents who completed the pre- and posttests, 510 (44.7%) reported that they did not engage with the game (hereafter referred to as nongamers). The primary reason for nonengagement was technical problems (233/510, 45.7%) such as inability to download the game. Other reasons included a lack of time or forgetfulness (198/510, 38.8%), a disinterest in the game (51/510, 10%), and other unspecified reasons (59/510, 5.5%). In contrast, 630 (55.3%) of the 1140 adolescents indicated that they played the game (hereafter referred to as gamers). There were no significant differences in all baseline sociodemographic characteristics between gamers and nongamers except for type of education curriculum (P=.02; see Table 2). Additionally, adolescents who played the game scored significantly higher on recognizing automatic thoughts (mean 13.09, SD 4.08 vs mean 12.08, SD 4.65; t_{1020}=–3.86, P<.001) and significantly higher on the Children's Negative Cognitive Error Questionnaire–Revised (mean 41.43, SD 10.90 vs mean 40.05, SD 11.01; t_{1138}=–2.10, P=.04) at baseline (see Table 2).

Table . Sociodemographic characteristics and baseline outcome measures of gamers and nongamers in schools.

Characteristics		Gamers (n=630)	Nongamers (n=510)	*P* value
Sex, n (%)				
	Female	371 (58.9)	329 (64.5)	.053
	Male	251 (39.8)	178 (34.9)	.09
	Other	8 (1.3)	3 (0.6)	.24
Age (y), mean (SD)		13.43 (1.30)	13.36 (1.33)	.40
Education curriculum, n (%)				
	First-grade GSE[a]	186 (29.5)	120 (23.5)	.02
	First-grade VE[b]	25 (4)	30 (5.9)	.13
	Second-grade GSE	264 (41.9)	249 (48.8)	.02
	Second-grade TSE[c]	99 (15.7)	74 (14.5)	.57
	Second-grade SEA[d]	14 (2.2)	6 (1.2)	.18
	Second-grade VE	42 (6.7)	31 (6.1)	.69
Treatment for psychological problems, n (%)				
	Never been to therapy	542 (86)	442 (86.7)	.76
	More than a year ago	39 (6.2)	29 (5.7)	.72
	Less than a year ago	31 (4.9)	19 (3.7)	.33
	In therapy	18 (2.9)	20 (3.9)	.32
Baseline outcome measures, mean (SD)				
	Recognizing automatic thoughts	13.09 (4.08)	12.08 (4.65)	<.001
	CNCEQ-R[e]	41.43 (10.90)	40.05 (11.01)	.04

[a]GSE: general secondary education.
[b]VE: vocational education.
[c]TSE: technical secondary education.
[d]SEA: secondary education in the arts.
[e]CNCEQ-R: Children's Negative Cognitive Error Questionnaire–Revised.

Media and Game Use

Table 3 shows media and game use of the gamers versus the nongamers. The nongamers were significantly less likely to play games currently (270/510, 52.9% vs 405/630, 64.3%; χ^2_1=15.02, *P*<.001) and were significantly more likely to have "no knowledge" about games than the gamers (69/510, 13.5% vs 54/630, 8.6%; χ^2_1=7.20, *P*=.007).

Table . Media and game use of gamers and nongamers in schools.

Media and game use		Gamers (n=630)	Nongamers (n=510)	*P* value
Media use, n (%)				
	Smartphone use	615 (97.6)	496 (97.3)	.70
	Tablet use	466 (74)	360 (70.6)	.20
	Desktop use	226 (35.9)	179 (35.1)	.79
	Laptop use	501 (79.5)	409 (80.2)	.78
	Game console use	367 (58.3)	279 (54.7)	.23
Game playing, n (%)				
	Ever	601 (95.4)	476 (93.3)	.13
	Currently	405 (64.3)	270 (52.9)	<.001
Game knowledge, n (%)				
	No knowledge	54 (8.6)	69 (13.5)	.007
	Beginner	260 (41.3)	211 (41.4)	.97
	Advanced	255 (40.5)	198 (38.8)	.57
	Expert	61 (9.7)	32 (6.3)	.04

Outcome Measures

Table 4 shows the mean changes in the outcome measures from pre- to posttest and its effect sizes for the group that played the game. The gamers significantly improved in recognizing automatic thoughts (*P*<.001) and had significantly fewer distorted cognitive errors (*P*<.001). On both measures, the difference represented a small effect size (0.16 and 0.23, respectively).

Table . Pre- and posttest scores on outcome measures and 2-tailed paired *t* test results.

Outcome measures	Pretest score, mean (SD)	Posttest score, mean (SD)	*t* test (*df*)	*d* (95% CI)	*P* value
Recognizing automatic thoughts	13.09 (4.08)	13.82 (5.09)	−4.00 (629)	0.16 (−1.09 to −0.37)	<.001
CNCEQ-R[a]	41.43 (10.90)	38.73 (12.79)	7.98 (629)	0.23 (3.36 to 7.98)	<.001

[a]CNCEQ-R: Children’s Negative Cognitive Error Questionnaire–Revised.

Game Evaluation

A total of 610 gamers gave a score on the various evaluation items (see Table 5). Besides the high number of neutral responses, they generally moderately or highly agreed with the items. The median overall satisfaction rating of 599 gamers, which was scored on a scale of 1 to 10, was 6 and the mean was 5.51 (SD 2.30).

Table . Game evaluation ratings (n=610).

Statements	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Score, mean (SD)
By playing the game I learned about different ways of thinking.	134 (22)	278 (45.6)	198 (32.5)	2.10 (0.73)
By playing the game I will recognize my unhelpful thoughts.	123 (20.2)	244 (40)	243 (39.8)	2.20 (0.75)
The game helps me to reflect more upon my thoughts.	164 (26.9)	272 (44.6)	174 (28.5)	2.02 (0.75)
I can empathize with the stories in the game.	202 (33.1)	204 (33.4)	204 (33.4)	2.00 (0.82)
I think the game is beautifully made.	96 (15.7)	220 (36.1)	294 (48.2)	2.32 (0.73)
I find the game difficult.	283 (46.4)	198 (32.5)	129 (21.1)	1.75 (0.78)
I find the game boring.	122 (20)	171 (28)	317 (52)	2.32 (0.79)
I would like it if the game was not only about thoughts but also about feelings and relaxation.	102 (16.7)	209 (34.3)	299 (49)	2.32 (0.74)

Discussion

Principal Findings

The aim of this pilot study was to test the efficacy and usability of the prototype of Silver, a serious game aimed at reducing cognitive distortions in adolescents and thus decreasing cognitive vulnerability to mental health problems. The trial supports our hypothesis that automatic thoughts are better recognized after playing Silver. This enhancement in cognitive awareness was also reflected in the evaluation of the game, wherein gamers indicate that playing the game helps in recognizing unhelpful thoughts. Furthermore, after having played the game, adolescents showed fewer cognitive distortions than before playing. These findings are in line with prior studies, underscoring the positive impact of serious games on cognitive beliefs and modification [19,34-37]. This study contributes to the growing evidence on digital interventions that incorporate core components of CBT, such as cognitive restructuring, and their beneficial effects on well-being and mental health issues [38]. Silver's emphasis on identifying and mitigating cognitive distortions aligns closely with the principles of rational emotive behavior therapy (REBT), a type of CBT [39]. This therapeutic approach focuses on the identification, challenge, and substitution of irrational beliefs with rational counterparts, alongside learning to manage emotions and behavior in a more helpful way. Literature suggests that serious games focusing on REBT techniques seem to have a strong positive effect in mitigating symptoms of depression and anxiety [19]. Considering the significant outcomes associated with the current version of Silver, which primarily targets cognitive distortions, an expanded version of the game that includes elements intended to address emotional and behavioral aspects could potentially have a greater positive impact on mental health. The inclusion of these additional components could further enhance the game's therapeutic effectiveness, aligning with REBT's approach to mental health improvement [19].

Regarding the evaluation of Silver, the majority (294/610, 48.2%) of gamers indicated that Silver is appealing. Research has shown that nonappealing interfaces may be off-putting and may cause adolescents to disengage from the game. The cocreation of Silver with the target audience probably has ensured that it has attractive aesthetics [40]. In contrast, the majority (317/610, 52%) also perceived Silver as boring. Adolescents may experience serious games in such a manner since they often have very didactic content that does not match commercial, off-the-shelf games, and as a result, the adolescents may cease playing these games [40]. This dichotomy can be attributed to the prototype's focus on a single aspect of cognitive vulnerability (ie, cognitive distortions), which, although important, may lack the variety necessary to sustain players' interest over time. Therefore, this may be a critical area for further development in diversifying the game's content and mechanics. The game could be broadened to include a range of elements. This could mitigate the issue of monotony, thereby improving overall engagement and effectiveness [19]. Furthermore, the gamers indicated that they would like a more comprehensive game that also deals with feelings and relaxation. Such a game may be more eventful, more fascinating, and less boring.

Preceding a discussion of potential implications of the study findings for the development of serious games, methodological issues need to be addressed. First, as this was a pre- and posttest study, a control group was lacking. Therefore, we were unable to compare the effects on the gamers with those in a random control group that did not play the game. A convenience sample was used, which can lead to a selection bias and consequently underrepresent or overrepresent particular groups. Efforts were made to counter this as much as possible by recruiting a large number of adolescents from 8 different schools with various curricula. However, it is also unclear why some adolescents agreed to take part in the study but others did not. As the sample was not chosen at random, the inherent bias in convenience sampling means that any generalizations of findings must be made with caution. Second, a large group did not adhere to the study protocol. The main reason for dropout were technical difficulties. These were largely due to the prototypic nature of our app; as it was not readily available on app stores, it required adolescents to undertake multiple steps before receiving access to the game. Moreover, participants' feedback showed time constraints and forgetfulness as additional factors for not engaging with the game. To mitigate these issues in future studies, it is imperative to streamline the app's accessibility, potentially by securing its availability on common digital distribution platforms. Furthermore, incorporating human support may serve to enhance participant engagement and possibly the overall effectiveness of the intervention [41]. Addressing these aspects is critical for improving study adherence and ensuring the robust evaluation of the app's therapeutic potential. Furthermore, participants who dropped out were older and had a lower education level. Additionally, they may have experienced more mental health problems, as baseline measurements showed that they were less skilled at recognizing automatic thoughts and were currently more likely to be in treatment for psychological problems. The nongamer group encompassed more adolescents with little interest in games, as they currently played no games or had less knowledge about them. In addition, they were less skilled at recognizing automatic thoughts but had fewer cognitive distortions themselves. However, high attrition rates are not uncommon when studying e-mental health interventions. A systematic review and meta-analysis of computer-based psychological treatments showed an overall dropout rate of 57%, which further increased to 74% in unsupported digital programs [6,42]. In this study, the adolescents were also not offered any support, and this may have had a major effect on adherence. Adding human support could decrease the attrition rate and may even increase the effectiveness of the serious game [43,44]. Moreover, the effectiveness of serious games is often assessed in pragmatic study trials. The real-life settings in which these studies are carried out can also have an impact on the attrition rate. Nevertheless, this type of study can improve the generalizability of the results, as the environments in which they will be implemented are similar to the ones in the trials [45]. Third, no standardized questionnaire for the recognition of automatic thoughts was used. The questionnaire was based on an existing standardized questionnaire [29,30] but was adapted for this study. Fourth, participants often responded with "neutral" in the game evaluation. Although they may have used the neutral midpoint response because they did not comprehend the items or were undecided, offering these neutral responses may decrease the quality and reliability of a questionnaire, particularly in adolescents, since they are more sensitive to pleasing by selecting a neutral answer. Future studies should consider omitting the neutral midpoint [46,47]. Fifth, the study's emphasis on quantitative measures may have introduced acquiescence bias. Adding qualitative methodologies may provide a more nuanced perspective of the participants' experiences with the prototype. Future studies should consider using a mixed methods approach to enhance understanding of the intervention's impact [48].

Lastly, the eligibility criterion requiring participants to possess a smartphone or tablet may have introduced a selection bias, potentially excluding adolescents without access to such technology or those reluctant to use it. This could inadvertently reinforce the digital divide, that is, inequalities in accessing and using information and communication technologies [49], and limit the generalizability of the findings. Future research should address this limitation by using more inclusive recruitment strategies to minimize technological barriers and ensure broader participation.

It is difficult to assess the effect of these methodological issues on the validity of the current findings, more so as the current results are difficult to compare to those from similar previous studies. The few available studies of serious games aimed at cognitive training had targeted adults or children with particular mental health problems such as anxiety, alcohol use disorder, or attention-deficit/hyperactivity disorder [11,20]. To the best of the authors' knowledge, only 2 serious games were studied regarding effects on emotional resilience or mental health promotion in a community sample of adolescents. The results of these studies are comparable to our findings [20,35,36].

Conclusions

In conclusion, and keeping the limitations mentioned above in mind, this pilot study demonstrates the promising effects of the Silver prototype. Notably, participants exhibited not only an enhanced ability to recognize cognitive distortions but also a significant decrease in their own experiences of such distortions after engaging with the prototype. This observation suggests a potential positive influence on cognitive characteristics, which are commonly associated with mental health issues. It is therefore recommended that a serious game aimed at decreasing cognitive vulnerability and therefore improving mental health in a general population of adolescents should be developed further and that its efficacy should be studied in future research. This study provides a few cues for further research. The dropout of adolescents who may have the greatest need for cognitive restructuring is a matter of concern, and reasons and remedies for this worrisome issue should be targeted in future research. Randomized controlled trials should be used to further explore the effects of serious games on adolescents in the general population, preferably using an active control group that engages with a different type of digital intervention [50]. Follow-up periods should be sufficiently long to study potential preventative effects among adolescents (Reynard et al [20]).

Acknowledgments

Funding for this study was provided by the Flemish government. We are very grateful to the participants, teachers, school directors, and game company whose time and efforts made this work possible.

Conflicts of Interest

None declared.

References

1. Kuehn BM. Lack of adolescents' mental health care is a global challenge. JAMA 2021 Nov 16;326(19):1898. [doi: 10.1001/jama.2021.20064] [Medline: 34783832]
2. United Nations Children's Fund. The state of the world's children 2021: on my mind – promoting, protecting and caring for children's mental health. : UNICEF; 2021 Oct URL: https://www.unicef.org/media/114636/file/SOWC-2021-full-report-English.pdf [accessed 2024-04-26]
3. Bridge JA, Goldstein TR, Brent DA. Adolescent suicide and suicidal behavior. J Child Psychol Psychiatry 2006;47(3-4):372-394. [doi: 10.1111/j.1469-7610.2006.01615.x] [Medline: 16492264]
4. Hawton K, Saunders KEA, O'Connor RC. Self-harm and suicide in adolescents. Lancet 2012 Jun 23;379(9834):2373-2382. [doi: 10.1016/S0140-6736(12)60322-5] [Medline: 22726518]
5. NICEimpact mental health. : National Institute for Health and Care Excellence (NICE); 2019 URL: https://www.issup.net/files/2019-05/NICEimpact-mental-health.pdf [accessed 2024-04-26]
6. Hollis C, Falconer CJ, Martin JL, et al. Annual research review: digital health interventions for children and young people with mental health problems – a systematic and meta-review. Child Psychology Psychiatry 2017 Apr;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
7. Perry Y, Werner-Seidler A, Calear AL, Christensen H. Web-based and mobile suicide prevention interventions for young people: a systematic review. J Can Acad Child Adolesc Psychiatry 2016;25(2):73-79. [Medline: 27274742]
8. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021 Oct;20(3):318-335. [doi: 10.1002/wps.20883] [Medline: 34505369]
9. Bell IH, Nicholas J, Alvarez-Jimenez M, Thompson A, Valmaggia L. Virtual reality as a clinical tool in mental health research and practice. Dialogues Clin Neurosci 2020 Jun;22(2):169-177. [doi: 10.31887/DCNS.2020.22.2/lvalmaggia] [Medline: 32699517]
10. Dellazizzo L, Potvin S, Luigi M, Dumais A. Evidence on virtual reality–based therapies for psychiatric disorders: meta-review of meta-analyses. J Med Internet Res 2020 Aug 19;22(8):e20889. [doi: 10.2196/20889] [Medline: 32812889]
11. Dewhirst A, Laugharne R, Shankar R. Therapeutic use of serious games in mental health: Scoping review. BJPsych open 2022 Feb 2;8(2):e37. [doi: 10.1192/bjo.2022.4] [Medline: 35105418]
12. Freeman D, Reeve S, Robinson A, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400. [doi: 10.1017/S003329171700040X] [Medline: 28325167]
13. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
14. Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on digital mental health interventions for adolescents and young people: systematic overview. JMIR Ment Health 2021 Apr 29;8(4):e25847. [doi: 10.2196/25847] [Medline: 33913817]
15. Radovic A, Vona PL, Santostefano AM, Ciaravino S, Miller E, Stein BD. Smartphone applications for mental health. Cyberpsychol Behav Soc Netw 2016 Jul;19(7):465-470. [doi: 10.1089/cyber.2015.0619] [Medline: 27428034]
16. Abd-Alrazaq A, Al-Jafar E, Alajlani M, et al. The effectiveness of serious games for alleviating depression: systematic review and meta-analysis. JMIR Serious Games 2022 Jan 14;10(1):e32331. [doi: 10.2196/32331] [Medline: 35029530]
17. Abd-Alrazaq A, Alajlani M, Alhuwail D, et al. The effectiveness of serious games in alleviating anxiety: systematic review and meta-analysis. JMIR Serious Games 2022 Feb 14;10(1):e29137. [doi: 10.2196/29137] [Medline: 35156932]
18. David OA, Costescu C, Cardos R, Mogoaşe C. How effective are serious games for promoting mental health and health behavioral change in children and adolescents? A systematic review and meta-analysis. Child Youth Care Forum 2020 Dec;49(6):817-838. [doi: 10.1007/s10566-020-09566-1]
19. Eve Z, Turner M, Di Basilio D, et al. Therapeutic games to reduce anxiety and depression in young people: a systematic review and exploratory meta-analysis of their use and effectiveness. Clin Psychology Psychoth 2024;31(1):e2938. [doi: 10.1002/cpp.2938]
20. Reynard S, Dias J, Mitic M, Schrank B, Woodcock KA. Digital interventions for emotion regulation in children and early adolescents: systematic review and meta-analysis. JMIR Serious Games 2022 Aug 19;10(3):e31456. [doi: 10.2196/31456] [Medline: 35984681]

21. Hjemdal O, Friborg O, Stiles TC, Martinussen M, Rosenvinge JH. A new scale for adolescent resilience: grasping the central protective resources behind healthy development. Meas Eval Couns Dev 2006;39(2):84-96. [doi: 10.1080/07481756.2006.11909791]
22. Beck JS. Cognitive Behavior Therapy: Basics and Beyond, 2nd edition: Guilford Press; 2011:xix-391.
23. Marton P, Churchard M, Kutcher S. Cognitive distortion in depressed adolescents. J Psychiatry Neurosci 1993 May;18(3):103-107. [Medline: 8499425]
24. Muris P, Field AP. Distorted cognition and pathological anxiety in children and adolescents. Cogn Emot 2008;22(3):395-421. [doi: 10.1080/02699930701843450]
25. Smith EM, Reynolds S, Orchard F, Whalley HC, Chan SW. Cognitive biases predict symptoms of depression, anxiety and wellbeing above and beyond neuroticism in adolescence. J Affect Disord 2018 Dec 1;241:446-453. [doi: 10.1016/j.jad.2018.08.051] [Medline: 30145516]
26. James AC, Reardon T, Soler A, James G, Creswell C. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev 2020 Nov 16;11(11):CD013162. [doi: 10.1002/14651858.CD013162.pub2] [Medline: 33196111]
27. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. Psychol Bull 2006 Jan;132(1):132-149. [doi: 10.1037/0033-2909.132.1.132] [Medline: 16435960]
28. Powell J, Hamborg T, Stallard N, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. J Med Internet Res 2012 Dec 31;15(1):e2. [doi: 10.2196/jmir.2240] [Medline: 23302475]
29. Raes F, Hermans D, van den Broeck K, Eelen P. Kort Instrumenteel: de Nederlandstalige versie van de Dysfunctional Attitude Scale – vorm A (DAS-A-NL). Gedragstherapie 2005;38:285-294 [FREE Full text]
30. Weissman AN, Beck AT. Development and validation of the Dysfunctional Attitude Scale: a preliminary investigation. Presented at: Annual Meeting of the American Educational Research Association; Mar 27 to 31, 1978; Toronto, ON URL: https://eric.ed.gov/?id=ED167619 [accessed 2024-04-26]
31. Maric M, Heyne DA, van Widenfelt BM, Westenberg PM. Distorted cognitive processing in youth: the structure of negative cognitive errors and their associations with anxiety. Cognit Ther Res 2011 Feb;35(1):11-20. [doi: 10.1007/s10608-009-9285-3] [Medline: 21461039]
32. Schwartz JS, Maric M. Negative cognitive errors in youth: specificity to anxious and depressive symptoms and age differences. Behav Cogn Psychother 2015 Sep;43(5):526-537. [doi: 10.1017/S1352465814000228] [Medline: 26202072]
33. De Jaegere E, van Landschoot R, van Heeringen K, et al. The online treatment of suicidal ideation: a randomised controlled trial of an unguided web-based intervention. Behav Res Ther 2019 Aug;119:103406. [doi: 10.1016/j.brat.2019.05.003] [Medline: 31176889]
34. David OA, Cardo RAI, Matu S. Changes in irrational beliefs are responsible for the efficacy of the REThink therapeutic game in preventing emotional disorders in children and adolescents: mechanisms of change analysis of a randomized clinical trial. Eur Child Adolesc Psychiatry 2019 Mar;28(3):307-318. [doi: 10.1007/s00787-018-1195-z] [Medline: 29987559]
35. David OA, Predatu R, Cardo RAI. Effectiveness of the REThink therapeutic online video game in promoting mental health in children and adolescents. Internet Interv 2021 Apr 20;25:100391. [doi: 10.1016/j.invent.2021.100391] [Medline: 33996508]
36. Huen JM, Lai ES, Shum AK, et al. Evaluation of a digital game-based learning program for enhancing youth mental health: a structural equation modeling of the program effectiveness. JMIR Ment Health 2016 Oct 7;3(4):e46. [doi: 10.2196/mental.5656] [Medline: 27717921]
37. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012 Apr 18;344:e2598. [doi: 10.1136/bmj.e2598] [Medline: 22517917]
38. Babbage CM, Jackson GM, Davies EB, Nixon E. Self-help digital interventions targeted at improving psychological well-being in young people with perceived or clinically diagnosed reduced well-being: systematic review. JMIR Ment Health 2022 Aug 26;9(8):e25716. [doi: 10.2196/25716] [Medline: 36018675]
39. Ellis A, Dryden W. The Practice of Rational Emotive Behavior Therapy, 2nd edition: Springer; 2007. [doi: 10.1891/9780826122179]
40. Garrido S, Millington C, Cheers D, et al. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019 Nov 13;10:759. [doi: 10.3389/fpsyt.2019.00759] [Medline: 31798468]
41. Tremain H, McEnery C, Fletcher K, Murray G. The therapeutic alliance in digital mental health interventions for serious mental illnesses: narrative review. JMIR Ment Health 2020 Aug 7;7(8):e17204. [doi: 10.2196/17204] [Medline: 32763881]
42. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
43. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions - a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
44. Shum AK, Lai ES, Leung WG, et al. A digital game and school-based intervention for students in Hong Kong: quasi-experimental design. J Med Internet Res 2019 Apr 5;21(4):e12003. [doi: 10.2196/12003] [Medline: 30950795]

45. Baumel A, Edan S, Kane JM. Is there a trial bias impacting user engagement with unguided e-mental health interventions? A systematic comparison of published reports and real-world usage of the same programs. Transl Behav Med 2019 Nov 25;9(6):1020-1033. [doi: 10.1093/tbm/ibz147] [Medline: 31689344]
46. Borgers N, Sikkel D, Hox J. Response effects in surveys on children and adolescents: the effect of number of response options, negative wording, and neutral mid-point. Qual Quant 2004 Feb;38:17-33. [doi: 10.1023/B:QUQU.0000013236.29205.a6]
47. Thornton L, Gardner LA, Osman B, et al. A multiple health behavior change, self-monitoring mobile app for adolescents: development and usability study of the Health4Life app. JMIR Form Res 2021 Apr 12;5(4):e25513. [doi: 10.2196/25513] [Medline: 33843590]
48. Bryman A. Integrating quantitative and qualitative research: how is it done? Qual Res J 2006 Feb;6(1):97-113. [doi: 10.1177/1468794106058877]
49. Scheerder A, van Deursen A, van Dijk J. Determinants of internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. Telemat Inform 2017 Dec;34(8):1607-1624. [doi: 10.1016/j.tele.2017.07.007]
50. Zayeni D, Raynaud JP, Revet A. Therapeutic and preventive use of video games in child and adolescent psychiatry: a systematic review. Front Psychiatry 2020 Feb 6;11:36. [doi: 10.3389/fpsyt.2020.00036] [Medline: 32116851]

Abbreviations

CBT: cognitive behavioral therapy
REBT: rational emotive behavior therapy

Edited by G Eysenbach; submitted 22.03.23; peer-reviewed by J Chu, P Matthews; revised version received 08.03.24; accepted 24.03.24; published 09.05.24.

Please cite as:
De Jaegere E, van Heeringen K, Emmery P, Mommerency G, Portzky G
Effects of a Serious Game for Adolescent Mental Health on Cognitive Vulnerability: Pilot Usability Study
JMIR Serious Games 2024;12:e47513
URL: https://games.jmir.org/2024/1/e47513
doi:10.2196/47513

Original Paper

Feasibility and Acceptability of a Mobile Game to Support Smoking Cessation: Repeated Measures Study

Shelby Goodwin[1], MPS; Jessica A Nastasi[2], PhD; Schyler T Newman[3], BA; Darion Rapoza[4,5], PhD; Bethany R Raiff[1], PhD

[1]Department of Psychology, College of Science and Mathematics, Rowan University, Glassboro, NJ, United States
[2]Department of Psychology, University of Florida, Gainesville, FL, United States
[3]Department of Psychology, University of Missouri, St. Louis, MO, United States
[4]Division of Cardiology, Department of Medicine, Duke University School of Medicine, Durham, NC, United States
[5]Entertainment Science, Inc, Durham, NC, United States

Corresponding Author:
Bethany R Raiff, PhD
Department of Psychology
College of Science and Mathematics
Rowan University
201 Mullica Hill Road
Glassboro, NJ, 08028
United States
Phone: 1 856 256 4500 ext 53782
Email: raiff@rowan.edu

Abstract

Background: Approximately half of smokers attempt to quit, but 90% of these attempts fail. Video game–based interventions have the potential to address common barriers to evidence-based smoking cessation treatment, including high cost, lack of health care access, and low engagement.

Objective: The purpose of this study was to evaluate the feasibility and acceptability of a video game–based smoking cessation intervention called Inspired and its impact on the 7-day smoking point prevalence at the 30-day follow-up.

Methods: US adults (n=48) were recruited via the web to use Inspired on their smartphones for 7 weeks. The object of the game was to defend a healing tree against attackers. Levels of the game were unlocked twice daily when participants self-reported the number of cigarettes they smoked since the previous entry. Completion of the levels awarded players in-game currency, which could strengthen in-game abilities. Participants received additional in-game rewards to aid gameplay by submitting either smoking self-reports only or self-reports indicating abstinence, determined through random assignment. In addition, participants completed a web-based survey at intake, week 4, week 7, and the 30-day follow-up.

Results: Of the 48 participants, who had an average age of 39.8 (SD 10.7) years, 27 (56%) were female, 4 (8%) Hispanic, 37 (77%) White, and 27 (56%) employed; 26 (54%) earned <US $40,000 a year; and 14 (29%) lived in nonurban areas (eg, rural and suburban). There were no significant differences between the groups, so all outcome data were combined. Participants averaged 20.6 (SD 15.3) days of gameplay and reached a mean highest game level of 10.7 (SD 8.4), although there was a high degree of variability. Participants reported abstinence on 31.4% (SD 38.2%) of all cessation phase reports and averaged 5.4 (SD 9.8) consecutive abstinent smoke reports. For every 1 SD increase in the highest level achieved, there was a 27% increase in the percentage of abstinent samples and a 405% increase in longest continuous abstinence. At the 30-day follow-up, 23% (11/48) of the participants reported having not taken a cigarette puff in the prior 7 days and 31% (15/48) had spent at least 24 hours without smoking in the prior 14 days. On an 11-point scale, participants rated the intervention moderately favorably: if they had to do it again, they would use Inspired to help them quit (mean 6.4, SD 3.4), and Inspired was helpful in their current attempt to quit (mean 5.4, SD 3.6).

Conclusions: These results support the acceptability of Inspired. Although high dropout rates prevent conclusions on feasibility, a subset of the participants responded favorably. Scalable and accessible video game–based smoking cessation interventions could be the key to addressing the foremost cause of preventable morbidity and mortality in the United States.

Trial Registration: ClinicalTrials.gov NCT03929003; https://clinicaltrials.gov/ct2/show/NCT03929003

(*JMIR Serious Games 2024;12:e54684*) doi:10.2196/54684

KEYWORDS

mHealth; mobile health; smoking cessation; video game intervention; mobile phone

Introduction

Background

Although the prevalence of smoking has steadily declined over the past several years, cigarette smoking remains the foremost cause of preventable morbidity and mortality in the United States, accounting for approximately 480,000 deaths each year [1]. Most smokers express an interest in quitting, and approximately half of smokers attempt to quit; however, approximately 90% of quit attempts fail [2]. Evidence-based interventions such as cognitive behavioral therapy and pharmacotherapy are most likely to result in smoking cessation; however, only 31% of smokers used counseling or medication in their quit attempt, and fewer still used these strategies if they were of low socioeconomic status [2-4]. Therefore, the development of effective interventions must include efforts to optimize their accessibility and appeal to all smokers who wish to quit.

Mobile health phone apps have substantially greater reach than in-person health visits and may reach populations of smokers who wish to quit but would not otherwise access in-person smoking cessation assistance. Mobile apps hold promise in their flexibility, patient-directed treatment options, and the ability to reach individuals who are not comfortable discussing smoking cessation with a health care professional. These apps vary widely in their level of interactivity, and participant retention remains low [5,6]. Attempts to increase app use have included gamification (ie, applying game-like elements without the context of an actual game) such as goal setting, progress tracking, and earning markers such as badges [7]. Use of these features has been associated with increased motivation to quit [8], but efforts to increase subjective enjoyment of the interventions may further expand their use.

One approach that may increase subjective enjoyment while maintaining accessibility is the use of video game–based smoking cessation interventions, the main component of which consists of a playable game [9-17]. In the United States, approximately 62% of adults report playing video games, and 75% of players spend ≥3 hours playing games per week [17]. One survey conducted among smokers indicated that 75% of participants already played video games, and 65% to 67% of participants believed that a video game–based smoking cessation intervention could help them or someone they knew quit smoking [18]. With 64% of video game players in the United States reporting the use of their smartphone to play games [19], smartphones provide an accessible platform on which to develop an intervention that is scalable, acceptable, and enjoyable.

Multiple video game–based smartphone smoking cessation interventions have been developed and tested, including Inner Dragon [16], Cigbreak [17], Tobstopp [12], QuitIt [15], and Quittr [13]. A recent review examined other smoking intervention games, highlighting 7 studies that featured digital cessation games. These games drew on a range of theories including social cognitive theory and cognitive behavioral therapy, and they had a wide variety of game features including earning points and other rewards as well as interactive storytelling. Results were mixed, with approximately half of the studies achieving significant effects on either the smoking status or the number of cigarettes smoked, and most studies displayed high rates of dropout [14].

In an attempt to improve cessation outcomes seen in video game–based smartphone interventions, we previously developed and evaluated a prototype of a stand-alone video game called Inspired [9]. Inspired was unique because it drew from the science of contingency management, which is a powerful intervention consisting of delivering incentives when individuals meet operationally defined and objectively verified behavioral goals [20]. One common criticism of contingency management is the cost of these typically monetary incentives, so Inspired used in-game rewards whose only value was in the context of the game [21,22]. The original intention of the game was to reward users for providing objective evidence in the form of breath carbon monoxide (CO) for meeting abstinence goals, with the goal that this would function as a positive reinforcer for smoking abstinence [23]. Adult smokers recruited via the web (N=28) viewed the prototype favorably and indicated that they thought a fully developed version of the game had the potential to be fun (100%) and help them quit smoking (71%). Participants also provided feedback for improving the gameplay, such as adding variation to game levels to reduce monotony and adding narrative to engage the player.

To continue this work, a beta version of Inspired was designed and developed by an interdisciplinary team of researchers and game designers based on the prototype evaluation. The narrative of the game explained that the players were tasked with saving a village and its people from invading enemies known as "creatures of darkness." At each level of the game, players needed to build defensive structures to protect a healing tree from the creatures. Game design elements included a core game system (ie, levels where the creatures were battled and earnings could be obtained), a metasystem (ie, where earnings could be spent to customize villages and upgrade structures), and a smoking cessation program in which players could earn in-game rewards for complying with the smoking cessation program's objectives.

Objectives

The original aim of the study was to examine the impact of in-game rewards dependent on objectively verified smoking abstinence on the 7-day smoking point prevalence at the 30-day follow-up. However, this objective was shifted due to technical difficulties with objective verification measures and insufficient impact of abstinence-dependent rewards (explained in the Methods section). Therefore, the purpose of this study was to evaluate the feasibility and acceptability of the beta version of

Inspired as a smartphone-based smoking cessation intervention to inform future research on smoking cessation games. The main outcome remained the 7-day smoking point prevalence at the 30-day follow-up; secondary outcomes included the longest self-reported abstinence streak and percentage of abstinent self-reported submissions during the abstinence phase.

Methods

Ethics Approval

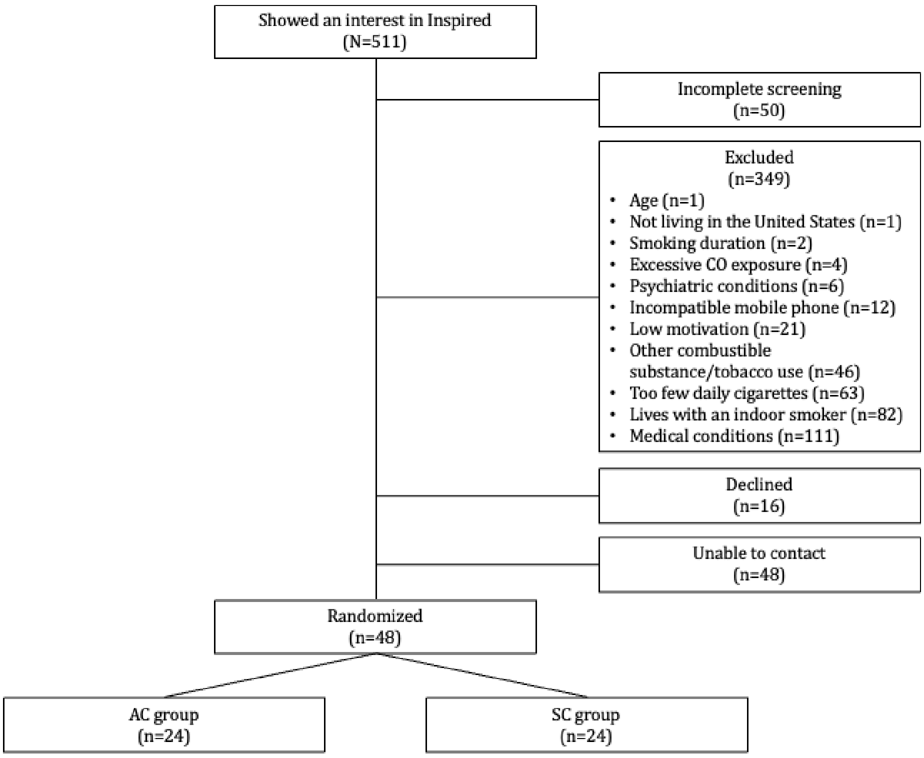
All procedures were reviewed and approved by the Rowan University Institutional Review Board (IRB-2014-170).

Participants

Participants were recruited on the web through Craigslist, Facebook (Meta Platforms, Inc), Google Ads (Google LLC), and Mechanical Turk (MTurk, Amazon). Multiple advertisements with similar but varying text (eg, "test out a mobile video game designed to help you quit"), images (eg, a screenshot of gameplay), and a link to a screening survey were included. The study was submitted to ClinicalTrials.gov on September 26, 2017, and was published on the website on April 24, 2019. Participants were recruited between May 3, 2019, and July 25, 2019. The screening survey was used to obtain information on age, smoking history, desire to quit, prior experience with video games, access to mobile devices, and contact information (refer to the study by Upton et al [11] for a secondary analysis of the screening data). Inclusion criteria required participants to (1) be of legal age to purchase cigarettes in their state or jurisdiction, (2) smoke ≥10 cigarettes per day, (3) report smoking for at least 2 years, (4) report a desire to quit of ≥8 on a 10-point Likert scale, and (5) own a smartphone using the Android operating system and have an active data plan or access to Wi-Fi to connect to the internet. Exclusion criteria were (1) smoking cannabis or using other tobacco products >2 times per month, (2) having unstable medical or psychiatric conditions that may interfere with submitting accurate CO samples (eg, asthma and lactose intolerance), and (3) living with someone who smokes in the house and (4) having a career that results in excessive exposure to CO (both of which would interfere with CO sample accuracy). Respondents were also asked whether they played video games and the types of games they played. If respondents reported that they did not play games, the researcher followed up by asking whether they were interested in playing a game designed to help them quit smoking.

A total of 511 respondents completed the screening survey, of whom 113 (22.1%) met the inclusion criteria. Respondents who met the inclusion criteria were contacted by email or phone so that study staff could explain additional information about the study and the individual could be fully informed before deciding to participate. A total of 48 participants were enrolled (see Figure 1 for a CONSORT [Consolidated Standards of Reporting Trials] diagram). Analyses were conducted in SPSS (IBM Corp) and R software (R Foundation for Statistical Computing).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram. AC: abstinence contingent; CO: carbon monoxide; SC: submission contingent.



The original recruitment goal was 114 participants based on a power analysis using an α level of .05 and an anticipated odds ratio of 6.68 to achieve 80% power when analyzing the effect of group on the dichotomous 7-day smoking point prevalence. However, due to considerable technical difficulties with collecting objective measures of smoking, recruitment was delayed, and a final sample of 48 participants was achieved. Initially, participants were required to submit breath CO samples to verify smoking abstinence within the last 12 hours or more using a portable breath CO detector (iCO) [24], which was attached to the participant's phone via the headphone jack. Furthermore, to verify the identity of the participant submitting the CO sample and subsequent smoking self-report, participants were required to use facial recognition software that was built into the game. Due to technical difficulties with the vendor-supplied hardware and software, both breath CO monitoring and facial recognition technologies were removed from the procedures after a pilot trial of participants. We transitioned to relying exclusively on smoking self-reports to determine smoking status for the purposes of in-game rewards, which is the standard practice for most app-based smoking cessation trials [25]. It should be noted that the final analytical sample comprised participants who exclusively reported their smoking through self-report as opposed to CO monitoring.

Participants completed the psychosocial history questionnaire at intake, which included questions about demographic information, smoking history, and cessation history. Questions on smoking and cessation history included current tobacco use patterns (eg, the number of cigarettes smoked per day and use of other tobacco products), smoking onset (eg, age at which the first cigarette was smoked and age at which the participants started smoking daily), and quitting behavior (eg, number of previous quit attempts). Participants rated how much they wanted to quit smoking on an 11-point scale (0-10), with higher ratings representing a greater desire to quit.

Participants completed the Fagerström Test for Nicotine Dependence during intake, which consisted of 6 items. The total score summed the responses and represented the participant's current level of cigarette dependence. A total score of 1 to 2 indicated low dependence, 3 to 4 indicated low to moderate dependence, 5 to 7 indicated moderate dependence, and ≥8 indicated high dependence [26].

The Video Game Use Survey is a 19-item measure that was administered during intake [18]. Questions assessed game playing history (eg, duration played per week) and preference for games (eg, action and adventure; strategy and puzzle) in an open-ended response format.

Study Phases

The game was designed for 7 weeks of gameplay. Participants completed web-based survey assessments at intake, week 4, week 7, and the 30-day follow-up. Assessment content included psychosocial history, video game use, and perceptions of the intervention. Data from intake, week 7, and the 30-day follow-up were used in these analyses. Survey links were sent via email, and compensation was loaded after response completion was confirmed. Participants were mailed debit cards (CT Payer), through which compensation was delivered after completion of each battery of surveys (ie, setup: US $10, week 4: US $15, week 7: US $20, and 30-day follow-up: US $30). The compensation provided for survey completion was not connected to participant cessation efforts. Each participant was told they had the opportunity to earn a total of US $75 in compensation if all surveys were completed, regardless of whether they quit smoking.

During intake, the researcher walked each participant through downloading the Inspired app and submitting their first smoking self-report. Smoking reports, explained in more detail in the Smoking Self-Report in the Methods section, consisted of reporting the number of cigarettes smoked since the last report. Participants had the opportunity to submit 2 reports a day, waiting at least 8 hours between reports.

In the baseline phase (days 1 to 3), all participants received in-game rewards (detailed in the Gameplay section) for submitting each smoking self-report and a silver multiplier for every 3 consecutive smoking self-report submissions; however, participants were not given any smoking goals. After baseline, graduated smoking reduction and cessation goals for the remainder of the study were provided to all participants. During the tapering phase (days 4 to 7), participants were given decreasing daily target smoking caps, calculated based on the mean number of cigarettes reported during baseline, such that the final goal of the tapering phase would be 0 (eg, mean 20 cigarettes divided by 8=goal reduced by approximately 2 to 3 cigarettes per required smoke report). During the abstinence phase (weeks 2 to 4) and thinning phase (weeks 5 to 7), smoking goals were always set to 0. Initially, the schedule of rewards was programmed to end in the thinning phase. A thinning phase is common in reinforcement-based interventions and consists of gradually reducing the frequency of reinforcer delivery to reduce dependence on the reinforcer while maintaining high rates of the target behavior. This helps an individual transition to maintaining abstinence, even when the reinforcer is no longer available [27]. However, due to low rates of engagement in the thinning phase, participants received rewards at the same frequency as the abstinence phase until the end of the thinning phase.

A repeated measures group design was used. Upon enrollment, participants were randomly assigned to either a submission contingent (SC) or an abstinence contingent (AC) experimental group. All participants were given goals for smoking reduction and eventual abstinence. The SC group earned "passes" for submitting smoking self-reports on schedule, regardless of whether they reported meeting their assigned smoking reduction or abstinence goals. The AC group earned "passes" only for submitting self-reports that met their assigned smoking reduction or abstinence goals. Submissions reporting cigarette use that did not meet their smoking reduction or abstinence target were scored as a miss. For both SC and AC groups, failure to submit a self-report by the reporting deadline was scored as a miss. Thus, the only difference between the SC and AC groups was what qualified as "passing" to earn the in-game rewards. The groups were designed to compare the impact of in-game reinforcement on cessation efforts and to determine if requiring abstinence to receive rewards would be more effective than the act of reporting smoking behaviors alone, as we have done with monetary-based contingency management [28,29]. We hypothesized that the group that received in-game rewards for their abstinence would be more likely to quit smoking than the group that received in-game rewards for submitting their smoke reports only.

We compared the SC and the AC groups at intake to detect systematic group differences based on reported demographics, smoking history, cessation history, nicotine dependence, and game playing history. The 2 groups were also compared on the proportion of participants completing baseline, tapering, abstinence, and thinning phases to determine differences in dropouts.

Gameplay

Each level in the core game began with planting a seed to grow a new healing tree. Players needed to repeatedly build and activate defensive structures to fight incoming enemies and to activate healing structures to help the tree grow. At the start of each level, players were given a bank of gems, which varied in color and shape. Players could choose to activate either a defensive or healing structure for a period of time by placing 3 gems from their bank into each structure, after which their bank would replenish with 3 randomly generated gems. Players could either place 3 random gems to power the structure for a short time or match the gem type to the "recipe" shown on the structure to power the structure for a longer period, creating a larger effect for that structure. The level was completed when the player placed an adequate number of gem sequences to sufficiently power the defense structures to push back enemies and let the healing structures fully grow their tree. Players could earn 1, 2, or 3 stars depending on how quickly the level was completed, which contributed to the player's total score, representing the accumulated points earned across overall levels. The number of stars earned also corresponded with the amount of silver, which functioned as the in-game currency gained upon level completion.

The first 3 levels of the game served as tutorials with instructions about how to play, and they were designed to be easy to win to engage players early on. As the game continued, players were exposed to levels with progressing difficulty, which required more advanced defense structures and strategic decision-making to defeat the new enemies.

In addition to playing the levels of the core game, players could access a metasystem, which was an area for activities performed outside of the core game levels. All players had a village that they could customize in the metasystem. Silver was the primary currency, earned through core level completion, that could be used to purchase cosmetic additions to customize the village and upgrade the strength of the defensive and healing structures used in the core game. Village customization was for aesthetics only and served no functional purpose, but defensive and healing structure upgrades were critical to the players' progression through more challenging levels in the core game (see Figures 2A-2C for screenshots of each game component).

Figure 2. Screenshots of Inspire. (A) core game level, (B) core game level, (C) metasystem, (D) smoking report results, and (E) silver multiplier for passing smoking report (abstinence-contingent [AC] group).

Game use outcomes included the total numerical score achieved in the game, with higher numbers indicating higher game use and achievement in completing levels. The total number of stars earned was calculated by adding the 1, 2, or 3 stars granted for the completion of each level (including replayed levels). The highest level achieved in the game and days of gameplay were also reported. Finally, the total number of defense and healing structures built in the metasystem was summed.

The Game Addiction Scale [30] assessed for indicators of gaming addiction. A total of 9 questions from the problems, conflict, and relapse subscales were rated on a 5-point Likert scale and assessed during participant intake and week 7. Scores were summed to indicate a total score of game addiction, with higher scores denoting more frequently endorsed gaming-related addictive behaviors (range 9-45). The SC and the AC groups were compared on game use (ie, the highest level, total score, total number of stars, and total number of buildings) and game addiction total scores.

Smoking Self-Reports

As mentioned previously, players were told to submit a smoking self-report twice per day. These reports were separated by at least 8 hours and stated whether they had smoked since submitting their last report; if so, participants reported the number of cigarettes they had smoked [29]. Participants were sent push notifications that signaled when it was time to submit each report and the time by which the reports were due to avoid being recorded as “missed.”

Upon submission of each smoking self-report, the game displayed a bar graph showing the number of cigarettes reported in each prior submission so that the participants could track their smoking cessation progress over time (Figure 2D). Submission of each smoking self-report unlocked the next level in the core game system, regardless of self-reported smoking or abstinence. Therefore, players could unlock a maximum of 2 new levels per day. Participants could also replay any unlocked level as often as they liked to earn additional silver (ie, in-game currency).

Players earned in-game rewards for submitting a "passing" smoking self-report, the criteria for which depended on their group assignment (refer to the Study Phases in the Methods section). These rewards consisted of a gem prism and a silver multiplier. The gem prism was earned for every pass and could be used to select a gem color of their choice in the core game, which players could use to more easily meet the "recipes" required to power their defensive or healing structure for a longer period if their randomly generated gem bank had insufficient gems available for the recipe. Although players could match gems at random to power their structures for a shorter time, gem prisms earned through passes made fulfilling gem recipes easier, allowing for a longer structure activation and making completion of the level more likely. A silver multiplier was earned for submitting 3 consecutive passes and increased the amount of silver earned by each completed level in the core game. Multiplier values included 1×, 2×, 4×, or 7× and increased with every 3 consecutive passing smoke reports (Figure 2E). Larger quantities of silver earned through passing smoke reports could be used to more frequently upgrade the defensive and healing structures in the metasystem. This resulted in stronger structure performance, which was particularly important for later, more challenging levels.

Smoking self-report outcomes included the percentage of smoking self-reports submitted (out of the total requested reports over the course of gameplay) and the percentage of smoking self-reports indicating abstinence (out of the total submitted reports). The percentage of smoking self-reports indicating abstinence was determined for all reports throughout the gameplay and calculated for just the abstinence phase, where the goal for cigarette use was set to 0. The duration of the abstinence streak was calculated based on the highest number of consecutive reports of 0 cigarettes smoked throughout the study. The self-report submissions informed treatment dropout, which was based on the completion of each phase of the study (eg, baseline, tapering, abstinence, and thinning). Specifically, participants were scored as completing a phase if they submitted at least 1 smoking self-report in the final 2 days of the phase. The follow-up survey assessed whether participants continued to smoke once participation concluded (ie, 30 days after week 7). Participants were asked whether they had spent 24 hours without a cigarette in the previous 2 weeks and whether they had not taken a puff from a cigarette in the past 7 days. We compared the SC and the AC groups on all smoking cessation outcomes (ie, percentage of abstinent submissions during the abstinence phase and the longest duration of abstinence streak).

Game Use on Smoking Abstinence

The proposed analyses originally planned to compare the 2 study groups, SC and AC, in the relationship between game use and smoking abstinence. This would have tested whether in-game rewards for reporting smoking abstinence were more effective in promoting cessation than rewards for reporting smoking independent of meeting abstinence goals. However, we found no systematic differences between the groups on demographics, smoking history, game use, or self-reported abstinence. In this study, earning in-game rewards (submission only or meeting abstinence goals) did not seem to impact gameplay or cessation. Therefore, the 2 groups were combined into 1 analytical group for the remaining analyses.

A total of 2 gamma generalized linear models with a log link function were created to examine the impact of total game score and the highest level achieved on both the percentage of abstinent submissions during the abstinence phase and the longest duration of abstinence. The models were conducted both with and without an interaction term between the total game score and the highest level achieved to determine the necessity of its inclusion, using the Akaike Information Criterion, Bayesian Information Criterion, and Bayes Factor as selection criteria. A generalized linear model was used to account for data heteroscedasticity, and a gamma distribution was used to accommodate a positive skew on both models.

Cotinine Testing

Saliva cotinine testing was conducted at intake, week 2, week 4, week 7, and the 30-day follow-up to confirm smoking abstinence. Testing kits (NicAlert) [31] were mailed to all eligible participants to biologically verify self-reported tobacco use. If participants reported using cessation strategies that could elevate their cotinine levels (eg, nicotine replacement therapy or vaping nicotine), cotinine testing was not conducted. Once the kit was received, a member of the research team scheduled a video call to walk the participant through completing the test at home. After testing procedures were completed, participants submitted an image of the strip results to the researcher. A total of 2 observers independently scored all strip results and were unaware of the time point at which the test was conducted. The interobserver agreement was calculated by dividing the number of strip results agreed by the total number of strips obtained, multiplied by 100 to obtain a percentage. The interobserver agreement was 91.7% for all strip results. Participants who were eligible to submit cotinine samples received US $5 for each submission.

Verification of smoking abstinence was determined by comparing the cotinine results (positive or negative) with the most recently submitted smoking self-report on the same day the sample was collected. Participants met the correspondence criteria when both the cotinine test and the smoking self-report indicated tobacco abstinence or tobacco use. The percentage of corresponding results across all participants and time points was reported.

Treatment Acceptability

The Treatment Acceptability Questionnaire was a 38-item survey developed to assess the perceived effectiveness of the intervention, gameplay feedback, and a comparison of how Inspired was viewed relative to other cessation strategies. Statements were rated on an 11-point scale, with 0 meaning "definitely not" to 10 meaning "yes, absolutely." Participants were also asked to rank 10 features of the game from most to least valuable (1 being "most valuable" to 10 being "least valuable"). This questionnaire was administered at week 7.

Results

Participants

A total of 48 participants were randomized (n=24, 50% per group), and a summary of their characteristics can be found in Table 1. Of those participants, 27 (56%) were female, 4 (8%) were Hispanic, and 37 (77%) were White, and they had an average age of 39.8 (SD 10.7) years. A third of the sample (n=17, 35%) had never been married and most either reached high school graduation (n=15, 31%) or completed some college (n=19, 40%). Approximately half of the participants reported being employed (n=27, 56%), earning <US $40,000 a year (n=26, 54%), and renting a home (as opposed to living in a home they or someone else owned; n=26, 54%). Geographically, 14 (29%) participants lived in nonurban areas (eg, rural or suburban) [32].

Table 1. Participant demographics.

	Total sample (n=48)	AC[a] group (n=24)	SC[b] group (n=24)
Age (y), mean (SD)	39.8 (10.7)	43.0 (12.2)	36.7 (8.2)
Sex[c], n (%)			
Male	20 (42)	11 (46)	9 (38)
Female	27 (56)	13 (54)	14 (58)
Other	1 (2)	0 (0)	1 (4)
Ethnicity, n (%)			
Hispanic	4 (8)	2 (8)	2 (8)
Not Hispanic	44 (92)	22 (92)	22 (92)
Race, n (%)			
American Indian or Alaska Native	1 (2)	1 (4)	0 (0)
Black	7 (15)	3 (12)	4 (17)
White	37 (77)	19 (79)	18 (75)
Mixed race	3 (6)	1 (4)	2 (8)
Marital status, n (%)			
Currently married	15 (31)	8 (33)	7 (29)
Widowed	2 (4)	0 (0)	2 (8)
Divorced	9 (19)	4 (17)	5 (21)
Separated	5 (10)	4 (17)	1 (4)
Never married	17 (35)	8 (33)	9 (38)
Education, n (%)			
Less than high school	3 (6)	3 (12)	0 (0)
High school graduate or equivalent	15 (31)	6 (25)	9 (38)
Some college	19 (40)	10 (42)	9 (38)
Associate degree	8 (17)	3 (12)	5 (21)
Bachelor degree	3 (6)	2 (8)	1 (4)
Employment, n (%)			
Employed	27 (56)	13 (54)	14 (58)
Student	1 (2)	0 (0)	1 (4)
Not employed	20 (42)	11 (46)	9 (38)
Housing, n (%)			
Owned by the participant or household member with mortgage or loan	13 (27)	6 (25)	7 (29)
Owned by the participant or household member free and clear	7 (15)	3 (12)	4 (17)
Rented	26 (54)	14 (58)	12 (50)
Occupied without rent	2 (4)	1 (4)	1 (4)
Living in nonurban area, n (%)			
Yes	14 (29)	6 (25)	8 (33)
No	34 (71)	18 (75)	16 (67)
Yearly income (US $), n (%)			
20,000	12 (25)	6 (25)	6 (25)
20,000 to 39,999	14 (29)	8 (33)	6 (25)
40,000 to 59,999	11 (23)	6 (25)	5 (21)

	Total sample (n=48)	AC[a] group (n=24)	SC[b] group (n=24)
≥60,000	11 (23)	4 (17)	7 (29)

[a]AC: abstinence contingent.

[b]SC: submission contingent.

[c]Although gender would have been the more appropriate metric, participants were asked to indicate their sex as opposed to gender at the time of assessment.

Detailed smoking history of the participants is presented in Table 2. All participants reported currently smoking and 38% (18/48) of the participants said they also currently smoke other substances (including cannabis). On the Fagerström Test for Nicotine Dependence, participants scored an average of 6.5 (SD 1.5), indicating moderate nicotine dependence. Participants endorsed playing games (on any device, including their mobile phone) for an average of 27.1 (SD 25.4) hours per week at baseline.

Table 2. Smoking history.

	Total sample (n=48)	AC[a] group (n=24)	SC[b] group (n=24)
Smoking other than cigarettes, n (%)			
Never or extremely rarely	24 (50)	12 (50)	12 (50)
Quit, but used to smoke	6 (12)	3 (12)	3 (12)
Yes, currently smoking	18 (38)	9 (38)	9 (38)
Number of cigarettes smoked per day, mean (SD)	15.9 (8.8)	17.5 (9.7)	14.4 (7.7)
Duration of smoking (years), mean (SD)	20.2 (12.3)	23.9 (13.6)	16.5 (9.9)
Age at which the first cigarette was smoked (years), mean (SD)	16.7 (4.9)	15.4 (4.6)	17.9 (4.9)
Age of smoking regularly (years), mean (SD)	19.3 (4.9)	18.7 (4.7)	19.9 (5.1)
Desire to quit (0-10), mean (SD)	8.9 (1.3)	8.7 (1.1)	9.2 (1.4)
Number of quit attempts, mean (SD)	5.4 (5.5)	5.4 (4.3)	5.4 (6.5)
FTND[c] total, mean (SD)	6.5 (1.5)	6.6 (1.7)	6.4 (1.3)

[a]AC: abstinence contingent.

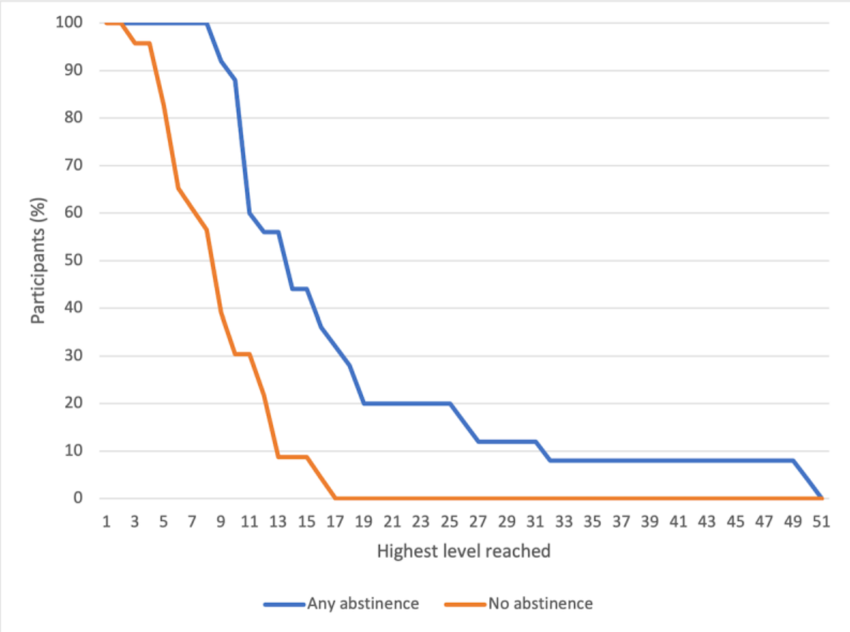
[b]SC: submission contingent.

[c]FTND: Fagerström Test for Nicotine Dependence.

Study Phases

Most participants completed the baseline (45/48, 94%) and tapering phases of the study (26/48, 75%). Of the total 48 participants, 20 (42%) continued to submit smoking self-reports through the end of the abstinence phase at week 4. Very few participants (n=7, 15%) submitted smoking self-reports by the end of the thinning phase (week 7). A survival curve of the highest level achieved by any abstinence reported in the abstinence phase is presented in Figure 3.

Figure 3. Survival curve of the highest level achieved by any abstinence reported in the abstinence phase.



Gameplay

The participants played the game for an average of 20.6 (SD 15.3) days over the 7 weeks of the study. They scored an average of 30,187.5 (SD 33,662.6; range 0-220,925) points and earned an average of 35.6 (SD 34.4) stars from level completions. The mean highest level reached in the game was 10.7 (SD 8.4). In the metasystem, players either built or upgraded an average of 16.4 (SD 17.6) structures. The average video game addiction score (on a scale from 9 to 45) remained stable, with scores of 14.6 (SD 4.5) at baseline versus 14.0 (SD 3.9) at week 7. Game use results are presented in Table 3.

Table 3. Game use.

	Total sample (n=48), mean (SD)	AC[a] group (n=24), mean (SD)	SC[b] group (n=24), mean (SD)	*P* value
Days of gameplay	20.6 (15.3)	18.4 (12.9)	22.8 (17.4)	.32
Highest level achieved in the game	10.7 (8.4)	8.6 (7.1)	12.8 (9.2)	.08
Total number of structures	16.4 (17.6)	14.2 (14.9)	18.6 (20.0)	.39
Total score on the game	30,287.5 (33,662.6)	22,122.9 (21,375.4)	38,452.1 (41,467.9)	.10
Total number of stars	35.6 (34.4)	26.7 (24.3)	44.5 (40.8)	.07

[a]AC: abstinence contingent.

[b]SC: submission contingent.

Smoking Self-Reports

A detailed summary of the smoking self-report results is presented in Table 4. Of the 98 total possible report submissions over the 7 weeks of the study, participants submitted an average of 34.7 (SD 29.1; 34.7/98, 35%) reports. Participants reported smoking abstinence on 28.7 (SD 32.5; 28.7/98, 29%) reports of all 98 smoking self-reports submitted over the entirety of the intervention and on 13 (SD 38.2; 13/42, 31%) reports of all 42 reports submitted in the abstinence phase only. During the abstinence phase, 25 (52%) of the 48 participants submitted at least 1 abstinent report, meaning that nearly half of the participants did not have any abstinent reports during this phase. On average, participants reported 5.4 (SD 9.8) consecutive abstinent smoke reports.

Table 4. Smoke report summary.

	Total sample (n=48), mean (SD)	AC[a] group (n=24), mean (SD)	SC[b] group (n=24), mean (SD)	*P* value
Longest abstinence streak	5.4 (9.8)	5.3 (8.3)	5.5 (11.2)	.94
Percent smoke report submitted	34.7 (29.1)	31.5 (30.3)	37.8 (28.0)	.46
Percentage of abstinent smoke reports (out of all smoke report submissions across all phases)	28.7 (32.5)	26.9 (31.3)	30.5 (24.3)	.70
Percentage of abstinent smoke report (out of all smoke report submissions in the abstinence phase only)	31.4 (38.2)	34.1 (41.7)	28.6 (34.9)	.63

[a]AC: abstinence contingent.

[b]SC: submission contingent.

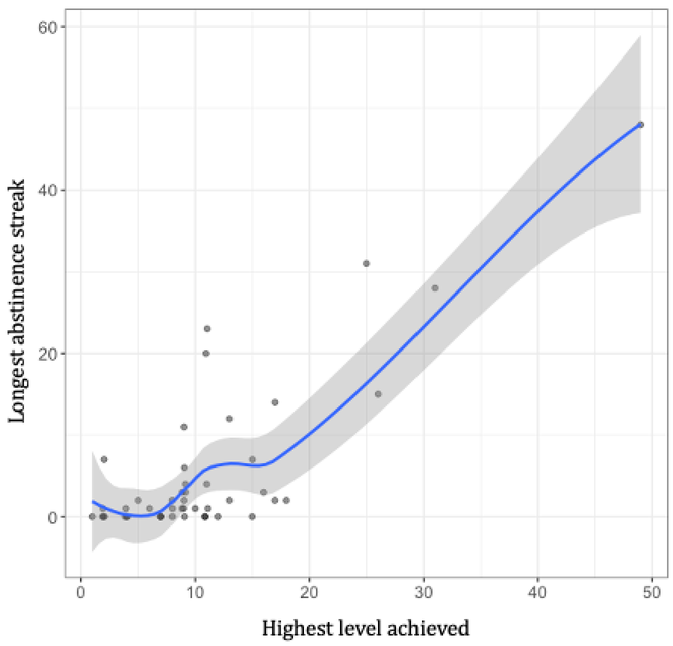
The main outcome of this examination was the 7-day smoking point prevalence at the 30-day follow-up. Of the 30 participants (across both groups) who completed the follow-up survey 30 days after the end of the intervention, 11 (37%) had not taken a puff of a cigarette in the previous 7 days and 15 (50%) said they spent at least 24 hours without smoking in the prior 2 weeks. Overall, 15 (50%) participants reported not smoking for 24 hours in the previous two weeks (AC=4, 36%; SC=11, 58%; *P*=.45) and 11 (37%) participants reported not taking a single puff from a cigarette in the previous 7-days (AC=4, 36%; SC=7, 37%; *P*=.65).

When individuals who did not respond to the survey were conservatively considered not abstinent from smoking, 11 (23%) of the 48 participants had not taken a puff of a cigarette in the previous 7 days and 15 (31%) said they spent at least 24 hours without smoking in the prior 2 weeks. Fifteen (31%) of participants reported not smoking for 24 hours in the previous 2 weeks (AC=4, 17%; SC=11, 46%; *P*=.06) and 11 (23%) reported not taking a single puff from a cigarette in the previous 7-days (AC=4, 20%; SC=7, 29%; *P*=.49).

Game Use on Smoking Abstinence

A total of 2 generalized linear models were conducted, first examining the impact of total game score and the highest level achieved on the percentage of abstinent submissions during the abstinence phase and the second examining the impact of total game score and the highest level achieved on the longest duration of continuous abstinence. Both models favored the exclusion of the interaction term, so only the individual contributions of the total game score and the highest level achieved were considered. Every 1 SD increase in total game score was associated with an 11% decrease in the percentage of abstinent submissions, whereas every 1 SD increase in the highest level achieved was associated with a 27% increase in the percentage of abstinent submissions. Similarly, an increase by 1 SD in the total score was associated with a 51% decrease in the longest abstinence streak, whereas an increase by 1 SD in the highest level achieved was associated with a 405% increase in the longest duration of continuous abstinence. In sum, the highest level achieved, but not the total game score, was positively related to both the percentage of abstinent submissions and the longest abstinence streak, with the stronger effect demonstrated between the highest level achieved and the longest duration of continuous abstinence (Figure 4).

Figure 4. The highest level achieved by the longest abstinence streak. The longest abstinence streak is calculated by determining the maximum number of continuously reported abstinent submissions over the course of the study.



Cotinine Testing

Throughout the study, 20 cotinine samples were collected across all participants and at all time points. These 20 samples were compared with the number of cigarettes smoked as self-reported by the participants, 7 (35%) of whom reported tobacco abstinence. There was a 95% (19/20) correspondence between self-reported tobacco use and biologically verified cotinine results across all samples and a 100% (20/20) correspondence for the 7 abstinent reports.

Treatment Acceptability

On an 11-point scale, ranging from 0 (“definitely not”) to 10 (“yes, absolutely”), participants rated the intervention moderately favorably. Specifically, they said that if they had to do it over again, they would use Inspired to help them quit smoking (mean 6.4, SD 3.4), that Inspired was helpful in their current attempt to quit (mean 5.3, SD 3.6), and that they would recommend Inspired to a friend (mean 6.5, SD 3.2). Participants felt the game was moderately fun to play (mean 6.7, SD 3.3), found the in-game rewards as motivating for them to submit tobacco-free smoke reports (mean 7.0, SD 3.6), and felt that the connection between in-game rewards and meeting smoking cessation goals was clear (mean 6.1, SD 3.9). More detailed results are presented in Table 5. Of the 10 game elements listed in the questionnaire, participants rated having specific-colored gems, being able to upgrade buildings to fight darkness, having waves at the end of the level that killed all remaining darkness, and being able to upgrade buildings to produce good health as the most valuable features. Apart from an open-ended “other” category, participants reported that the least valuable elements were silver to customize their village and the silver multiplier earned with consecutive tobacco-abstinent submissions.

Table 5. Treatment acceptability[a] (n=30).

	Total sample (n=30), mean (SD)	AC[b] group (n=11), mean (SD)	SC[c] group (n=19), mean (SD)	*P* value
If you had to do it over again, would you use Inspired to help you quit smoking?	6.4 (3.4)	7.1 (3.8)	6.1 (3.3)	.44
Was Inspired helpful to you in your attempt to quit smoking?	5.4 (3.6)	5.5 (3.6)	5.3 (3.6)	.92
Do you think the Inspired game is fun to play?	6.7 (3.3)	7.3 (3.8)	6.4 (3.0)	.47
Would you recommend Inspired to a friend?	6.5 (3.2)	7.5 (3.1)	6.0 (3.2)	.22
I felt motivated to submit clean samples in order to obtain cessation points[d]	7.0 (3.6)	7.6 (3.2)	6.6 (3.9)	.45
It was clear to me how I could use the rewards I earned from meeting my goals	6.1 (3.9)	6.4 (4.3)	6.0 (3.8)	.81

[a]Scored on an 11-point scale, with 0 meaning "definitely not" to 10 meaning "yes, absolutely."

[b]AC: abstinence contingent.

[c]SC: submission contingent.

[d]Although "abstinent" samples would be a more appropriate way to refer to these smoke reports, participants were asked to indicate their motivation to submit "clean" samples at the time of assessment.

Discussion

Principal Findings

The study results supported the preliminary treatment acceptability for the use of the Inspired mobile game as a smoking cessation intervention. Treatment dropout was high, and both game engagement and tobacco abstinence were highly variable, decreasing feasibility. Half of the participants (25/48, 52%) submitted at least 1 abstinent report during the abstinence phase, and game scores had a large range, suggesting that some participants may have been more receptive to the intervention than others. The main outcome, the 7-day smoking point prevalence at the 30-day follow-up, found that one-quarter (11/48, 23%) of all participants had not taken a puff of a cigarette in the prior week and a third (15/48, 31%) had spent at least 24 hours without a cigarette in the prior 2 weeks. However, it is important to note that during testing, we discovered that the game was imbalanced, and there were remaining bugs (discussed more later in the Limitations section), all of which likely prohibited game progression and led to premature dropout. These bugs likely made the in-game rewards earned through submitting smoke reports less valuable and, therefore, less likely to function as a reinforcer to facilitate smoking abstinence. Overall, participants played an average of almost 11 levels over 20 days of gameplay, and their video game addiction scores remained steady. One-third (31.4%, SD 38.2%) of submissions during the abstinence phase indicated tobacco abstinence, and participants averaged 5.4 (SD 9.8) submissions of continuous abstinence.

Of note, the AC and SC groups did not differ on gameplay and smoking cessation. It is possible that the SC group erroneously believed that they would earn in-game rewards through meeting the smoking goal requirement rather than submission of smoking reports alone. It is also possible the in-game rewards (gem prisms and silver multipliers) did not function as reinforcers. Gem prisms and silver multipliers were not required for playing the game, and their utility was most apparent for later, more challenging levels, which many participants did not reach. Participants ranked the silver multiplier among the least valuable game features in the Treatment Acceptability Questionnaire, though the helpfulness of gem prisms was not directly assessed. In addition, we do not have the data on earning and using the gem prisms and silver multipliers to determine whether the participants made contact with these rewards. Participants indicated, on the Treatment Acceptability Questionnaire, that they had a reasonably clear understanding of how the rewards in the game could motivate their smoking cessation, but participants were not asked about how they earned in-game rewards, what those rewards were, and how they could be valuable for their gameplay. Therefore, it is unclear whether the game-based incentives impacted participants' reports of smoking.

The generalized linear models found that the highest level achieved was positively associated with the percentage of abstinent submissions and the longest abstinence streak, which suggests promise for the potential efficacy of the game to support smoking abstinence. However, the total game score was not associated with reported abstinence. Participants unlocked levels through smoking self-report submissions only and not through gameplay, whereas participants increased their score by completing levels in a sequential fashion and replaying levels, which could be done independently of submitting smoking self-reports or meeting smoking abstinence goals. In addition, the total game score had a high range and a large SD, meaning that some participants scored very little, whereas some scored very high. It is possible that the highest-scoring individuals engaged with the game more due to finding it fun or challenging rather than being related to their efforts to achieve smoking abstinence.

It is also important to note that most participants (28/48, 58%) stopped submitting smoking reports by the thinning phase. This decrease in engagement aligns with the lack of continued engagement across app-based smoking cessation interventions more broadly, which continues to be a concern in the field [33]. High dropout likely contributed to the low percentage of smoking reports submitted out of the total possible reports, at around one-third (34.7%, SD 29.1%). Finding strategies to increase and maintain engagement will be critical for future video game–based smoking cessation interventions.

The 30 participants who completed the Treatment Acceptability Questionnaire at week 7 found the game to be relatively helpful, fun, and motivating to help them quit smoking. It should be noted that participants who dropped out of the study before week 7 or who otherwise did not complete the survey may have had different opinions of the treatment, particularly if those opinions contributed to their dropout. Those who did complete the survey (30/48, 63%) highly valued the more active gameplay elements (eg, colored gems and defensive buildings) as opposed to the optional village customization that was part of the metasystem.

Limitations

These results should be considered in the context of the study's limitations. Due to technical difficulties, CO collection could not be performed, and smoking abstinence was determined primarily from smoking self-report. It is possible that participants falsely reported abstinence for social desirability or to gain access to in-game rewards. Although not ideal, self-reporting is common in smoking cessation studies and has aligned closely with biologically verified smoking measures in prior research [25,34,35]. We were only able to collect 20 cotinine samples; however, these demonstrated a high level of agreement with self-reported tobacco use among both abstinent and nonabstinent individuals.

The difficulties with CO collection also hindered the final sample size. The power was originally calculated for a sample size of 114 as opposed to the final sample of 48, and these analyses were based on group membership as an independent variable and the 7-day smoking point prevalence as an outcome. The study was underpowered to detect group differences in the outcome. The 2 groups were also not found to systematically differ in demographics, smoking history, game use, or self-reported abstinence. These developments informed the choice to focus on feasibility and acceptability in these analyses.

In addition, as noted earlier, participants encountered a number of bugs in the gameplay because Inspired was still in an early phase of testing. Study staff worked closely to collect this information from participants and helped troubleshoot difficulties. The high level of dropout seen throughout the study may have been due to participants returning to smoking or due to frustrations with the bugs they encountered in the game. For example, we know that level 9 of the core game was unbalanced, and many participants (9/48, 19%) stopped engaging at this level because it was too difficult to complete. Limited information was collected on exact actions during gameplay, such as the number of attempts on each level, which prevented a more granular analysis. Future examinations should focus on the playability of the game and the reasons for discontinuation. Finally, the duration of follow-up was short (30 days), preventing the examination of the intervention's impact over a longer period.

Strengths

This study had many strengths. Participants were relatively diverse in terms of age, education, employment, and annual income. In addition, the game was designed to deliver rewards that were only valuable in the context of the game, which, if effective, could greatly reduce the costs associated with delivering interventions that typically rely on financial incentives to promote abstinence (ie, contingency management) [36]. Although Inspired was ineffective in promoting abstinence for most participants for several potential reasons already noted, if it could be improved to promote greater levels of abstinence, it could be expanded to reach many people. Scaling up a game-based intervention that could be used by thousands of users over an extended period would add little additional cost, especially when compared to monetary-based contingency management, where each successful quit attempt, as well as longer periods of sustained abstinence, increases costs [37].

A major strength of Inspired was its flexibility. Participants were able use their mobile phone to access the game at any time of day and from any location, maximizing convenience. Using participants' existing mobile phones made this intervention scalable and sustainable, requiring no additional hardware. Although some disparities do continue to exist concerning access to smartphones, especially among lower-income populations who are also more likely to smoke, this gap has been steadily closing, and approximately 76% of people living in households making ≤US $30,000 have access to smartphones [38]. Although participants were required to submit smoke reports twice per day, they were able to incorporate them into their existing schedule and play the game when they wished. The core game was intentionally designed to take approximately 5 minutes to complete (ie, casual game), which is about the duration of a craving arc and approximates how long it might take to smoke a cigarette, serving as a competitor to smoking [39-41]. Furthermore, common barriers to in-person smoking cessation treatment include high cost, lack of access to health care settings, and lack of cessation support from health care providers, with those of low socioeconomic status and those belonging to underrepresented groups being particularly vulnerable to disparities in care [42,43]. A mobile smartphone game such as Inspired could provide a cessation intervention regardless of the health care status or the ability to pay for treatment. The flexibility of a smartphone game–based intervention could allow for low-cost access to cessation services for a wider range of individuals.

Conclusions

Taken together, these results showed promise for Inspired as an acceptable video game–based smoking cessation intervention. Participants were highly motivated to quit but had unsuccessfully tried to do so an average of 5.4 (SD 5.5) times before enrolling in the study. Although engagement and abstinence were variable, a subset of the participants (11/48, 23%) appeared to respond well to this approach having not taken a puff of cigarette in the previous 7 days at the 30-day follow up, which may make the difference in their smoking cessation quit attempts. Inspired used the principles of contingency management, delivering incentives, the value of which only existed within the context of the game and did not require additional monetary resources to implement [44].

Although Inspired provides an exciting alternative to traditional, in-person smoking cessation interventions, much needs to be done to ensure feasibility and scalability. The bugs in the gameplay limited the potential reinforcing value of the in-game rewards. Future examinations of Inspired and other such interventions should work to ensure that the gameplay is easy to learn and enjoyable so that achieving abstinence goals results in immediate, meaningful rewards.

The incentives offered through Inspired did not have value outside of the game, but participants may have been motivated to conceal their smoking anyway. To be in alignment with the science of contingency management, objective measures of smoking abstinence are needed. The iCO portable breath CO detector and NicAlert cotinine test strips that we attempted to use in this study were flawed [24]. Other objective measures of smoking abstinence could be explored, such as Alere rapid oral cotinine tests, which we successfully used in another study, or changes in resting heart rate that can be measured from smartphone cameras or smart watches [45,46]. The limitations of this study can inform future video game–based smoking cessation interventions, helping them fulfill the promise of using contingency management as a scalable and fun way to enact behavior change. Such low-cost and accessible interventions could be the key to achieving a healthier, smoke-free life.

Acknowledgments

The authors would like to thank Connor Burrows for helping with early-stage testing and data management, Caitlyn Upton for helping to collect cotinine data, Dan Scherlis for helping with early-stage game conceptualizing, Nicolas Fortugno for game design and overseeing game development, and John Hewitt for helping with coding the game and database.

Conflicts of Interest

DR is in conflict because he has the potential to benefit from sales of the final Inspired game.

References

1. The health consequences of smoking - 50 year of progress: a report of the surgeon general. National Center for Chronic Disease Prevention and Health Promotion. 2014. URL: https://www.ncbi.nlm.nih.gov/books/NBK179276/ [accessed 2024-04-29]
2. Babb S, Malarcher A, Schauer G, Asman K, Jamal A. Quitting smoking among adults - United States, 2000-2015. MMWR Morb Mortal Wkly Rep 2017 Jan 06;65(52):1457-1464 [FREE Full text] [doi: 10.15585/mmwr.mm6552a1] [Medline: 28056007]
3. Chaney SE, Sheriff S. Evidence-based treatments for smoking cessation. Nurse Pract 2012 Apr 15;37(4):24-31. [doi: 10.1097/01.NPR.0000412892.27557.e8] [Medline: 22395743]
4. Christiansen BA, Reeder KM, TerBeek EG, Fiore MC, Baker TB. Motivating low socioeconomic status smokers to accept evidence-based smoking cessation treatment: a brief intervention for the community agency setting. Nicotine Tob Res 2015 Aug 15;17(8):1002-1011 [FREE Full text] [doi: 10.1093/ntr/ntu345] [Medline: 26180226]
5. Paige SR, Alber JM, Stellefson ML, Krieger JL. Missing the mark for patient engagement: mHealth literacy strategies and behavior change processes in smoking cessation apps. Patient Educ Couns 2018 May;101(5):951-955 [FREE Full text] [doi: 10.1016/j.pec.2017.11.006] [Medline: 29153592]
6. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. J Med Internet Res 2022 Apr 26;24(4):e35120 [FREE Full text] [doi: 10.2196/35120] [Medline: 35471414]
7. Rajani NB, Bustamante L, Weth D, Romo L, Mastellos N, Filippidis FT. Engagement with gamification elements in a smoking cessation app and short-term smoking abstinence: quantitative assessment. JMIR Serious Games 2023 Feb 01;11:e39975 [FREE Full text] [doi: 10.2196/39975] [Medline: 36724003]
8. Rajani NB, Mastellos N, Filippidis FT. Impact of gamification on the self-efficacy and motivation to quit of smokers: observational study of two gamified smoking cessation mobile apps. JMIR Serious Games 2021 Apr 27;9(2):e27290 [FREE Full text] [doi: 10.2196/27290] [Medline: 33904824]
9. Raiff BR, Fortugno N, Scherlis DR, Rapoza D. A mobile game to support smoking cessation: prototype assessment. JMIR Serious Games 2018 Jun 07;6(2):e11 [FREE Full text] [doi: 10.2196/games.9599] [Medline: 29880466]
10. Schlam TR, Baker TB. Playing around with quitting smoking: a randomized pilot trial of mobile games as a craving response strategy. Games Health J 2020 Feb 01;9(1):64-70 [FREE Full text] [doi: 10.1089/g4h.2019.0030] [Medline: 31536384]
11. Upton CR, Nastasi JA, Raiff BR. Identifying video game preferences among adults interested in quitting smoking cigarettes: survey study. JMIR Serious Games 2022 Mar 24;10(1):e30949 [FREE Full text] [doi: 10.2196/30949] [Medline: 35323116]
12. Marin-Gomez FX, Garcia-Moreno Marchán R, Mayos-Fernandez A, Flores-Mateo G, Granado-Font E, Barrera Uriarte ML, et al. Exploring efficacy of a serious game (Tobbstop) for smoking cessation during pregnancy: randomized controlled trial. JMIR Serious Games 2019 Mar 27;7(1):e12835 [FREE Full text] [doi: 10.2196/12835] [Medline: 30916655]

13. Bindoff I, de Salas K, Peterson G, Ling T, Lewis I, Wells L, et al. Quittr: the design of a video game to support smoking cessation. JMIR Serious Games 2016 Dec 01;4(2):e19 [FREE Full text] [doi: 10.2196/games.6258] [Medline: 27908844]
14. Derksen ME, van Strijp S, Kunst AE, Daams JG, Jaspers MW, Fransen MP. Serious games for smoking prevention and cessation: a systematic review of game elements and game effects. J Am Med Inform Assoc 2020 May 01;27(5):818-833 [FREE Full text] [doi: 10.1093/jamia/ocaa013] [Medline: 32330255]
15. Krebs P, Burkhalter JE, Snow B, Fiske J, Ostroff JS. Development and alpha testing of QuitIT: an interactive video game to enhance skills for coping with smoking urges. JMIR Res Protoc 2013 Sep 11;2(2):e35 [FREE Full text] [doi: 10.2196/resprot.2416] [Medline: 24025236]
16. White JS, Salem MK, Toussaert S, Westmaas JL, Raiff BR, Crane D, et al. Developing a game (Inner Dragon) within a leading smartphone app for smoking cessation: design and feasibility evaluation study. JMIR Serious Games 2023 Aug 11;11:e46602 [FREE Full text] [doi: 10.2196/46602] [Medline: 37566442]
17. Edwards EA, Caton H, Lumsden J, Rivas C, Steed L, Pirunsarn Y, et al. Creating a theoretically grounded, gamified health app: lessons from developing the cigbreak smoking cessation mobile phone game. JMIR Serious Games 2018 Nov 29;6(4):e10252 [FREE Full text] [doi: 10.2196/10252] [Medline: 30497994]
18. Raiff BR, Jarvis BP, Rapoza D. Prevalence of video game use, cigarette smoking, and acceptability of a video game-based smoking cessation intervention among online adults. Nicotine Tob Res 2012 Dec;14(12):1453-1457 [FREE Full text] [doi: 10.1093/ntr/nts079] [Medline: 22422929]
19. Essential facts about the U.S. game industry. Entertainment Software Association. 2023. URL: https://www.theesa.com/wp-content/uploads/2024/02/2023-EF-FINAL.pdf [accessed 2024-04-29]
20. Higgins ST, Silverman K, Heil SH. Contingency Management in Substance Abuse Treatment. New York, NY: The Guilford Press; 2007.
21. Ledgerwood D. Contingency management for smoking cessation: where do we go from here? Curr Drug Abuse Rev 2008 Nov 01;1(3):340-349. [doi: 10.2174/1874473710801030340] [Medline: 19630730]
22. Kirby KC, Benishek LA, Dugosh KL, Kerwin ME. Substance abuse treatment providers' beliefs and objections regarding contingency management: implications for dissemination. Drug Alcohol Depend 2006 Oct 15;85(1):19-27. [doi: 10.1016/j.drugalcdep.2006.03.010] [Medline: 16650657]
23. Skinner BF. The Behavior of Organisms: An Experimental. New York, NY: Appleton-Century; 1938.
24. iCOquit® pilot study utilizing personal Bluetooth® co devices for smoking cessation during pregnancy. Bedfont. 2023. URL: https://www.bedfont.com/icoquit-pilot-study-utilizing-personal-bluetooth-co-devices-for-smoking-cessation-during-pregnancy/ [accessed 2023-04-29]
25. Bricker JB, Watson NL, Mull KE, Sullivan BM, Heffner JL. Efficacy of smartphone applications for smoking cessation: a randomized clinical trial. JAMA Intern Med 2020 Nov 01;180(11):1472-1480 [FREE Full text] [doi: 10.1001/jamainternmed.2020.4055] [Medline: 32955554]
26. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström test for nicotine dependence: a revision of the Fagerström tolerance questionnaire. Br J Addict 1991 Sep 13;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
27. Browning KO, Sutton GM, Nist AN, Shahan TA. The effects of large, small, and thinning magnitudes of alternative reinforcement on resurgence. Behav Processes 2022 Feb;195:104586 [FREE Full text] [doi: 10.1016/j.beproc.2022.104586] [Medline: 35065243]
28. Dallery J, Raiff BR, Grabinski MJ. Internet-based contingency management to promote smoking cessation: a randomized controlled study. J Appl Behav Anal 2013 Dec;46(4):750-764 [FREE Full text] [doi: 10.1002/jaba.89] [Medline: 24114862]
29. Dallery J, Raiff BR, Kim SJ, Marsch LA, Stitzer M, Grabinski MJ. Nationwide access to an internet-based contingency management intervention to promote smoking cessation: a randomized controlled trial. Addiction 2017 May;112(5):875-883 [FREE Full text] [doi: 10.1111/add.13715] [Medline: 27923264]
30. Lemmens JS, Valkenburg PM, Peter J. Development and validation of a game addiction scale for adolescents. Media Psychol 2009 Mar 05;12(1):77-95. [doi: 10.1080/15213260802669458]
31. NicAlert. Nymox Corporation. URL: https://nymox.com [accessed 2024-04-29]
32. Defining rural population. Health Resources and Services Administration. URL: https://www.hrsa.gov/rural-health/about-us/what-is-rural [accessed 2024-04-29]
33. Regmi K, Kassim N, Ahmad N, Tuah N. Effectiveness of mobile apps for smoking cessation: a review. Tob Prev Cessat 2017 Apr 12;3(April):12 [FREE Full text] [doi: 10.18332/tpc/70088] [Medline: 32432186]
34. Crutzen R, Göritz AS. Social desirability and self-reported health risk behaviors in web-based research: three longitudinal studies. BMC Public Health 2010 Nov 23;10(1):720 [FREE Full text] [doi: 10.1186/1471-2458-10-720] [Medline: 21092267]
35. Yeager DS, Krosnick JA. The validity of self-reported nicotine produce use in the 2001-2008 National health and Nutrition Examination Survey. Med Care 2010;48(12):1132. [doi: 10.1097/mlr.0b013e3181ef9948]
36. Coughlin LN, Salino S, Jennings C, Lacek M, Townsend W, Koffarnus MN, et al. A systematic review of remotely delivered contingency management treatment for substance use. J Subst Use Addict Treat 2023 Apr;147:208977. [doi: 10.1016/j.josat.2023.208977] [Medline: 36804352]

37. González-Roz A, Weidberg S, García-Pérez Á, Martínez-Loredo V, Secades-Villa R. One-year efficacy and incremental cost-effectiveness of contingency management for cigarette smokers with depression. Nicotine Tob Res 2021 Jan 22;23(2):320-326. [doi: 10.1093/ntr/ntaa146] [Medline: 32772097]
38. Perrin A. Mobile technology and home broadband 2021. Pew Research Center. URL: https://www.pewresearch.org/internet/2021/06/03/mobile-technology-and-home-broadband-2021/ [accessed 2024-04-29]
39. Hajek P, Pittaccio K, Pesola F, Myers Smith K, Phillips-Waller A, Przulj D. Nicotine delivery and users' reactions to Juul compared with cigarettes and other e-cigarette products. Addiction 2020 Jun 29;115(6):1141-1148 [FREE Full text] [doi: 10.1111/add.14936] [Medline: 31994254]
40. Hoffmann DH, Hoffmann I. The changing cigarette, 1950-1995. J Toxicol Environ Health 1997 Mar;50(4):307-364. [doi: 10.1080/009841097160393] [Medline: 9120872]
41. Shiffman S, Dunbar M, Kirchner T, Li X, Tindle H, Anderson S, et al. Smoker reactivity to cues: effects on craving and on smoking behavior. J Abnorm Psychol 2013 Feb;122(1):264-280 [FREE Full text] [doi: 10.1037/a0028339] [Medline: 22708884]
42. van Rossem C, Spigt MG, Kleijsen JR, Hendricx M, van Schayck CP, Kotz D. Smoking cessation in primary care: exploration of barriers and solutions in current daily practice from the perspective of smokers and healthcare professionals. Eur J Gen Pract 2015 Jun 04;21(2):111-117. [doi: 10.3109/13814788.2014.990881] [Medline: 25649048]
43. Twyman L, Bonevski B, Paul C, Bryant J. Perceived barriers to smoking cessation in selected vulnerable groups: a systematic review of the qualitative and quantitative literature. BMJ Open 2014 Dec 22;4(12):e006414 [FREE Full text] [doi: 10.1136/bmjopen-2014-006414] [Medline: 25534212]
44. Secades-Villa R, Aonso-Diego G, García-Pérez Á, González-Roz A. Effectiveness of contingency management for smoking cessation in substance users: a systematic review and meta-analysis. J Consult Clin Psychol 2020 Oct;88(10):951-964. [doi: 10.1037/ccp0000611] [Medline: 33048571]
45. Raiff BR, Newman ST, Upton CR, Burrows CA. The feasibility, acceptability, and initial efficacy of a remotely delivered, financial-incentive intervention to initiate vaping abstinence in young adults. Exp Clin Psychopharmacol 2022 Oct;30(5):632-641. [doi: 10.1037/pha0000468] [Medline: 34251841]
46. Herbec A, Parker E, Ubhi HK, Raupach T, West R. Decrease in resting heart rate measured using smartphone apps to verify abstinence from smoking: an exploratory study. Nicotine Tob Res 2020 Jul 16;22(8):1424-1427 [FREE Full text] [doi: 10.1093/ntr/ntaa021] [Medline: 31971595]

Abbreviations

AC: abstinence contingent
CO: carbon monoxide
CONSORT: Consolidated Standards of Reporting Trials
SC: submission contingent

Edited by R Baranyi; submitted 18.11.23; peer-reviewed by J White, D Sharma, U Flato, TAR Sure; comments to author 30.01.24; revised version received 15.03.24; accepted 16.04.24; published 21.08.24.

Please cite as:
Goodwin S, Nastasi JA, Newman ST, Rapoza D, Raiff BR
Feasibility and Acceptability of a Mobile Game to Support Smoking Cessation: Repeated Measures Study
JMIR Serious Games 2024;12:e54684
URL: https://games.jmir.org/2024/1/e54684
doi:10.2196/54684
PMID:

Impact of Fruit and Vegetable Enzyme Supplementation on Aerobic Performance and Lactate Response in Older Adults Following High-Intensity Interval Exercise Through Exergaming: Randomized Experimental Matched-Pair Study

Shu-Cheng Lin[1], PhD; Chien-Yen Wang[2], MA; Tien-Hung Hou[3], PhD; Hong-Ching Chen[1], BA; Chia-Chi Wang[4], PhD

1

2

3

4

Corresponding Author:
Chia-Chi Wang, PhD

Abstract

Background: Exercise offers substantial health benefits but can induce oxidative stress and inflammation, especially in high-intensity formats such as high-intensity interval exercise (HIIE). Exergaming has become an effective, enjoyable fitness tool for all ages, particularly older adults. Enzyme supplements may enhance exercise performance by improving lactate metabolism and reducing oxidative stress.

Objective: This study investigates the efficacy of fruit and vegetable enzyme supplementation in modulating fatigue and enhancing aerobic capacity in older adults following HIIE through exergaming.

Methods: The study recruited 16 older adult female participants and allocated them into 2 distinct groups (enzyme and placebo) based on their pretest lactate levels. This division used pairwise grouping to guarantee comparability between the groups, ensuring the integrity of the results. They engaged in HIIE using Nintendo Switch Ring Fit Adventure, performing 8 sets of 20 seconds of maximum effort exercise interspersed with 30 seconds of rest, totaling 370 seconds of exercise. Key metrics assessed included blood lactate levels, heart rate, rating of perceived exertion, and training impulse. Participants in the enzyme group were administered a fruit and vegetable enzyme supplement at a dosage of 30 mL twice daily over a period of 14 days.

Results: The enzyme group showed significantly lower blood lactate levels compared to the placebo group, notably after the fourth (mean 4.29, SD 0.67 vs mean 6.34, SD 1.17 mmol/L; *P*=.001) and eighth (mean 5.84, SD 0.63 vs mean 8.20, SD 1.15 mmol/L; *P*<.001) exercise sessions. This trend continued at 5 minutes (mean 6.85, SD 0.82 vs mean 8.60, SD 1.13 mmol/L; *P*=.003) and 10 minutes (mean 5.91, SD 1.16 vs mean 8.21, SD 1.27 mmol/L; *P*=.002) after exercise. Although both groups exceeded 85% of their estimated maximum heart rate during the exercise, enzyme supplementation did not markedly affect the perceived intensity or effort.

Conclusions: The study indicates that fruit and vegetable enzyme supplementation can significantly reduce blood lactate levels in older adults following HIIE through exergaming. This suggests a potential role for these enzymes in modulating lactate production or clearance during and after high-intensity exercise. These findings have implications for developing targeted interventions to enhance exercise tolerance and recovery in older adults.

Trial Registration: ClinicalTrials.gov NCT06466408; https://clinicaltrials.gov/study/NCT06466408

(JMIR Serious Games 2024;12:e52231) doi:10.2196/52231

KEYWORDS

Ring Fit Adventure; training load; older adult training; training impulse; food supplement; older adults; exergames; exergame; Taiwan; female; fruits; vegetables; blood lactate; exercise; feasibility; aerobic; enzymes; enzyme; female older adults; fitness; food intake; diet; exergaming; enzyme supplements; older adults training; female older adult; older adult

Introduction

Exercise represents a paradoxical element in health management, offering substantial benefits yet posing potential risks if not properly moderated [1,2]. High-intensity exercise, although efficacious in improving various health parameters, can lead to oxidative stress, muscle damage, and inflammation [3,4]. The oxidative stress primarily arises from increased reactive oxygen species production during intensive physical activities [5]. Moreover, exercise-induced fatigue serves as a protective mechanism against overexertion and consequent injuries [6,7]. In contemporary fitness regimes, high-intensity interval exercise (HIIE), particularly the Tabata training method, has gained prominence for its effectiveness in enhancing aerobic power, fat oxidation, and muscular endurance [8-10]. These attributes are especially crucial for the older adult population, a demographic that significantly benefits from regular physical activity [11-13].

Exergaming, an innovative blend of physical exercise and interactive gaming, has emerged as a transformative approach to fitness, especially in engaging diverse age groups in regular physical activity. Its efficacy in enhancing key fitness parameters such as aerobic capacity, agility, and coordination, coupled with its ability to make exercise more enjoyable, has been well documented [14-16]. This fusion of technology and exercise not only caters to the digital age but also opens avenues for personalized fitness experiences, which are adaptable to various demographic needs [17,18]. Although exergaming has been effective across a range of ages, its application in older adult populations presents unique opportunities and challenges. As the older adult population seeks safe, engaging, and effective exercise methods, exergaming could offer a solution that aligns with these requirements. However, integrating HIIE concepts into exergaming for older adults remains a relatively uncharted territory. HIIE, known for its efficiency in improving cardiovascular health and metabolic function, could significantly benefit older adults, particularly in terms of enhancing functional capacity and overall quality of life [12,19].

The potential of HIIE within exergaming for older adults hinges on the balance between intensity and safety. Although HIIE is beneficial, it is crucial to adapt its intensity to suit the physiological capabilities and limitations of older individuals. Research indicates that tailored HIIE programs can be both feasible and beneficial for older adults, leading to improvements in cardiovascular health, muscle strength, and metabolic function [20,21]. Integrating these concepts into exergaming could further enhance adherence and enjoyment, which are crucial factors in maintaining regular exercise habits in this demographic. Furthermore, the interactive and immersive nature of exergaming can address common barriers to exercise among older adults, such as the lack of motivation or fear of injury. By providing a safe, controlled environment for engaging in HIIE, exergaming can potentially transform the perception and experience of high-intensity workouts for older adults. This is particularly pertinent given the increasing need for innovative exercise interventions that cater to the aging global population [11].

Nutritional supplementation, especially with natural fruit and vegetable enzymes, presents a promising avenue in augmenting exercise performance through their antioxidant, anti-inflammatory, and metabolic benefits [21-28]. Such supplementation could potentially optimize lactate metabolism and enhance muscle function during exercise. Recent advancements in nutritional science have highlighted the substantial role of natural fruit and vegetable enzymes in enhancing exercise performance. These enzymes are increasingly recognized for their multifaceted health benefits, including their antioxidant, anti-inflammatory, and metabolism-enhancing properties [21,22]. Notably, their potential impact on exercise physiology, particularly in the context of high-intensity workouts, offers a new perspective on improving athletic performance and recovery.

One of the critical areas where these enzymes show promise is in the modulation of lactate metabolism. Lactate, often produced in higher quantities during intense physical activity, can lead to fatigue and decreased muscle efficiency. The traditional view of lactate as merely a byproduct of anaerobic metabolism has evolved, with current research acknowledging its role as a valuable energy source during prolonged exercise [23]. This shift in understanding opens up new avenues for using enzyme supplementation to optimize lactate use. Enzymes such as bromelain and papain, found in pineapples and papayas, respectively, have been studied for their potential in improving lactate metabolism. These enzymes are known to facilitate faster clearance of lactate from the bloodstream, thereby enhancing recovery and reducing fatigue [26,28]. Furthermore, the antioxidant properties of these enzymes play a crucial role in combating oxidative stress, which is often elevated during intense exercise regimens [24,25]. This reduction in oxidative stress is not only beneficial for immediate recovery but also contributes to long-term muscle health and function. Moreover, the anti-inflammatory actions of these natural enzymes can mitigate the inflammatory response often triggered by high-intensity exercise [27]. By reducing inflammation, these enzymes may enhance muscle recovery and function, thus allowing for more efficient and prolonged exercise performance. This aspect is particularly relevant in training regimens where recovery is as crucial as the exercise itself.

The primary aim of this feasibility study is to examine the effects of fruit and vegetable enzyme supplementation on aerobic capacity and blood lactate response in older adults engaged in HIIE through an exergaming framework. This study is dual faceted, focusing on (1) the physiological responses and feasibility of an exergaming HIIE regimen tailored for older adults and (2) the impact of enzyme supplementation on enhancing these exercise outcomes.

Methods

Sample Size

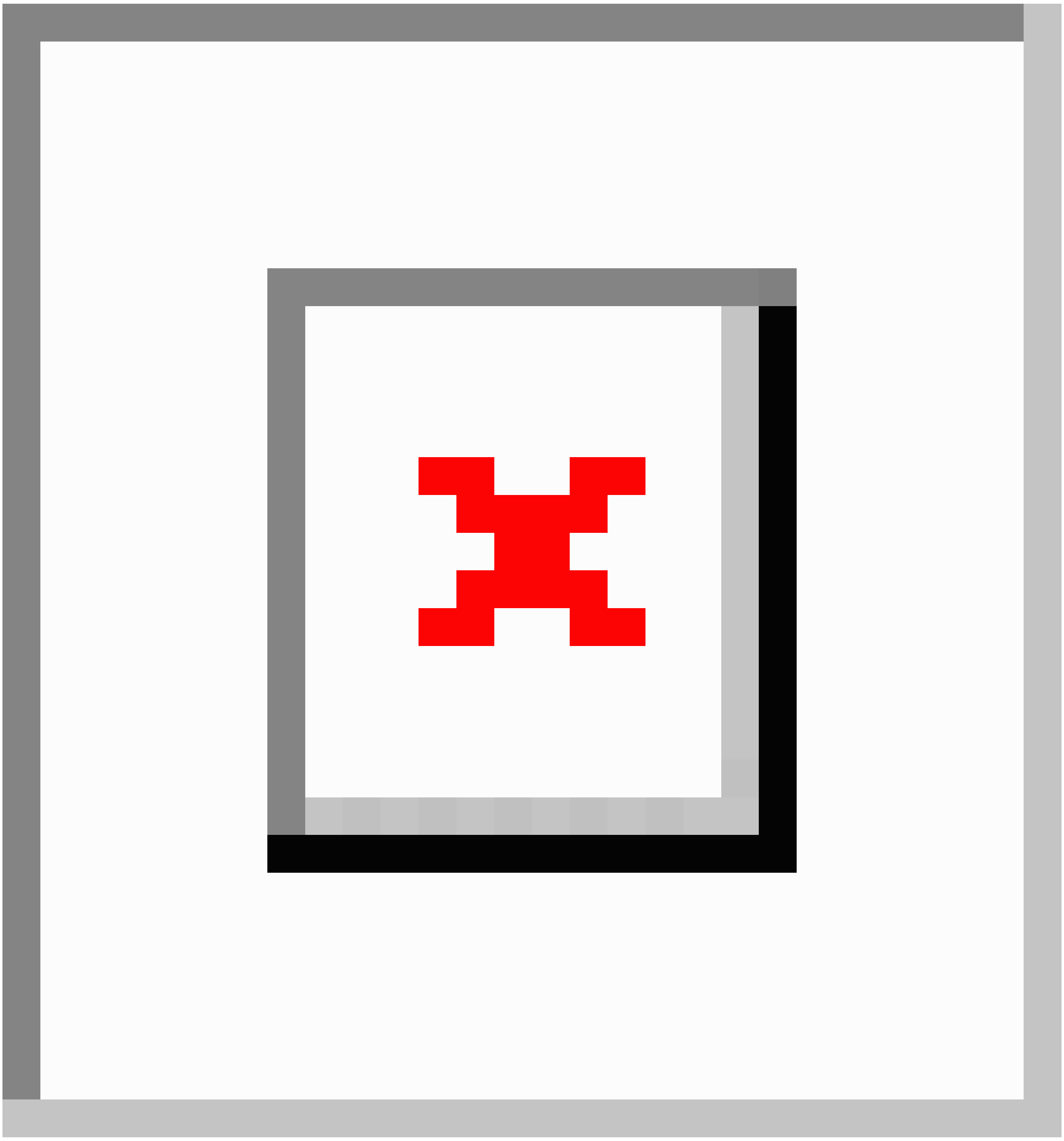
The sample size computation was based on the study by Flanagan and Jakeman [29]. Based on a statistical power analysis, a total sample size of 16 participants (8 per group) was needed to achieve a statistical power of 0.8 to detect a large effect size (ES) for supplement-time interaction at an α level of .05 [30].

Participants and Experimental Design

After recruiting a total of 30 healthy older adult participants, the study proceeded with screenings and initial explanations. Subsequently, 12 individuals were excluded as they did not meet the inclusion criteria, and 2 declined to participate. Ultimately, 16 female older adult participants were enrolled in the study. These participants were then divided into 2 distinct groups (enzyme and placebo) based on their pretest lactate levels. Pairwise grouping was used to ensure comparability between the groups, thereby preserving the integrity of the results. All participants reported a regular exercise habit (3 times per week within the past year). They also completed the Physical Activity Readiness Questionnaire and confirmed no history of upper-limb skeletal muscle injury or major injury. Participants were instructed to avoid strenuous activities and the intake of caffeine or muscle-enhancing supplements for 24 hours prior to the experiment. Before the study commenced, all participants provided personal information, completed health questionnaires, disclosed personal medical history, and signed informed consent forms.

The 16 participants underwent the exergaming HIIE test as an initial assessment (pretest). Participants engaged in a 5-minute warm-up on a stationary bike, followed by HIIE using Nintendo Switch Ring Fit Adventure. The training method was adapted from previous research [8,9] and consisted of 8 sets of 20 seconds of maximum effort exercise with 30 seconds of complete rest between each set, resulting in a total exercise time of 370 seconds. The HIIE design incorporated training modes targeting the deltoid, pectoralis major, latissimus dorsi, and quadricep muscles in Nintendo Switch Ring Fit Adventure. Blood lactate levels, heart rate (HR), and ratings of perceived exertion (RPE) were recorded before, during, and after exercise, and training load was quantified using training impulse (TRIMP). Participants were matched and divided into 2 groups, the enzyme group and the placebo group, based on their blood lactate levels during HIIE. Each group comprised 8 individuals. Supplementation with vegetable and fruit enzymes or maltodextrin commenced 3 days after the pretest and lasted for a total of 14 days. On the 14th day, following the completion of supplementation, the participants underwent the exergaming HIIE test as a posttest (Figure 1). This study was not preregistered as it was considered a feasibility study.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) and experimental procedure diagram. HIIE: high-intensity interval exercise; HR: heart rate; RPE: rating of perceived exertion.



Ethical Considerations

The human research ethics committee of the local university approved this study, which was also approved by the human research ethics committee of the National Cheng Kung University, Taiwan (approval NCKU HREC-E-112-419-2). Users volunteered for this study and agreed to participate by signing an informed consent form. To protect the personal data of participants, all participant information has been anonymized and assigned identification numbers. Participation was voluntary following recruitment, and participants were given a small gift at the conclusion as a token of appreciation.

Supplementation Protocol

After the pretest, the enzyme group consumed 30 mL of vegetable and fruit enzymes (the contents included needle-leaf cherries, cherries, apples, cranberries, blackberries, black currants, blueberries, beets, broccoli, cabbage, carrots, Concord grapes, cranberries, elderberries, kale, oranges, peaches, papayas, parsley, pineapples, raspberries, red currants, spinach, and tomatoes, etc; Enzyme Village) mixed with 150 mL of water twice a day (at breakfast and dinner) for 14 consecutive days. The placebo group followed the same protocol but consumed malt syrup (Amazon) instead until the end of the study. Participants returned to the laboratory each morning to receive the daily supplement, which was administered on site.

Following supplementation, participants reported their dinner intake to the researchers, ensuring compliance with the prescribed supplementation regimen.

Exergaming HIIE Test: Combination of Exergaming and HIIE

Participants in this experiment engaged in HIIE using the Nintendo Switch Ring-Con within a laboratory environment. All participants completed pre- and posttest assessments on the same day. The exergame used in this study was Nintendo Switch Fitness Adventure, which ingeniously blends exercise with an adventure narrative to deliver both physical workouts and gaming enjoyment concurrently. This game is noted for its intuitive, user-friendly interface that accommodates players of all ages. It incorporates a specialized fitness ring—a smart accessory that connects to the Nintendo Switch console. The sensor system used 2 Nintendo controllers: 1 mounted on the exercise ring and the other secured to the participant's thigh to enhance gameplay interaction. Through the Ring-Con, participants engaged in diverse physical activities such as weightlifting, yoga, and aerobic exercises. The fitness ring sensor accurately captures and integrates players' movements into the game. The gameplay involves unlocking levels and engaging in fitness challenges that are achieved through actual physical activities. It offers a wide range of exercise routines targeting various muscle groups and provides engaging gaming challenges. The exercise protocol included 8 sets of 20-second, high-effort exercises, interspersed with 30-second rest intervals, totaling 370 seconds of active exercise time. Specifically, the fitness game mode used was the Adventure Mode in Ring Fit Adventure, comprising exercises targeting the pectoralis major, latissimus dorsi, deltoid, and quadricep muscles (Figures 2 and 3).

Figure 2. Experimental flowchart. * indicates lactate test. # indicates tests for heart rate and rating of perceived exertion. Ex-: bouts of HIIE; HIIE: high-intensity interval exercise; post–5 min: after 5 minutes of HIIE; post–10 min: after 10 minutes of HIIE.

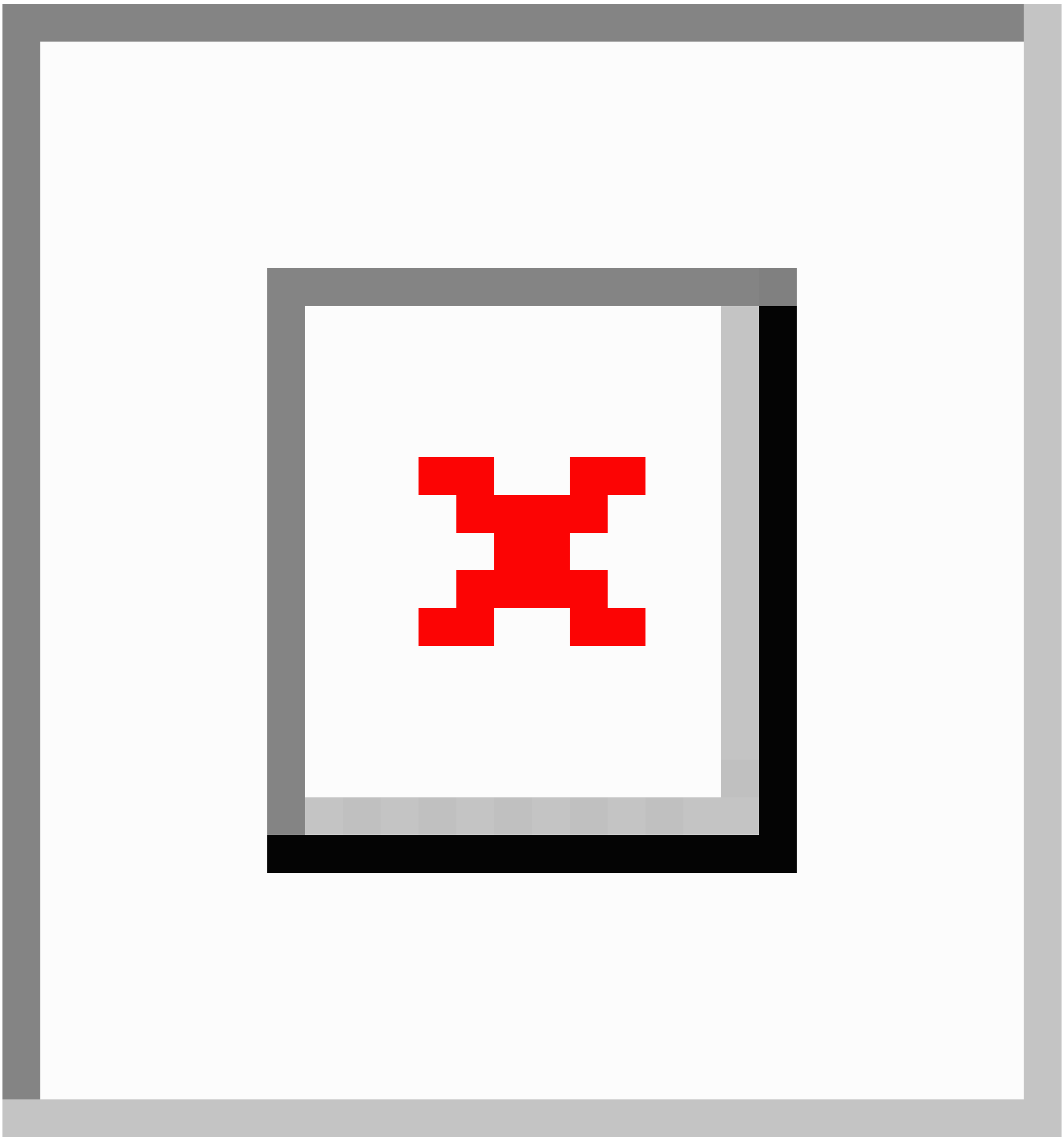
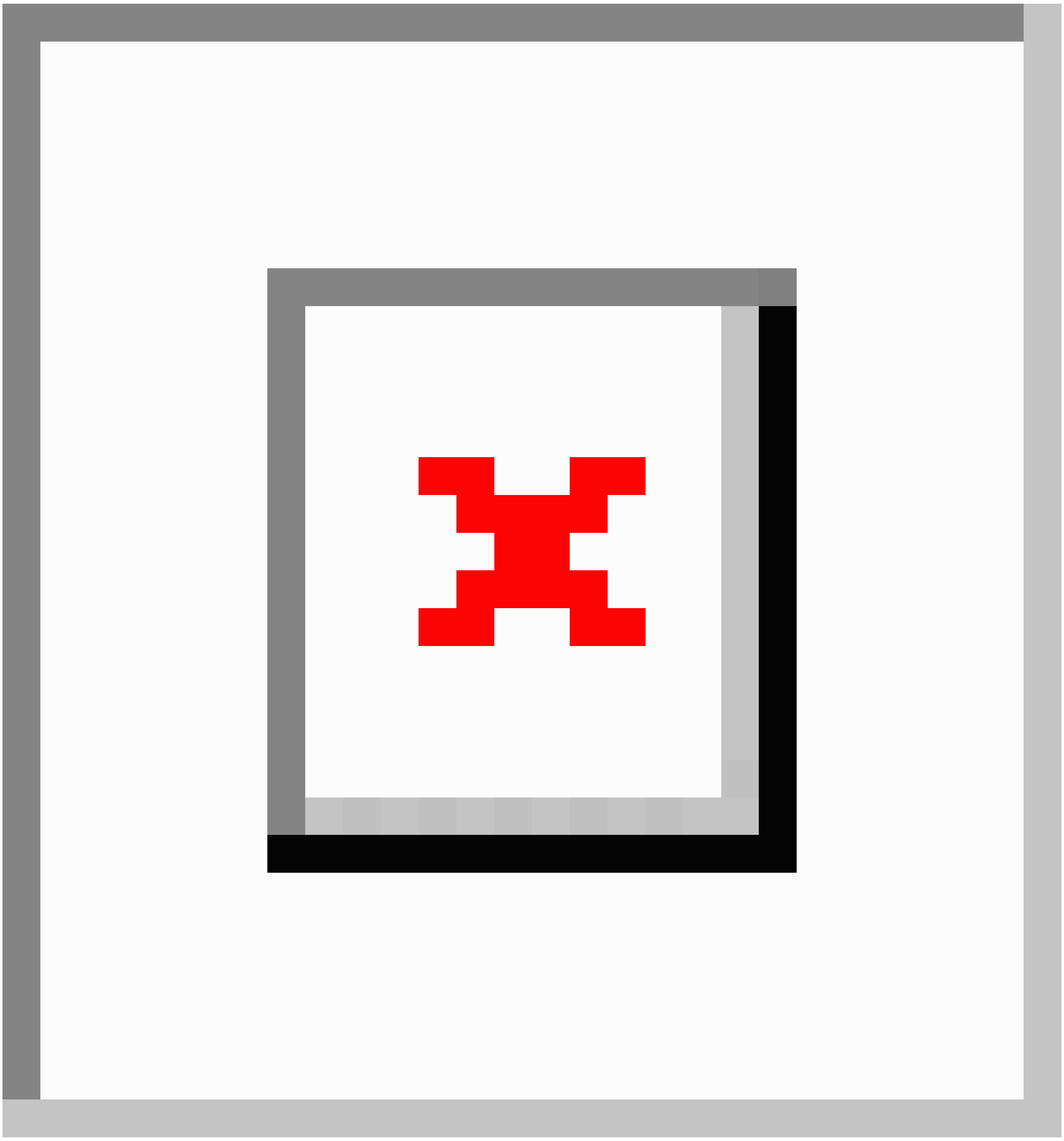


Figure 3. Exercise training model: (A) pectoralis major, (B) deltoid, (C) latissimus dorsi, and (D) quadricep muscles. The images represent the 4 exercises used in this study. (A) shows pressing the fitness ring inward; (B) and (C) depict pulling the fitness ring outward; and (D) illustrates mounting the sensor device on the thigh, which should be raised to approximately 90°.



Blood Lactate Test

Blood lactate was measured at 5 time points: before exercise, after the fourth and eighth bouts of exercise, and at 5 and 10 minutes after exercise. Blood lactate was analyzed using a Biosen Cline blood analysis system (EKF-diagnostic). Capillary blood samples of 10 μL were collected, added to red blood cell lysis reagent, and stored at low temperature until analyzed. Prior to analysis, instrument standardization and test calibration were performed, and the coefficient of variation was determined to be ≤1.5%. The detection range for blood lactate was 0.5-40 mM [31].

Exercise Load (TRIMP)

Overview

In this study, the exercise load was represented by the TRIMP [32], which was calculated as the product of exercise intensity and duration. To accommodate the convenience of the experiment, 2 different TRIMP calculation methods were used, including % maximum HR (HRmax; objective) and RPE (subjective). At the end of each exercise bout (8 bouts in total) and during the recovery period before the next bout (7 bouts in total), participants were asked to report their RPE, and their HR was recorded. This process was repeated 8 times.

% HRmax Calculation Method

During the entire HIIE, the participant's HR was recorded every 5 seconds using a HR monitor (iHeart Polar) to calculate % HRmax. The block TRIMP method developed by Edwards [33] was used, which divides the exercise intensity into 5 blocks with corresponding weighting factors (Table 1). The weighted score of each block was multiplied by the exercise time (min) and then summed to obtain the exercise load (arbitrary unit [AU]). The calculation formula was as follows: *Exercise load = (Z1 exercise time × 1) + (Z2 exercise time × 2) + (Z3 exercise time × 3) + (Z4 exercise time × 4) + (Z5 exercise time × 5).*

Table . The Edwards [33] block training intensity calculation method.

Zone	Intensity (% HRmax[a]), range	Weighted score
Z1	50-60	1
Z2	60-70	2
Z3	70-80	3
Z4	80-90	4
Z5	90-100	5

[a]HRmax: maximum heart rate.

RPE Calculation Method

The TRIMP calculation method of Foster et al [34,35] was used to calculate the exercise load, by multiplying the RPE value of each exercise segment by the exercise time and summing them up. The RPE scale used in this method was the CR-10 version modified by Foster et al [35] based on Borg et al [36] (Table 2). The calculation formula was as follows: *Exercise load (AU) = Borg CR-10 RPE score × exercise time (min).*

Table . Borg CR-10 rating of perceived exertion (RPE).

Borg CR-10 RPE score	Level of exertion
0	Rest
1	Very, very easy
2	Easy
3	Moderate
4	Somewhat hard
5	Hard
6	Hard
7	Very hard
8	Very hard
9	Very hard
10	Maximal

Statistical Analysis

All the data were analyzed by SPSS for Windows 20.0 (IBM Corp). Data are expressed as mean (SD) and 95% CI. A mixed design 2-way ANOVA (group×time) was used to compare the variables of lactate response, HR, and TRIMP between 2 groups before and after the 14 days of supplementation. Graphs were generated using GraphPad Prism 8.0 (GraphPad Software). Cohen conventions for ES (Cohen *d*) were calculated by the G*Power 3.1 software program (Heinrich-Heine-Universität), where the ESs of 0.2, 0.5, and 0.8 are considered small, medium, and large, respectively. Statistical significance was set as $P<.05$.

Results

Overview

Table 3 outlines the baseline characteristics of participants in the study, divided into the enzyme and placebo groups. The average age of participants was slightly higher in the placebo group (66.50, SD 1.31 y) than the enzyme group (65.75, SD 0.88 y). Heights were similar across both groups, with the enzyme group averaging 160.50 (SD 2.67) cm and the placebo group averaging 160.13 (SD 2.75) cm. The enzyme group members were slightly heavier (mean 56.75, SD 4.27 kg) than those in the placebo group (mean 53.5, SD 3.42 kg), which was also reflected in a higher average BMI (22.02, SD 1.41 kg/m² in the enzyme group vs 20.89, SD 1.71 kg/m² in the placebo group). Regarding exercise habits, both groups engaged in regular physical activity, with the enzyme group exercising on average 3.75 (SD 0.71) days per week and the placebo group exercising slightly more at 4.00 (SD 0.76) days per week. The daily exercise duration was comparable between groups, with the enzyme group averaging 76.25 (SD 41.04) minutes and the placebo group averaging 78.75 (SD 31.82) minutes.

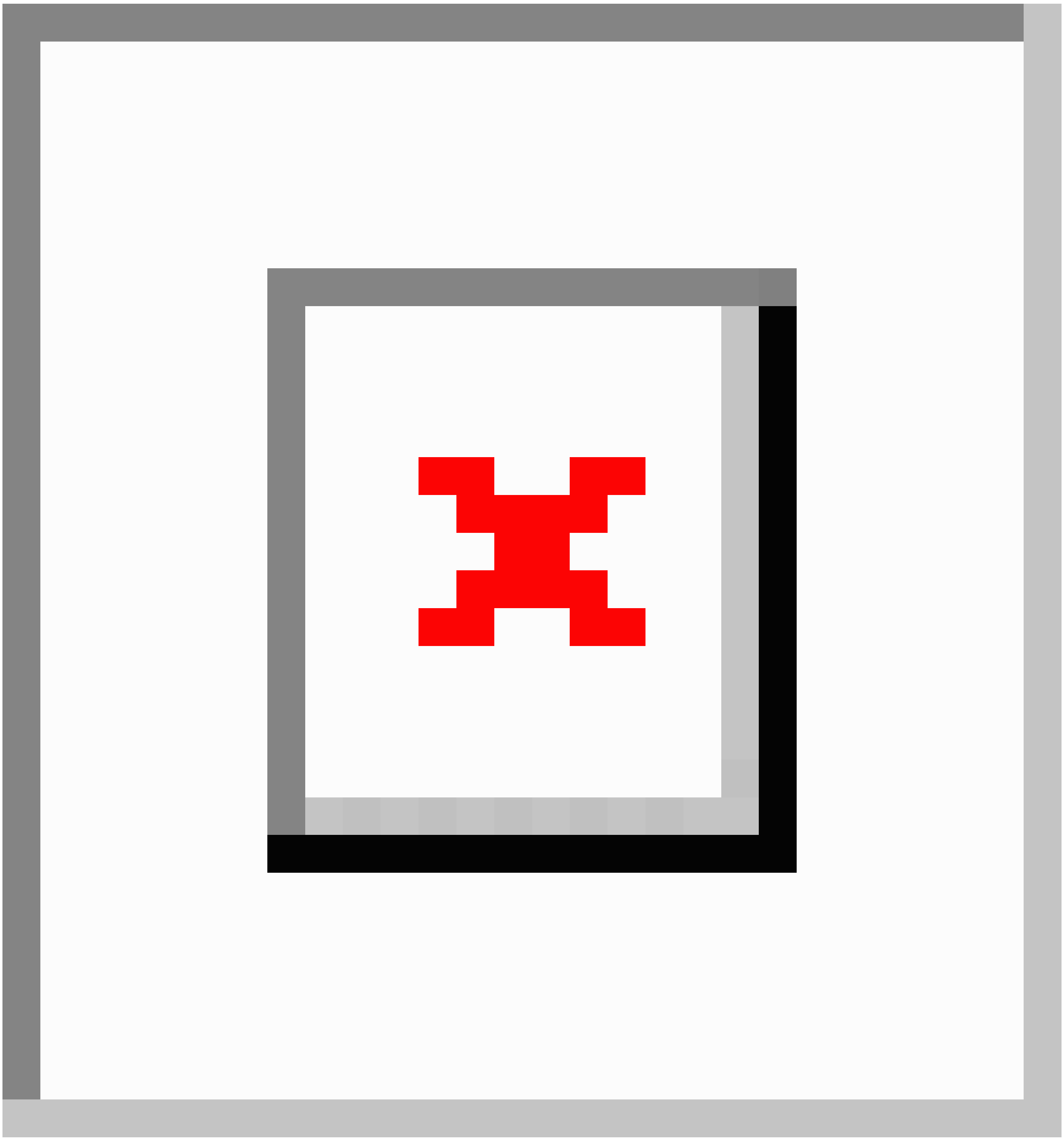
Table . Participants' baseline characteristics.

Characteristics	Enzyme group, mean (SD)	Placebo group, mean (SD)
Age (y)	65.75 (0.88)	66.50 (1.31)
Height (cm)	160.50 (2.67)	160.13 (2.75)
Weight (kg)	56.75 (4.27)	53.5 (3.42)
BMI (kg/m^2)	22.02 (1.41)	20.89 (1.71)
Frequency of regular exercise habits (d/wk within the past year)	3.75 (0.71)	4.00 (0.76)
Daily exercise duration (min)	76.25 (41.04)	78.75 (31.82)

Enzyme Supplementation's Impact on Lactate Response in Exergaming Combined With HIIE

The results demonstrated that blood lactate levels surpassed 4 mmol/L after the fourth exercise bout, indicating the presence of high-intensity exercise. Additionally, the study examined the effects of 14 days of enzyme or placebo supplementation on blood lactate levels ($F_{1,14}$=6.99; *P*=.001). The enzyme group exhibited significantly lower blood lactate levels than the placebo group after the fourth (mean 4.29, SD 0.67; 95% CI 3.56-5.01 vs mean 6.34, SD 1.17; 95% CI 5.61-7.06 mmol/L; ES=−2.14; *P*=.001) and eighth (mean 5.84, SD 0.63; 95% CI 5.14-6.54 vs mean 8.20, SD 1.15; 95% CI 7.50-8.90 mmol/L; ES=−2.56; *P*=.001) exercise bouts, as well as at 5 minutes (mean 6.85, SD 0.82; 95% CI 6.10-7.60 vs mean 8.60, SD 1.13; 95% CI 7.85-9.35 mmol/L; ES=−1.78; *P*=.003) and 10 minutes (mean 5.91, SD 1.16; 95% CI 4.99-6.84 vs mean 8.21, SD 1.27; 95% CI 7.29-9.14 mmol/L; ES=−1.89; *P*=.002) after exercise (Figure 4). These findings suggest that the combination of HIIE and exergaming can lead to high-intensity exercise, and enzyme supplementation can contribute to a reduction in lactate levels.

Figure 4. Blood lactate response (A) before and (B) after 14 days of enzyme or placebo supplementation. Data are presented as mean (SD). * indicates a significant difference (*P*<.05) from the pre-exercise value within the group. # indicates a significant difference (*P*<.05) between the groups. Ex-4: fourth bout of HIIE; Ex-8: eighth bout of HIIE; HIIE: high-intensity interval exercise; post–5 min: after 5 minutes of HIIE; post–10 min: after 10 minutes of HIIE.

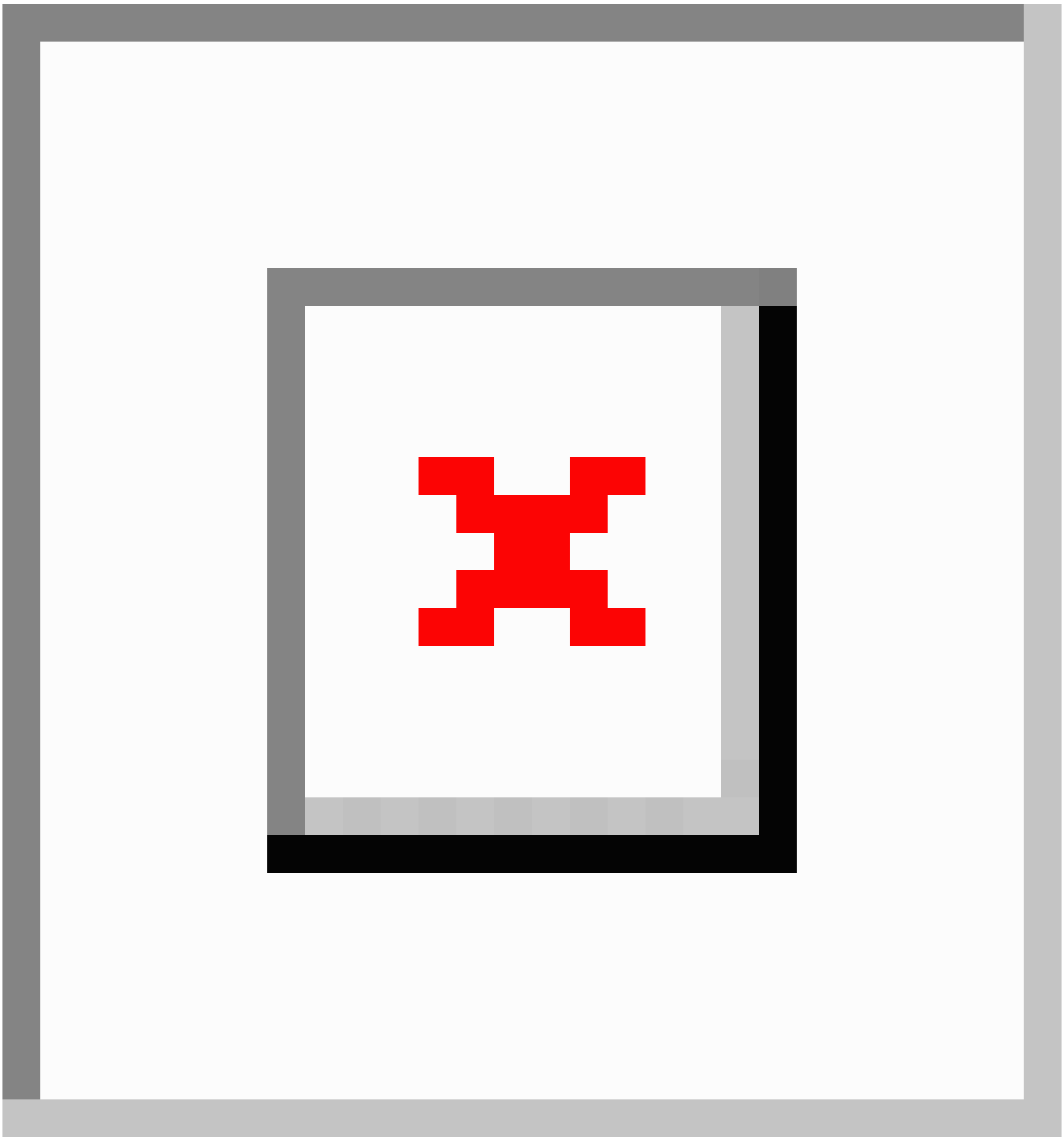


Enzyme Supplementation’s Impact on HR in Exergaming Combined With HIIE

The results demonstrated that during exergaming combined with HIIE, older adult participants experienced a significant increase in HR compared with before exercise (*P*<.05). The estimated HRmax (*220 – age, SD 10*) for older adults was 155 (SD 10) beats per minute (bpm), and the observed HRs during exercise exceeded 85% of the estimated HRmax for both groups. However, there was no significant difference in the average HR of the older adults between the enzyme and placebo groups before and after supplementation. Before supplementation, there was no significant difference in the HRs of the older adult participants between the enzyme and placebo groups during the first (mean 104.63, SD 24.71; 95% CI 86.77-122.48 vs mean 104.63, SD 22.32; 95% CI 86.77-122.48 bpm; ES=0; *P*>.99), second (mean 116.13, SD 21.81; 95% CI 101.54-141.72 vs mean 114.38, SD 16.26; 95% CI 99.79-128.96 bpm; ES=0.09; *P*=.86), third (mean 126.13, SD 24.30; 95% CI 110.53-141.72 vs mean 118.38, SD 15.96; 95% CI 102.78-133.97 bpm; ES=0.38; *P*=.46), fourth (mean 130.13, SD 21.97; 95% CI 114.63-145.62 vs mean 121.38, SD 18.77; 95% CI 105.88-136.87 bpm; ES=0.43; *P*=.41), fifth (mean 124.63, SD 20.49; 95% CI 109.37-139.88 vs mean 125.63, SD 19.73; 95%

CI 110.37-140.88 bpm; ES=−0.05; *P*=.92), sixth (mean 128.88, SD 23.34; 95% CI 112.71-145.04 vs mean 127.38, SD 19.06; 95% CI 111.21-143.54 bpm; ES=0.07; *P*=.89), seventh (mean 131.63, SD 22.01; 95% CI 116.56-146.69 vs mean 129.75, SD 17.45; 95% CI 114.69-144.81 bpm; ES=0.09; *P*=.85), and eighth (mean 137.75, SD 23.60; 95% CI 121.78-153.73 vs mean 135.63, SD 18.19; 95% CI 119.65-151.60 bpm; ES=0.10; *P*=.84) sets (Figure 2). Similarly, after supplementation, there was no significant difference in the HRs of the enzyme and placebo groups during the first (mean 102.13, SD 21.21; 95% CI 88.87-115.67 vs mean 109.88, SD 13.64; 95% CI 96.33-123.42 bpm; ES=−0.43; *P*=.40), second (mean 111.13, SD 18.04; 95% CI 98.32-123.93 vs mean 114.88, SD 15.63; 95% CI 102.07-127.68 bpm; ES=−0.22; *P*=.66), third (mean 121.13, SD 19.38; 95% CI 108.63-133.62 vs mean 115.38, SD 12.95; 95% CI 102.88-127.87 bpm; ES=0.35; *P*=.50), fourth (mean 128.25, SD 18.75; 95% CI 116.57-139.93 vs mean 118.5, SD 11.10; 95% CI 106.82-130.18 bpm; ES=0.63; *P*=.23), fifth (mean 129.25, SD 18.12; 95% CI 115.29-143.21 vs mean 128.75, SD 18.68; 95% CI 114.79-142.71 bpm; ES=0.03; *P*=.96), sixth (mean 132.00, SD 19.79; 95% CI 118.70-145.30 vs mean 125.50, SD 14.95; 95% CI 112.20-138.80 bpm; ES=0.37; *P*=.47), seventh (mean 133.88, SD 20.84; 95% CI 120.50-147.25 vs mean 127.63, SD 13.70; 95% CI 114.25-141.00 bpm; ES=0.35; *P*=.49), and eighth (mean 137.75, SD 21.18; 95% CI 124.42-151.08 vs mean 134.63, SD 13.02; 95% CI 121.30-147.95 bpm; ES=0.18; *P*=.73) sets (Figure 5). In summary, the findings indicate that exergaming combined with HIIE leads to a significant increase in HR among older adults. However, there was no significant difference in HR between the enzyme and placebo groups before and after supplementation.

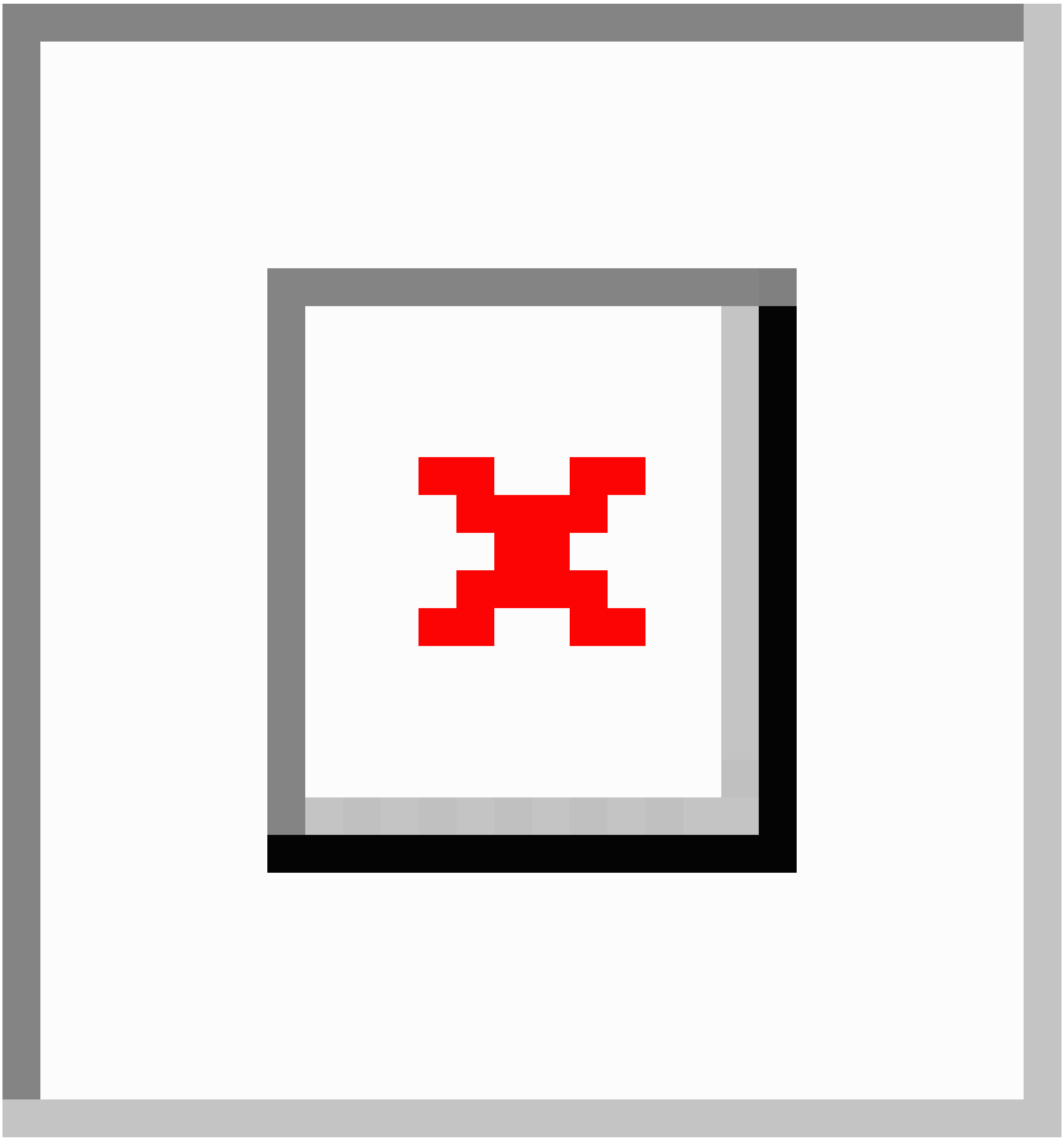
Figure 5. Heat rate response (A) before and (B) after 14 days of enzyme or placebo supplementation. Data are presented as mean (SD). * indicates a significant difference (*P*<.05) from the pre-exercise value within the group. Ex-: bouts of HIIE; HIIE: high-intensity interval exercise; post–5 min: after 5 minutes of HIIE; post–10 min: after 10 minutes of HIIE.



TRIMP in Enzyme Versus Placebo Groups After Supplementation in Exergaming Combined With HIIE

The TRIMP, representing both objective and subjective training loads, was compared between the enzyme and placebo groups after supplementation. Analysis revealed no significant differences in either the objective (mean 542.5, SD 172.19 vs mean 531.25, SD 123.34 AU; ES=0.08; *P*=.88) or subjective training loads (mean 895, SD 143.73 vs mean 847.50, SD 223.46 AU; ES=0.25; *P*=.62) between the groups (Figure 6). This suggests that the supplementation did not significantly alter the perceived intensity or effort of the HIIE when combined with exergaming.

Figure 6. Comparison of (A) objective and (B) subjective training impulse (TRIMP) between the enzyme and placebo groups after supplementation during HIIE with exergaming. Data are presented as mean (SD). AU: arbitrary unit; Ex-: bout of HIIE; HIIE: high-intensity interval exercise; HRmax: maximum heart rate; RPE: rating of perceived exertion.



Discussion

Principal Findings

The study investigated the effects of enzyme supplementation on lactate response and HR in older adult individuals engaging in a combination of exergaming and HIIE. The results indicated that enzyme supplementation significantly reduced blood lactate levels after exercise, particularly after the fourth (*P*=.001) and eighth (*P*<.001) exercise bouts, demonstrating the potential of enzymes to mitigate exercise-induced lactate accumulation. Despite a notable increase in HR during the exercise sessions, which surpassed 85% of the estimated HRmax for older adult participants, there was no discernible difference in HR responses between the enzyme and placebo groups, either before or after supplementation. Furthermore, the analysis of TRIMP, encompassing both objective and subjective measures of training load, revealed no significant differences between the enzyme and placebo groups after supplementation. This suggests that although enzyme supplementation may aid in lactate management, it does not significantly impact the overall perceived intensity or cardiovascular demand of HIIE combined with exergaming in older adult individuals.

Lactate Response in Exergaming

This study contributes valuable insights into the efficacy of fruit and vegetable enzyme supplementation in optimizing exercise outcomes for older adults, particularly when combined with HIIE and exergaming. The notable finding that blood lactate levels surpassed the 4 mmol/L threshold after the fourth exercise bout underlines the high intensity and physiological rigor of the exercise protocol. This study's emphasis on enzyme supplementation's impact on blood lactate levels is especially pertinent. Enzyme supplementation significantly lowered blood lactate levels after exercise, as compared to the placebo, after both the fourth (P=.001) and eighth (P<.001) exercise bouts and at 5 and 10 minutes after exercise (P=.003 and P=.002, respectively). This observation suggests a potential role of enzyme supplementation in enhancing lactate metabolism, either through its reduction or improved clearance during and after high-intensity exercise. The metabolism-enhancing attributes of fruit and vegetable enzymes, such as bromelain and papain, may facilitate this reduction in lactate accumulation [22,37]. Furthermore, their antioxidant and anti-inflammatory properties could lead to enhanced muscle function, thereby contributing to lower lactate production [38].

Exergaming, when integrated with HIIE, presents an innovative and engaging exercise modality, particularly for older adults. It has been established as an effective and enjoyable exercise option, capable of achieving intensities comparable to traditional exercise forms [16]. This study reinforces the feasibility of exergaming combined with HIIE as a viable strategy for older adults, achieving substantial exercise intensity as evidenced by elevated lactate levels. However, the study is not without limitations. The relatively small sample size and focus on a specific demographic and exercise protocol may restrict the broader applicability of the findings. Further research with larger, more diverse populations is necessary to validate and extend these preliminary results.

HR Response in Exergaming

Interestingly, although exergaming combined with HIIE effectively elevated physiological parameters such as HR and lactate levels, no significant difference in HR response was observed between the enzyme and placebo groups. This suggests that the subjective perception of effort might not accurately reflect the actual physiological demands of the exercise, echoing previous research [32,33]. In summary, this study illustrates that enzyme supplementation can potentially reduce blood lactate levels during and after high-intensity exercise in an older adult cohort engaged in HIIE combined with exergaming. These findings underscore the value of enzyme supplementation in enhancing metabolic responses and optimizing exercise outcomes. Future research should aim to unravel the underlying mechanisms and investigate the long-term impacts of enzyme supplementation across diverse populations. A deeper understanding of the interplay between nutritional supplementation, exercise modality, and physiological responses is crucial in tailoring effective interventions for optimal exercise performance and overall health promotion.

TRIMP Response in Exergaming

An additional focal point of our study was the evaluation of TRIMP in relation to enzyme supplementation during HIIE combined with exergaming. TRIMP is a quantifiable measure of training load, incorporating both objective and subjective elements of exercise intensity [32]. In our study, the analysis revealed no significant differences in TRIMP between the enzyme and placebo groups after supplementation. This outcome suggests that enzyme supplementation does not significantly alter the perceived intensity or exertion levels during HIIE with exergaming. This finding aligns with previous studies that have explored the multifaceted nature of TRIMP. For instance, research by Laursen and Jenkins [39] highlighted the complexity of accurately measuring training load, emphasizing the need to consider both physiological and psychological factors. The lack of significant difference in TRIMP in our study could be attributed to the stable physiological responses (HRs and lactate levels) observed across both groups. This observation is consistent with the work of Manzi et al [40], who noted the importance of physiological markers in determining training load, particularly in endurance sports.

Furthermore, the subjective component of TRIMP, which relates to athletes' perceived exertion, is a crucial aspect of training load assessment [35]. Our study's findings, where the subjective perception of effort did not significantly differ between the enzyme and placebo groups, resonate with the notion that perceived exertion is a complex interplay of physical and psychological factors [33]. This complexity might explain why enzyme supplementation, primarily impacting physiological responses, did not significantly alter the subjective experience of the training load. The implication of these results is substantial for designing exercise programs for older adults. As suggested by Bethancourt et al [41], understanding and managing training load is crucial in preventing overtraining and optimizing exercise benefits, especially in older adults. The lack of difference in TRIMP between the groups in our study indicates that enzyme supplementation, although beneficial in reducing lactate levels, does not necessarily impact the overall training load as perceived by the participants. This insight is vital for practitioners and researchers in tailoring exercise regimens that are both physiologically effective and psychologically manageable for older adults.

In conclusion, our study contributes to the growing body of knowledge on TRIMP and its applications in exercise science. Although enzyme supplementation shows promise in reducing lactate levels, its impact on the overall training load, as measured by TRIMP, appears to be minimal. Future research should continue to explore this area, considering both physiological and psychological aspects of exercise, to develop comprehensive training strategies for various populations, including older adults.

Conclusions

This study aimed to evaluate the impact of fruit and vegetable enzyme supplementation on aerobic capacity and blood lactate response in older adults participating in HIIE combined with exergaming. The results demonstrate that enzyme supplementation significantly reduced blood lactate levels after exercise compared to a placebo. This finding is indicative of the potential role of such supplementation in enhancing lactate metabolism during and after high-intensity exercise. Additionally, the integration of HIIE with exergaming has proven to be a novel and effective approach to exercise for older adults, achieving significant physiological intensities while maintaining engagement and enjoyment. However, enzyme supplementation did not exhibit a noticeable effect on HR response or the overall perceived training load, as measured by TRIMP. This suggests that although enzyme supplementation may influence specific physiological responses, such as lactate production and clearance, it does not significantly alter the overall perceived exertion or exercise experience for participants.

These findings contribute to the growing body of literature on the synergistic effects of nutritional supplementation and innovative exercise modalities such as exergaming in the older adult population. They highlight the potential of enzyme supplementation in optimizing exercise outcomes, particularly in reducing lactate accumulation, which is a crucial aspect of high-intensity exercise tolerance. Moreover, the study underscores the feasibility and effectiveness of exergaming combined with HIIE as a strategy to enhance physical activity levels in older adults. The study's implications extend beyond exercise physiology, offering practical insights for health practitioners, fitness professionals, and researchers in the development of targeted, effective, and enjoyable exercise interventions for older adults. Future research should aim to further elucidate the mechanisms behind enzyme supplementation's impact on exercise performance and explore the long-term effects of such interventions in a wider demographic.

In summary, this research supports the notion that carefully tailored nutritional and exercise interventions, such as enzyme supplementation combined with HIIE and exergaming, can significantly enhance exercise outcomes in older adults. These interventions hold promise for improving overall health and well-being in this demographic, contributing to the growing field of serious games and their application in health and fitness.

Acknowledgments

We thank all members of our research team for their contributions to this study and all participants for taking part in this study. Financial support was provided by the Taiwan Ministry of Education's 'Industry Academy Program' and a grant from the Tainan University of Technology. We also extend our gratitude to Taiwan Enzyme Village Co., Ltd. for participating in the industry-academia collaboration, which facilitated the successful completion of the experiment.

Authors' Contributions

SCL carried out the laboratory experiments, analyzed the data, interpreted the results, prepared figures and tables, and prepared the manuscript. CYW, THH, CCW, and HCC assisted in the data collection and the discussion of the literature. CCW designed the study, supervised the experimental procedure, and reviewed the entire preparation of the manuscript.

Conflicts of Interest

None declared.

Editorial Notice

This randomized study was only retrospectively registered. It was not prospectively registered as the authors considered it to be a feasibility study. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Checklist 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File, 1123 KB - games_v12i1e52231_app1.pdf]

References

1. Duttaroy S, Thorell D, Karlsson L, Börjesson M. A single-bout of one-hour spinning exercise increases troponin T in healthy subjects. Scand Cardiovasc J 2012 Feb;46(1):2-6. [doi: 10.3109/14017431.2011.622783] [Medline: 22214280]
2. Wadley AJ, Turner JE, Aldred S. Factors influencing post-exercise plasma protein carbonyl concentration. Free Radic Res 2016 Feb 12;50(4):375-384. [doi: 10.3109/10715762.2015.1131824] [Medline: 26873473]
3. Muldoon MF, Laderian B, Kuan DCH, Sereika SM, Marsland AL, Manuck SB. Fish oil supplementation does not lower C-reactive protein or interleukin-6 levels in healthy adults. J Intern Med 2016 Jan;279(1):98-109. [doi: 10.1111/joim.12442] [Medline: 26497831]
4. Clarkson PM, Hubal MJ. Exercise-induced muscle damage in humans. Am J Phys Med Rehabil 2002 Nov;81(11 Suppl):S52-S69. [doi: 10.1097/00002060-200211001-00007] [Medline: 12409811]

5. Finaud J, Lac G, Filaire E. Oxidative stress: relationship with exercise and training. Sports Med 2006;36(4):327-358. [doi: 10.2165/00007256-200636040-00004] [Medline: 16573358]
6. Tom A, Nair KS. Assessment of branched-chain amino acid status and potential for biomarkers. J Nutr 2006 Jan;136(1 Suppl):324S-330S. [doi: 10.1093/jn/136.1.324S] [Medline: 16365107]
7. Del Coso J, González-Millán C, Salinero JJ, et al. Muscle damage and its relationship with muscle fatigue during a half-iron triathlon. PLoS One 2012 Aug 10;7(8):e43280. [doi: 10.1371/journal.pone.0043280] [Medline: 22900101]
8. Viana RB, de Lira CAB, Naves JPA, Coswig VS, Del Vecchio FB, Gentil P. Tabata protocol: a review of its application, variations and outcomes. Clin Physiol Funct Imaging 2019 Jan;39(1):1-8. [doi: 10.1111/cpf.12513] [Medline: 29608238]
9. Pearson RC, Olenick AA, Green ES, Jenkins NT. Tabata-style functional exercise increases resting and postprandial fat oxidation but does not reduce triglyceride concentrations. Exp Physiol 2020 Mar;105(3):468-476. [doi: 10.1113/EP088330] [Medline: 31916294]
10. Menz V, Marterer N, Amin SB, Faulhaber M, Hansen AB, Lawley JS. Functional vs. running low-volume high-intensity interval training: effects on VO2max and muscular endurance. J Sports Sci Med 2019 Sep;18(3):497-504. [Medline: 31427872]
11. Eckstrom E, Neukam S, Kalin L, Wright J. Physical activity and healthy aging. Clin Geriatr Med 2020 Nov;36(4):671-683. [doi: 10.1016/j.cger.2020.06.009] [Medline: 33010902]
12. Gist NH, Fedewa MV, Dishman RK, Cureton KJ. Sprint interval training effects on aerobic capacity: a systematic review and meta-analysis. Sports Med 2014 Feb;44(2):269-279. [doi: 10.1007/s40279-013-0115-0] [Medline: 24129784]
13. Kercher VM, Kercher K, Bennion T, et al. Fitness trends from around the globe. ACSMs Health Fit J 2022;26(1):21-37. [doi: 10.1249/FIT.0000000000000737]
14. Bock BC, Dunsiger SI, Ciccolo JT, et al. Mediators of physical activity between standard exercise and exercise video games. Health Psychol 2019 Dec;38(12):1107-1115. [doi: 10.1037/hea0000791] [Medline: 31512923]
15. Comeras-Chueca C, Marin-Puyalto J, Matute-Llorente A, Vicente-Rodriguez G, Casajus JA, Gonzalez-Aguero A. The effects of active video games on health-related physical fitness and motor competence in children and adolescents with healthy weight: a systematic review and meta-analysis. Int J Environ Res Public Health 2021 Jun 29;18(13):6965. [doi: 10.3390/ijerph18136965] [Medline: 34209767]
16. McDonough DJ, Pope ZC, Zeng N, Lee JE, Gao Z. Comparison of college students' energy expenditure, physical activity, and enjoyment during exergaming and traditional exercise. J Clin Med 2018 Nov 10;7(11):433. [doi: 10.3390/jcm7110433] [Medline: 30423805]
17. Park SB, Kim M, Lee E, et al. Energy system contributions and physical activity in specific age groups during exergames. Int J Environ Res Public Health 2020 Jul 7;17(13):4905. [doi: 10.3390/ijerph17134905] [Medline: 32646023]
18. Röglin L, Ketelhut S, Ketelhut K, et al. Adaptive high-intensity exergaming: the more enjoyable alternative to conventional training approaches despite working harder. Games Health J 2021 Dec;10(6):400-407. [doi: 10.1089/g4h.2021.0014] [Medline: 34558966]
19. Weston M, Taylor KL, Batterham AM, Hopkins WG. Effects of low-volume high-intensity interval training (HIT) on fitness in adults: a meta-analysis of controlled and non-controlled trials. Sports Med 2014 Jul;44(7):1005-1017. [doi: 10.1007/s40279-014-0180-z] [Medline: 24743927]
20. Batrakoulis A, Jamurtas AZ, Fatouros IG. High-intensity interval training in metabolic diseases: physiological adaptations. ACSMs Health Fit J 2021;25(5):54-59. [doi: 10.1249/FIT.0000000000000703]
21. Batrakoulis A, Fatouros IG. Psychological adaptations to high-intensity interval training in overweight and obese adults: a topical review. Sports (Basel) 2022 Apr 22;10(5):64. [doi: 10.3390/sports10050064] [Medline: 35622474]
22. Şanlier N, Gökcen BB, Sezgin AC. Health benefits of fermented foods. Crit Rev Food Sci Nutr 2019;59(3):506-527. [doi: 10.1080/10408398.2017.1383355] [Medline: 28945458]
23. Brooks GA. Lactate as a fulcrum of metabolism. Redox Biol 2020 Aug;35:101454. [doi: 10.1016/j.redox.2020.101454] [Medline: 32113910]
24. Lucertini F, Gervasi M, D'Amen G, et al. Effect of water-based recovery on blood lactate removal after high-intensity exercise. PLoS One 2017 Sep 6;12(9):e0184240. [doi: 10.1371/journal.pone.0184240] [Medline: 28877225]
25. Clifford MN. Diet-derived phenols in plasma and tissues and their implications for health. Planta Med 2004 Dec;70(12):1103-1114. [doi: 10.1055/s-2004-835835] [Medline: 15643541]
26. Tsao JP, Liu CC, Wang HF, Bernard JR, Huang CC, Cheng IS. Oral resveratrol supplementation attenuates exercise-induced interleukin-6 but not oxidative stress after a high intensity cycling challenge in adults. Int J Med Sci 2021 Mar 18;18(10):2137-2145. [doi: 10.7150/ijms.55633] [Medline: 33859520]
27. Ma GD, Chiu CH, Hsu YJ, Hou CW, Chen YM, Huang CC. Changbai mountain ginseng (Panax ginseng C.A. Mey) extract supplementation improves exercise performance and energy utilization and decreases fatigue-associated parameters in mice. Molecules 2017 Feb 5;22(2):237. [doi: 10.3390/molecules22020237] [Medline: 28165424]
28. Mohr T, Desser L. Plant proteolytic enzyme papain abrogates angiogenic activation of human umbilical vein endothelial cells (HUVEC) in vitro. BMC Complement Altern Med 2013 Sep 21;13:231. [doi: 10.1186/1472-6882-13-231] [Medline: 24053149]

29. Flanagan EP, Jakeman PM. Oral creatine supplementation and short-term dynamic power production in healthy young men. Presented at: 24 International Symposium on Biomechanics in Sports (2006); Jul 14 to 18, 2006; Salzburg, Austria URL: https://ojs.ub.uni-konstanz.de/cpa/article/view/218 [accessed 2024-05-21]
30. Crisafulli DL, Buddhadev HH, Brilla LR, Chalmers GR, Suprak DN, San Juan JG. Creatine-electrolyte supplementation improves repeated sprint cycling performance: a double blind randomized control study. J Int Soc Sports Nutr 2018 May 2;15:21. [doi: 10.1186/s12970-018-0226-y] [Medline: 29743825]
31. Hsu MC, Chien KY, Hsu CC, Chung CJ, Chan KH, Su B. Effects of BCAA, arginine and carbohydrate combined drink on post-exercise biochemical response and psychological condition. Chin J Physiol 2011 Apr 30;54(2):71-78. [doi: 10.4077/cjp.2011.amk075] [Medline: 21789887]
32. Banister EW, Calvert TW. Planning for future performance: implications for long term training. Can J Appl Sport Sci 1980 Sep;5(3):170-176. [Medline: 6778623]
33. Edwards S. High performance training and racing. In: The Heart Rate Monitor Book: Feet Fleet Press; 1993:113-123.
34. Foster C, Hector LL, Welsh R, Schrager M, Green MA, Snyder AC. Effects of specific versus cross-training on running performance. Eur J Appl Physiol Occup Physiol 1995;70(4):367-372. [doi: 10.1007/BF00865035] [Medline: 7649149]
35. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. J Strength Cond Res 2001 Feb;15(1):109-115. [Medline: 11708692]
36. Borg G, Ljunggren G, Ceci R. The increase of perceived exertion, aches and pain in the legs, heart rate and blood lactate during exercise on a bicycle ergometer. Eur J Appl Physiol Occup Physiol 1985;54(4):343-349. [doi: 10.1007/BF02337176] [Medline: 4065121]
37. Samtiya M, Aluko RE, Dhewa T, Moreno-Rojas JM. Potential health benefits of plant food-derived bioactive components: an overview. Foods 2021 Apr 12;10(4):839. [doi: 10.3390/foods10040839] [Medline: 33921351]
38. Lavefve L, Marasini D, Carbonero F. Microbial ecology of fermented vegetables and non-alcoholic drinks and current knowledge on their impact on human health. Adv Food Nutr Res 2019 Dec 7;87:147-185. [doi: 10.1016/bs.afnr.2018.09.001] [Medline: 30678814]
39. Laursen PB, Jenkins DG. The scientific basis for high-intensity interval training. Sports Med 2002 Jan;32:53-73. [doi: 10.2165/00007256-200232010-00003]
40. Manzi V, Iellamo F, Impellizzeri F, D'ottavio S, Castagna C. Relation between individualized training impulses and performance in distance runners. Med Sci Sports Exerc 2009 Nov;41(11):2090-2096. [doi: 10.1249/MSS.0b013e3181a6a959] [Medline: 19812506]
41. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014 Sep;12(1-2):10-20. [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]

Abbreviations

AU: arbitrary unit
bpm: beats per minute
ES: effect size
HIIE: high-intensity interval exercise
HR: heart rate
RPE: rating of perceived exertion
TRIMP: training impulse

Edited by G Eysenbach, T Leung; submitted 28.08.23; peer-reviewed by A Batrakoulis, C Aygün, S Sumartiningsih; revised version received 26.02.24; accepted 25.04.24; published 25.06.24.

Please cite as:
Lin SC, Wang CY, Hou TH, Chen HC, Wang CC
Impact of Fruit and Vegetable Enzyme Supplementation on Aerobic Performance and Lactate Response in Older Adults Following High-Intensity Interval Exercise Through Exergaming: Randomized Experimental Matched-Pair Study
JMIR Serious Games 2024;12:e52231
URL: https://games.jmir.org/2024/1/e52231
doi:10.2196/52231

Effect of Elastic Resistance on Exercise Intensity and User Satisfaction While Playing the Active Video Game BoxVR in Immersive Virtual Reality: Empirical Study

Jacek Polechoński[1], PhD; Alan Przepiórzyński[2], MSc; Piotr Polechoński[2]; Rajmund Tomik[3], PhD

1

2

3

Corresponding Author:
Jacek Polechoński, PhD

Abstract

Background: One of the main contemporary forms of physical activity (PA) involves exercises and games in an immersive virtual reality (VR) environment, which allows the user to practice various forms of PA in a small space. Unfortunately, most of the currently available VR games and workout applications are mostly based on upper body movements, especially the arms, which do not guarantee sufficiently high exercise intensity and health benefits. Therefore, it is worth seeking solutions to help increase the exercise load during PA in VR.

Objective: The main aim of this study was to evaluate the effect of elastic arm resistance in the form of latex resistance bands of different elasticity levels on the intensity of students' PA while playing the BoxVR game. We further assessed the satisfaction of this form of exercise and its associations with PA intensity.

Methods: A total of 21 healthy and physically fit men (mean age 22.5, SD 2.0 years) were included in the study. The tests consisted of 3 10-minute games. One game was run with no load and the other two were run with 1.5-meter latex resistance bands (low and high resistance). The order of the tests was randomized and the participants rested for 20 minutes after each exercise. Exercise intensity was estimated using objective (heart rate monitoring) and subjective (Borg scale) methods. The Physical Activity Enjoyment Scale was used to assess satisfaction with the PA. The effect of elastic resistance on exercise intensity and user enjoyment was estimated using ANOVA for repeated measures.

Results: The ANOVA results indicated that incorporation of elastic resistance caused a significant change ($F_{2,40}$=20.235, P<.001; $\eta^2 p$=0.503) in the intensity of PA in VR, which was low while playing without resistance and then increased to a moderate level with additional resistance. The use of elastic bands also changed participants' perceptions of the enjoyment of exercise in VR ($F_{2,40}$=9.259, P<.001; $\eta^2 p$=0.316). The students rated their satisfaction with PA in VR on a 7-point scale highly and similarly when exercising without an upper limb load (mean 6.19, SD 0.61) and with slight elastic resistance (mean 6.17, SD 0.66), whereas their satisfaction declined significantly (mean 5.66, SD 0.94) when incorporating a higher load.

Conclusions: The intensity of PA among students playing the BoxVR game is at a relatively low level. With the added resistance of elastic bands attached to the upper limbs, the intensity of the exercise increased to a moderate level, as recommended for obtaining health benefits. Participants rated the enjoyment of PA in VR highly. The use of slight elastic resistance did not negatively affect satisfaction with the BoxVR game, although user satisfaction declined with a higher load. Further research should be undertaken to increase the effectiveness of exercise in VR so that regular users can enjoy the health benefits.

(JMIR Serious Games 2024;12:e58411) doi:10.2196/58411

KEYWORDS

virtual reality; VR; game; gaming; immersive; immersion; health-related physical activity; physical activity; exercise; active video games; attractiveness; enjoyment scale; enjoyment; serious games; elastic resistance; resistance

Introduction

The past few years have seen a rapid development of technologies related to immersive virtual reality (VR). With immersive VR, the user is cut off from the visual and auditory stimuli of the surrounding reality and instead receives artificially produced images, sounds, and even tactile sensations using information technology, which is finding increasing applications in various areas of human life. In particular, VR is increasingly used for physical activity (PA) and in the development and diagnosis of physical fitness parameters. VR applications are being developed to shape and assess motor skills [1-5], research on motion analysis in VR is now actively carried out [6,7], and active virtual reality games (AVRGs) are becoming increasingly popular [8]. Feedback from users indicates the attractiveness of AVRGs, making them competitive with conventional forms of PA [9-11]. Some reports have also shown that VR can offer greater flow for PA than a similar form of exercise in the real world [12]. The great potential of AVRGs is also linked to the fact that these types of applications enable practicing different forms of exercise in a small space at home. However, one of the main limitations of VR technology in the context of its use for PA is that most of the exercises performed in the virtual environment are primarily based on movements of the upper body as the sensors placed in the VR headsets allow for tracking movements of the head and possibly the torso, while the movement sensors located in the controllers allow for tracking arm movements. Although some trainers currently work with VR headsets, such as virtual treadmills, flight simulators, and cycle ergometers, these are relatively expensive and take up space in living areas. Since PA performed in VR is mainly based on upper limb movements, there is a concern that this type of exercise may be characterized by relatively low intensity. Consequently, physical efforts practiced in a virtual environment may not be effective in terms of potential health benefits. According to World Health Organization (WHO) recommendations, PA should be characterized by moderate to high intensity to obtain health benefits [13]. Therefore, solutions should be sought to increase the body's workload when practicing PA in VR.

The first related studies are already being carried out. One of the proposals to increase the intensity of physical exercise in VR is to use an additional load in the form of handheld weights (HHWs). Based on the experiments carried out to date, this type of solution can be effective [10]. An alternative solution to Velcro-fastened weights could be elastic resistance in the form of rubber bands attached to the distal parts of the upper limbs. Indeed, resistance bands have been widely used in fitness classes and various sports to improve the effectiveness of training [14-17]. Such a solution may be used provided there is effective stretching of the elastic bands during arm movements in VR. Applications that meet this condition include the popular AVRGs based on boxing techniques. With appropriately fastened straps, boxing movements can be performed with elastic resistance, which should potentially increase the intensity of this type of PA. However, the specific amount of resistance that should be applied to increase the effectiveness of the exercise while not causing discomfort due to the excessive load remains unclear, as this could negatively affect the attractiveness of PA perceived by users.

Therefore, the main aim of this study was to evaluate the effect of elastic resistance in the form of latex resistance bands with different elastic properties on exercise intensity in young and physically fit adults while playing the popular AVRG game BoxVR [18]. The results obtained were related to the WHO health recommendations for PA. This study further assessed the attractiveness of such a form of exercise and the relationships between the use of elastic resistance and user satisfaction. It was hypothesized that the use of resistance bands would significantly increase the participants' exercise intensity and would not significantly affect their assessment of satisfaction with playing the AVRG.

Methods

Participants

The study involved 21 healthy and physically fit men studying at the Academy of Physical Education in Katowice, Poland (mean age 22.5, SD 2.0 years; mean body height 181.6, SD 7.3 centimeters; mean body weight 79.5, SD 11.0 kilograms). People with motion sickness, sensitivity to flashing lights, epileptic seizures, and balance disorders were excluded from the study. The research was carried out at the Jerzy Kukuczka Academy of Physical Education in Katowice, Poland, at a certified Laboratory of Research on Pro-Health Physical Activity (PN-EN ISO 9001:2015, certificate validity: 7.12.2021 - 16.12.2024).

Ethical Considerations

The study was conducted according to the guidelines of the Declaration of Helsinki, and was reviewed and approved by the Research Ethics Committee of the Jerzy Kukuczka Academy of Physical Education in Katowice (protocols: 9/2018; KB/27/2022). All participants took part in the study voluntarily and could discontinue their participation at any time. All participants were familiarized in detail with the purpose of the study, safety rules, the use of the VR equipment, and the course of the study. In addition, a written informed consent form was provided to all eligible study participants. The test results were secured in accordance with the security procedures in force at the Laboratory of Research on Pro-Health Physical Activity.

Research Tools and Procedures

An HTC Vive (HTC Corporation, New Taipei, Taiwan) kit was used for immersive VR, consisting of a headset, two base stations, two controllers, and a computer. The HTC Vive set is one of the popular VR systems on the market, which is characterized by high visual quality and allows for realistic VR experiences. This system was selected for this research since the motion-tracking system in HTC Vive is very precise and accurate, which allows the user to move smoothly and naturally in the virtual environment. Owing to a set of sensors and controllers, users can freely explore the virtual world. The BoxVR application was used with several training programs of varying difficulty. These programs involve boxing routines combined with music, reminiscent of shadow boxing. The game is based on basic boxing punches (ie, straight, hook, and undercut), which are performed on virtual objects coming from the depths of the room to the rhythm of the music in various combinations. There are also shapes the user has to avoid or block. The user hits targets with their hands, which they perceive in the virtual environment as boxing gloves in two colors: blue (left hand) and pink (right hand). An illustration of the playing environment is provided in Figure 1. To obtain a high score in the game, the user has to execute the punches correctly and hit the targets moving toward them that correspond to the appropriate glove color. The user scores points for every correct response. In the case of a series of several or more accurate hits, the score is further multiplied by an appropriate multiplier.

Information on the number of points scored and the duration of the game is displayed on virtual screens in front of the user.

The systsem has a panel that allows the user to select game modes with different durations, levels of difficulty, and nature of exercise. For the purposes of the study, mode “seventeen” was selected lasting 10 minutes, set at medium difficulty and with the no squat option. This mode was considered to be optimal for physically fit young adults. The squat option was disabled because the aim of the study was to assess the impact of the elastic resistance of the arms on the intensity of physical exercise; therefore, additional lower limb exercises would be a factor that could complicate the interpretation of the study results. It should be noted that many active video games (AVGs) are based solely on arm and torso movements. Before the study, participants were familiarized with the use of the application and took part in a short (2-minute) no-load trial.

The tests consisted of 3 10-minute games. One game was run with no load and the other two were run with 1.5-meter latex resistance bands, including green (low resistance) and silver (high resistance) bands. The elasticity characteristics of the resistance bands provided by the manufacturer (Thera-Band) are presented in Table 1. The bands were attached to the ground on one side and to tactical gloves (M-Tac) on the other. When starting the exercise with the bands, users held their hands in a guard position and were positioned at an appropriate distance from the point of attachment ensuring that the bands were taut but not stretched. When the punches were performed, the bands stretched, causing resistance (Figure 2). The order of the tests was randomized, and the participant rested for 20 minutes after each exercise before starting the next game.

While playing the VR game, participants’ heart rates were monitored using a Vantage V heart rate monitor (Polar Electro Oy, Kempele, Finland) coupled with a chest strap (Polar H10). Based on the average exercise heart rate (HR_{ave}), the PA intensity was estimated as the average percentage of maximum heart rate (% HR_{max}). The HR_{max} value was first estimated from the formula 208 – 0.7 × age (years) [19]. The results obtained were compared to the PA intensity standards recommended by the American College of Sports Medicine [20]. According to this classification, it is assumed that during low-intensity exercises, HR_{ave}<64% of HR_{max}, high-intensity PA occurs when HR_{ave}≥77% of HR_{max}, and moderate exercise is defined in the condition of HR_{ave}≥64% of HR_{max} with less than 77% of HR_{max}. Furthermore, the average absolute duration of PA (in seconds) was estimated for the following exercise intensity zones: 0, less than 50% HR_{max}; 1, 50% - 59% HR_{max}; 2, 60% - 69% HR_{max}; 3, 70% - 79% HR_{max}; 4, 80% - 89% HR_{max}; and 5, ≥90% HR_{max}. These intensity zones were selected because they are used to report the results in the software of the heart rate monitor (Vantage V) used in the study.

At the end of each test, the participants also self-assessed their perceived exertion using the Borg Rating of Perceived Exertion (RPE), which ranges from 6 to 20 [21,22]. According to this scale, a score of 10-11 indicates low-intensity exercise, a score of 12-13 indicates moderate-intensity exercise, and a score of 14 - 16 indicates high-intensity exercise [23]. The RPE scores were compared with the objective measurements to determine the correlation of PA intensity with the subjective perceptions of the participants.

Subsequently, the participants assessed their satisfaction with PA in VR using the long version of the Physical Activity Enjoyment Scale (PACES), consisting of 18 items [24], which were answered after each test on a 7-point Likert scale. The average calculated from all responses was used for analysis.

Figure 1. Screenshot showing a view of the BoxVR game environment from the user's perspective.

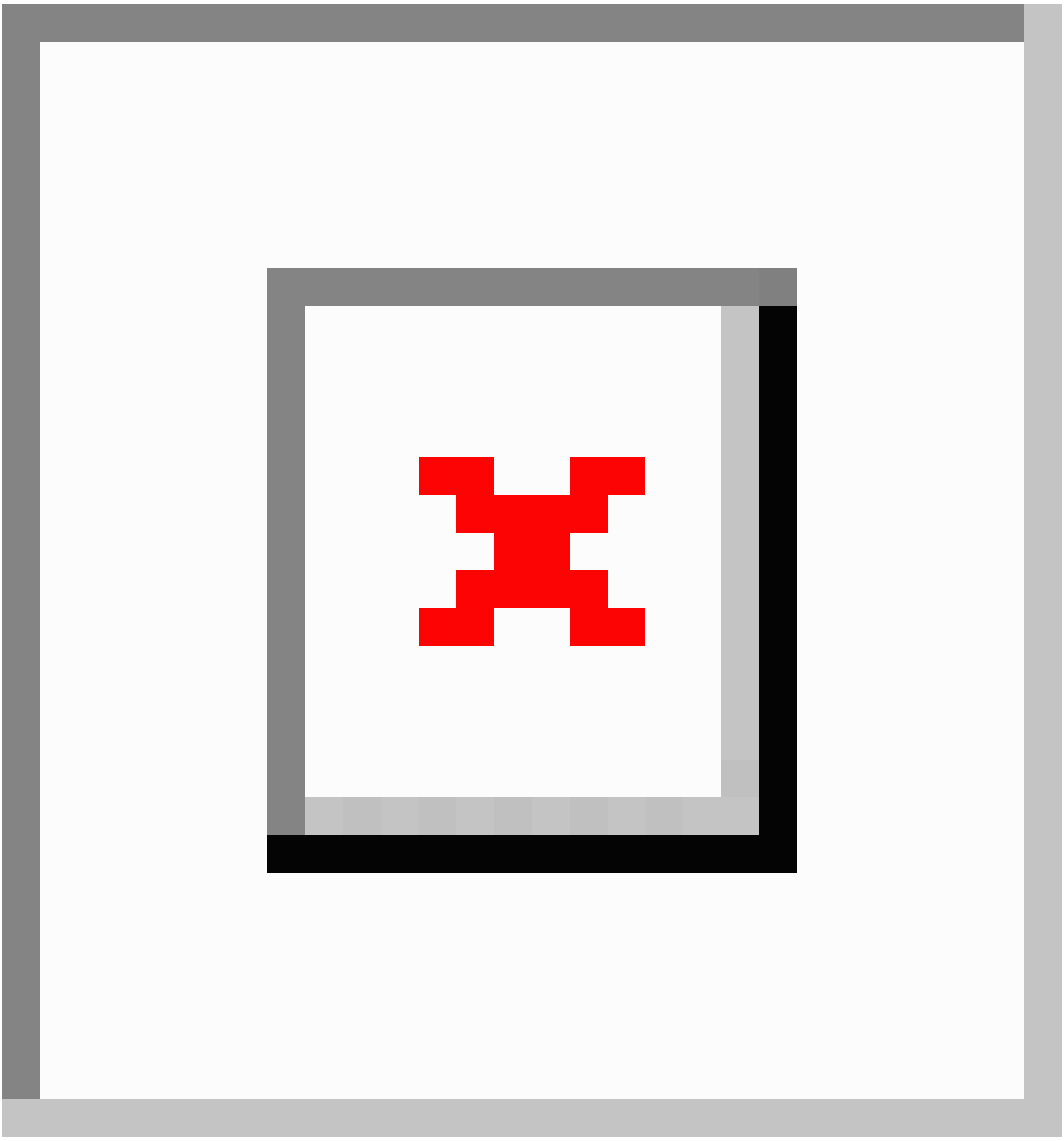


Table . Specifications of the elastic bands (Thera-Band) used during testing.

Stretch, %	Resistance, kilograms	
	Green band	Silver band
25	0.9	2.3
50	1.5	3.9
75	1.9	5.0
100	2.3	6.0
125	2.6	6.9
150	3.0	7.8
175	3.3	8.6
200	3.6	9.5
225	4.0	10.5
250	4.4	11.5

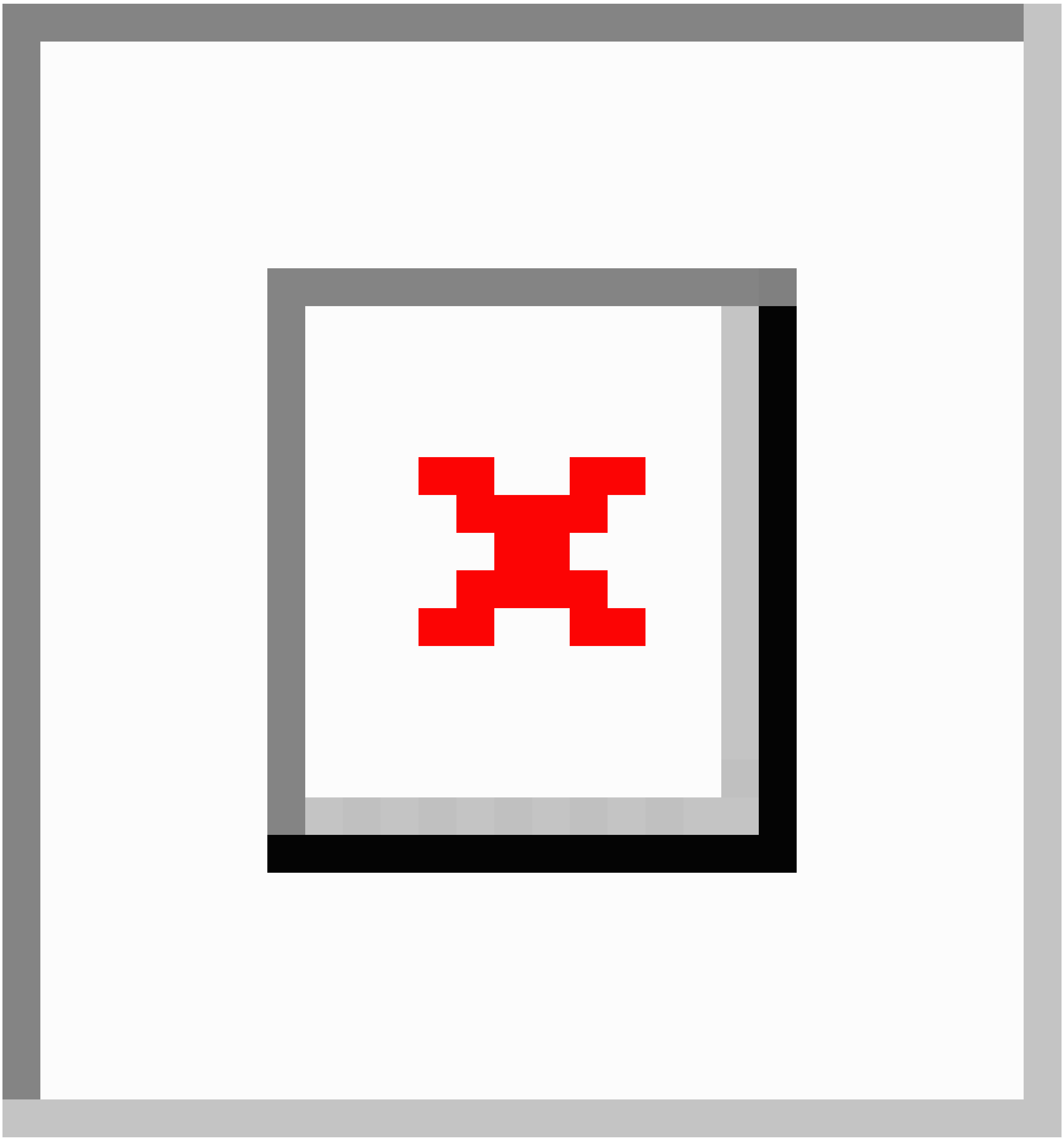


Figure 2. A participant during a resistance exercise with the green band.

Statistical Analysis

Basic descriptive statistics (arithmetic means and SDs) were calculated. The Shapiro-Wilk test was used to assess whether the data followed a normal distribution, whereas sphericity was assessed using the Mauchley test. The effect of elastic resistance on exercise intensity was estimated using ANOVA for repeated measures or Friedman ANOVA, depending on the distribution of the data. The repeated-measures ANOVA was supplemented with Tukey posthoc tests, whereas Friedman ANOVA was followed by the Dunn posthoc test. The level of statistical significance was set at α=.05. The effect size was estimated using η^2p or the Kendall coefficient (*W*). The Spearman rank correlation coefficient (r_S) was used as a measure of the relationship between objective and subjective intensity measures. Statistical analyses were performed using Statistica v.13 (TIBCO Software Inc) and Jamovi v. 2.2.3.0 software.

Results

Exercise Intensity in VR Without Upper Limb Loading and With Elastic Resistance

Repeated-measures ANOVA showed that elastic resistance significantly affected the participants' heart rate ($F_{2,40}$=20.151, *P*<.001; η^2p=0.503). When playing without external resistance, the heart rate was the lowest, with a mean of 117.33 (SD 21.21) beats per minute (bpm). The heart rate increased to a mean of 124.43 (SD 20.62) bpm during play with the green elastic band

and increased further to a mean of 134.90 (SD 20.33) bpm during exercise with the silver band. Posthoc tests showed statistically significant differences between the results of all measurements taken according to the level of resistance applied (Figure 3).

Elastic resistance also resulted in a significant change ($F_{2,40}$=20.235, P<.001; η²p=0.503) in exercise intensity as expressed by the mean %HR_{max}. While playing without resistance, the intensity of physical effort was low, with a mean %HR_{max} of 61.27% (SD 11.21%). With elastic resistance, physical effort (%HR_{max}) increased to a moderate level, as recommended for health benefits, for both the green (mean 64.97%, SD 10.86%) and silver (mean 70.43%, SD 10.68%) bands. Posthoc tests revealed statistically significant differences between the results of all tests according to varying levels of resistance (Figure 4).

Under conditions of no external load, the participants' heart rates lasted the longest in zone 1 (50% - 59% of HR_{max}), whereas in both cases of exercise with elastic bands, the heart rates remained in zone 2 (60% - 69% of HR_{max}). The Friedman ANOVA showed significant variation in results for zones: 0 (χ^2_2=14.711, P<.001; W=0.350), 3 (χ^2_2=8.954, P=.01; W=0.213), and 5 (χ^2_2=9.333, P=.009; W=0.222). No significant effect of arm loading on PA intensity was found for the other zones: 1 (χ^2_2=3.610; P=.17; W=0.086), 2 (χ^2_2=2.913; P=.33; W=0.052), and 4 (χ^2_2=3.360; P=.19; W=0.080) (Figure 5).

We further analyzed the exertion perceived by the users after each test based on the RPE scale (6-20). Friedman ANOVA demonstrated that elastic resistance loading significantly (χ^2_2=36.861, P<.001; W=0.878) altered users' perceptions of exertion. The lowest fatigue was declared by those exercising without additional load. The exercise intensity in this case was rated a mean of 11.19 (SD 2.54) points. In contrast, significantly more exertion was reported for PA with elastic resistance. The mean scores for the intensity of exercise were 13.67 (SD 2.15) for the green band and 16.62 (SD 1.72) for the silver band. Posthoc tests for the pairwise comparisons of the results revealed statistically significant differences (Figure 6).

Spearman correlation analysis between subjective and objective measures of exercise intensity showed a statistically significant positive relationship between intensity evaluated based on %HR_{max} and the Borg RPE scale (6-20) for PA in VR without load (r_S=0.504, P=.02) and with green elastic band resistance (r_S=0.45, P=.04). No significant correlation was found for physical exercise with the silver band (r_S=0.356, P=.11).

Figure 3. Average heart rate while playing BoxVR depending on upper limb load. bpm: beats per minute; HR_{ave}: average heart rate.

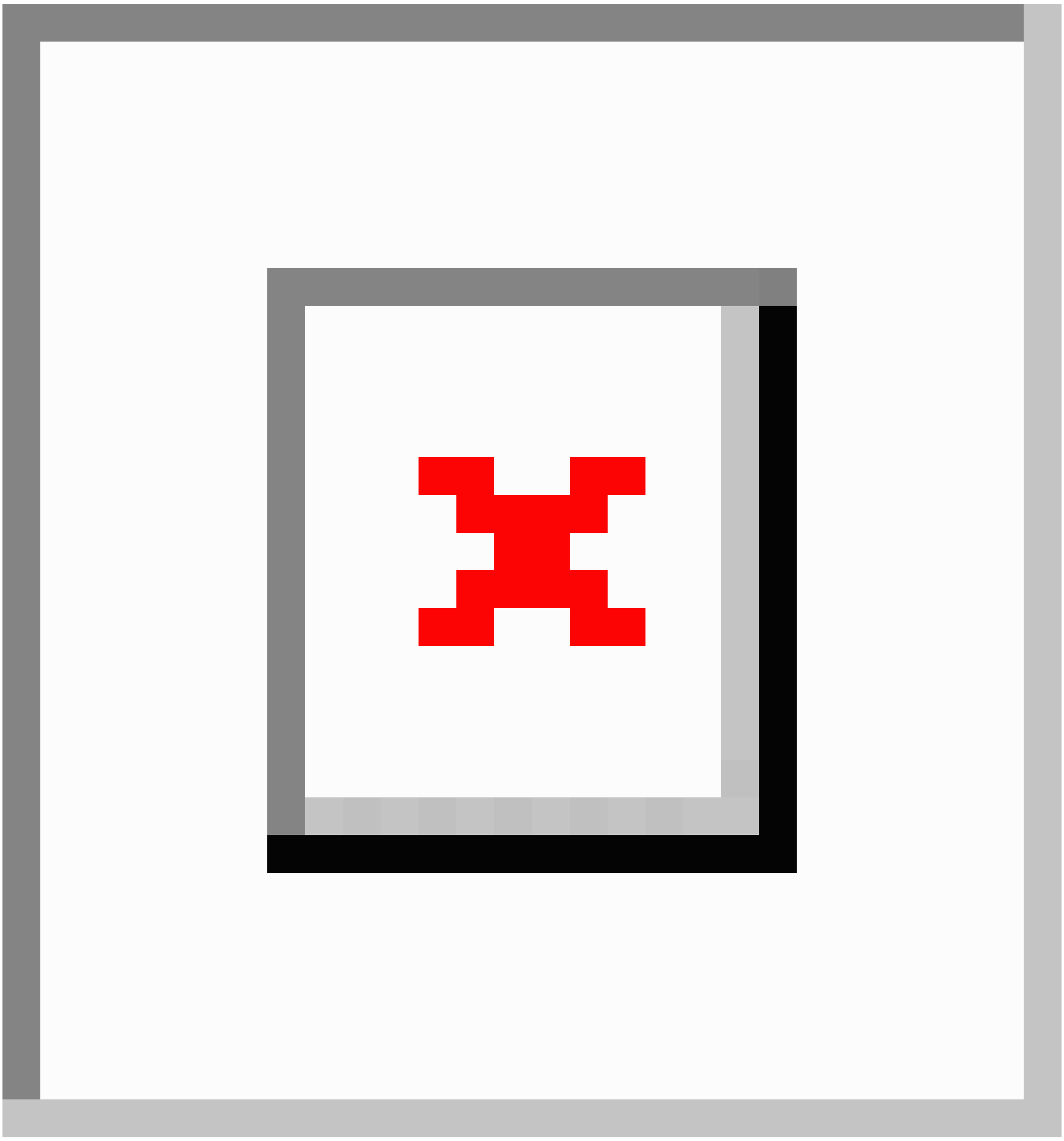


Figure 4. Intensity of physical exercise while playing BoxVR depending on upper limb load. % HR_{max}: percentage of maximum heart rate.

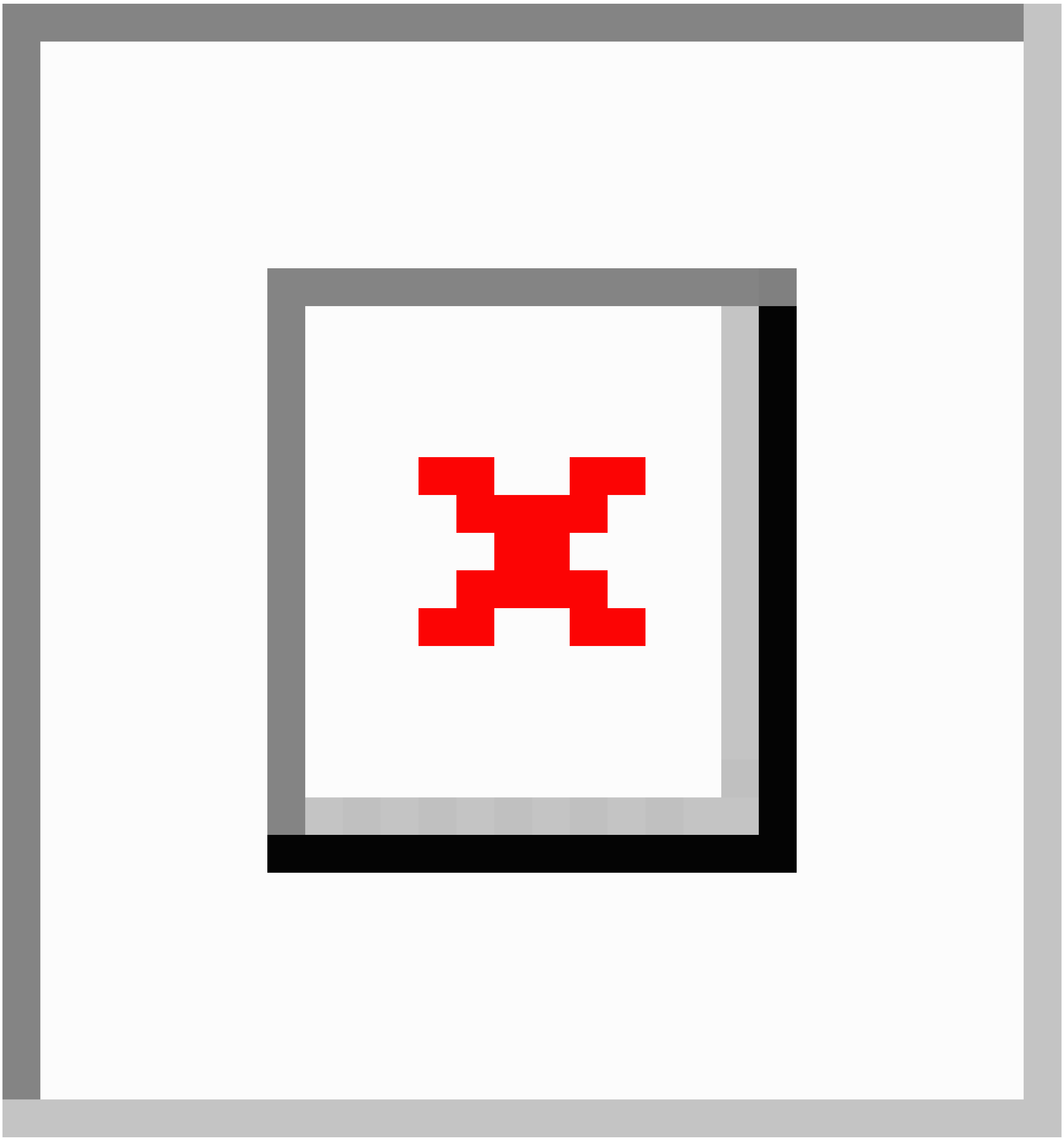


Figure 5. Average time spent in different heart rate zones by participants depending on upper limb load while playing BoxVR; HR_{max}: maximum heart rate.

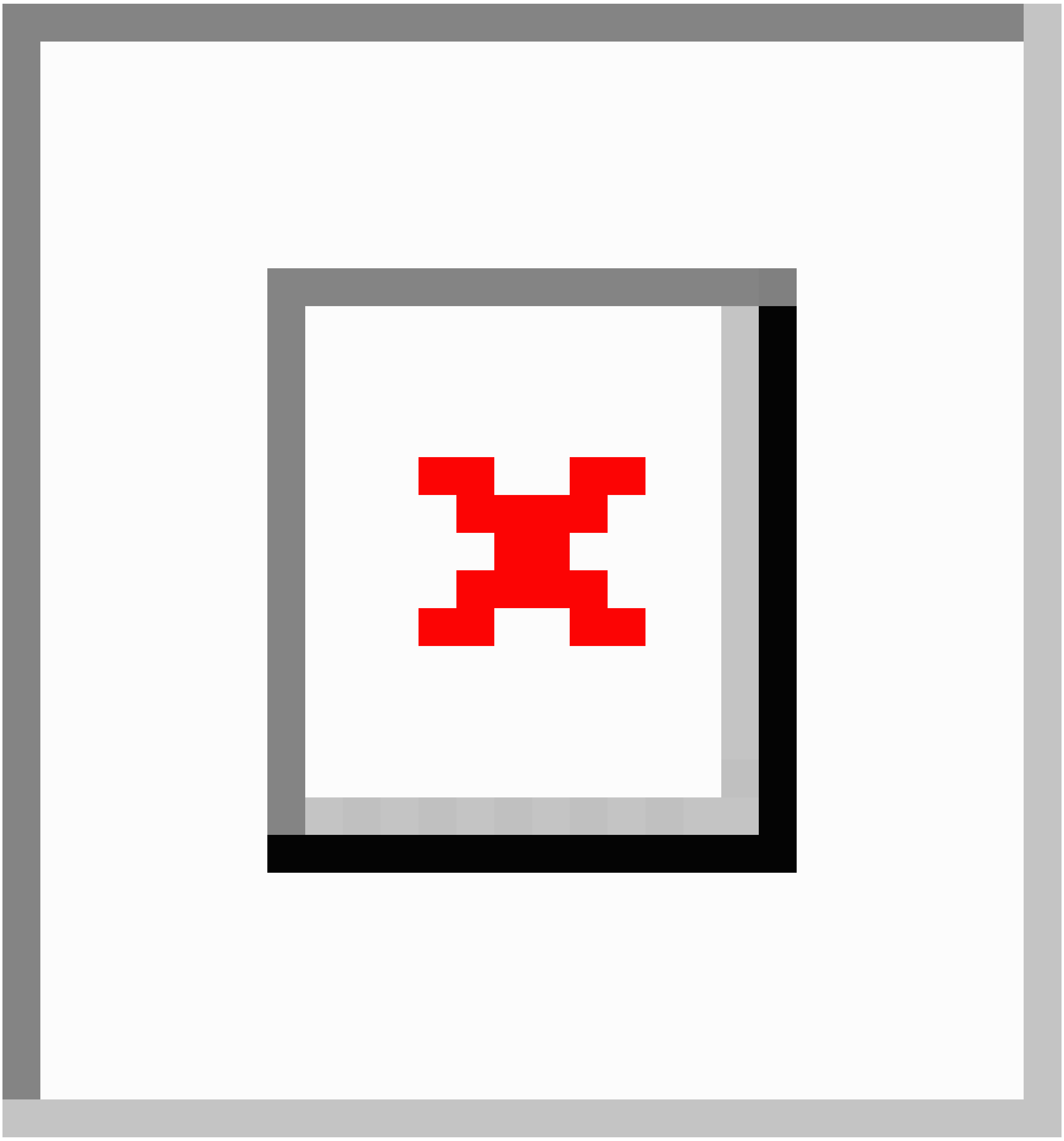
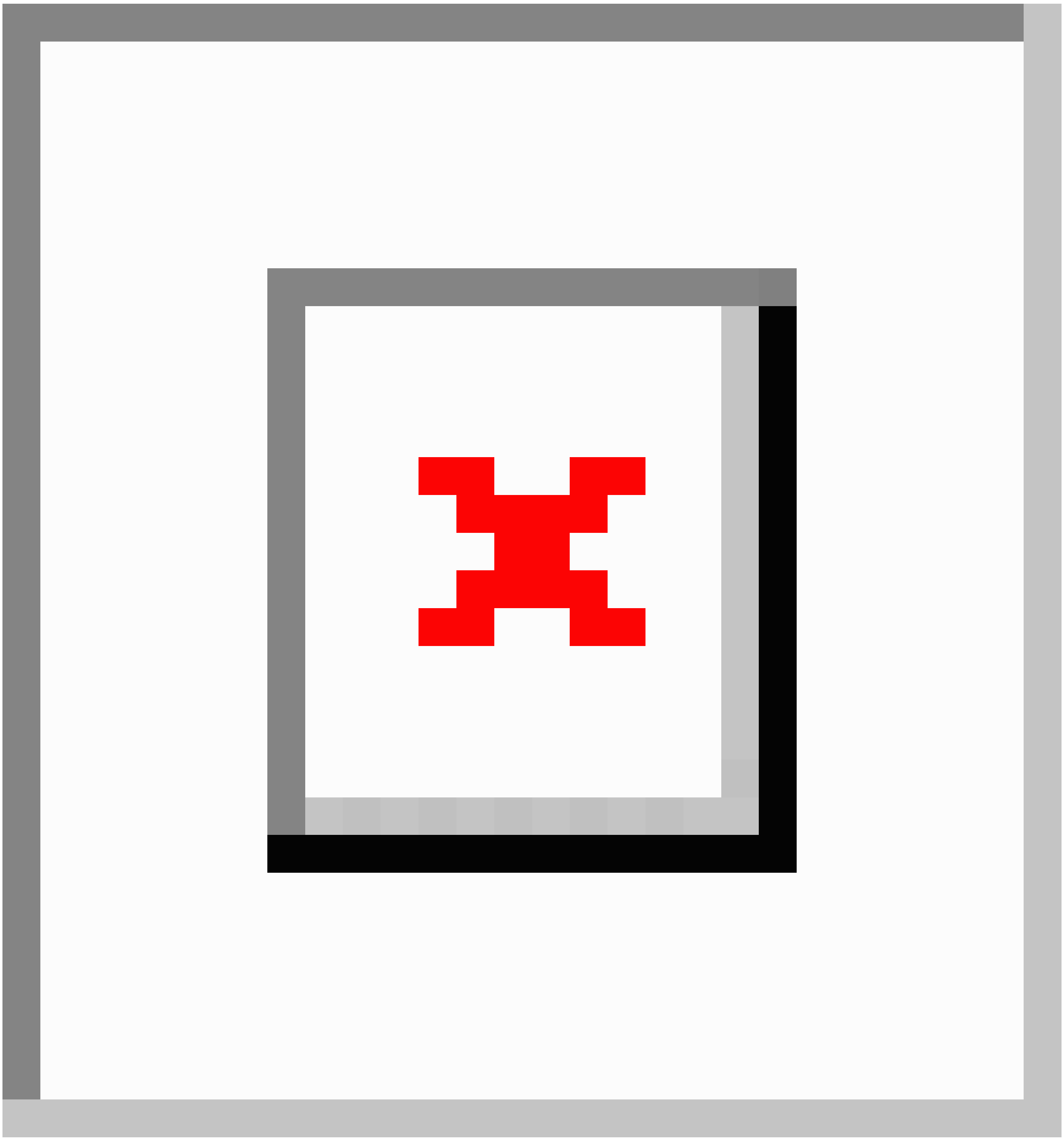


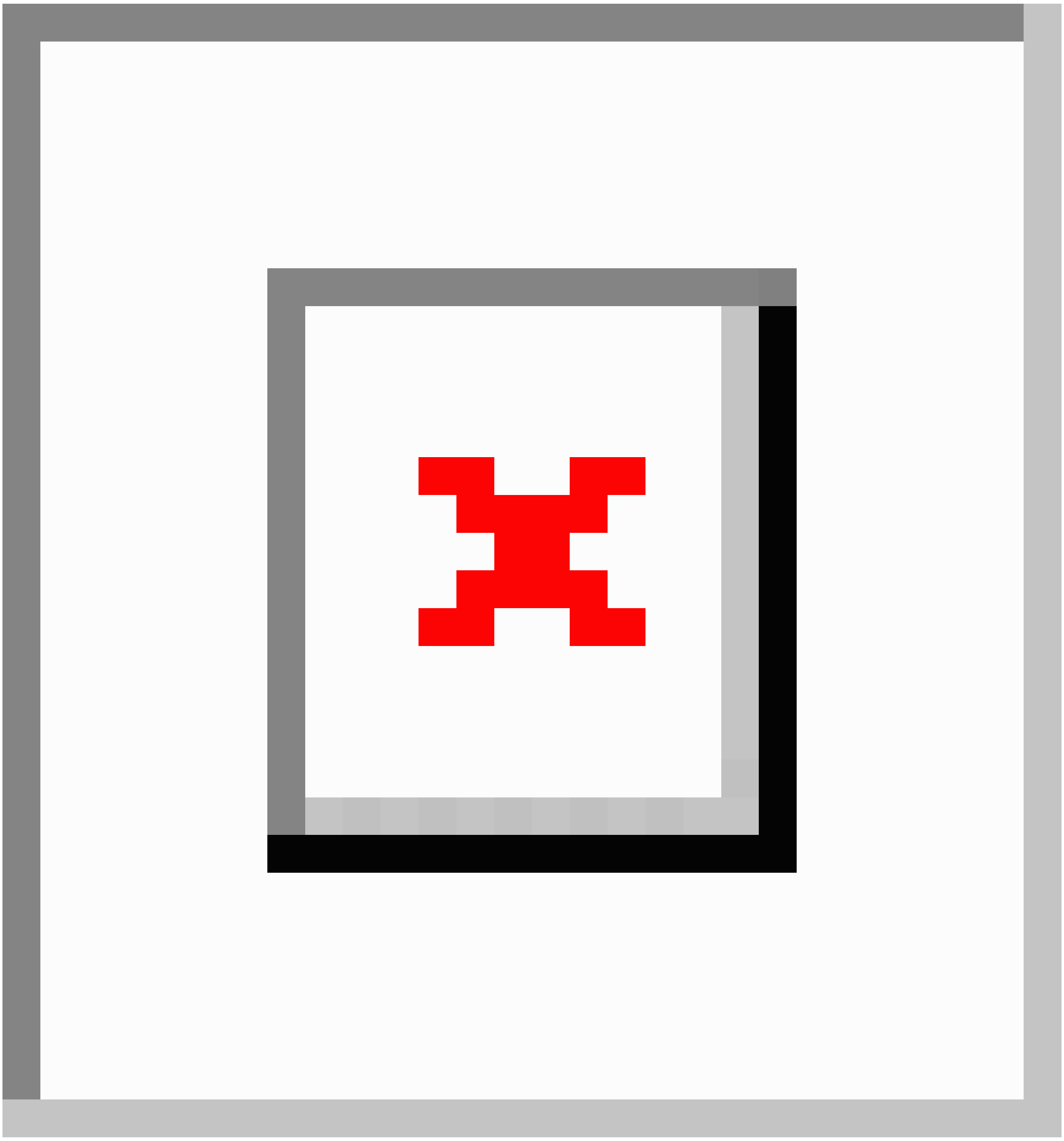
Figure 6. Rating of Perceived Exertion (RPE) scale versus upper limb load while playing BoxVR.



Satisfaction of Study Participants With Exercise in VR

The ANOVA of the PACES questionnaire results showed that additional elastic resistance significantly ($F_{2,40}$=9.259, P<.001; η^2p=0.316) influenced participants' perceptions of the attractiveness of exercise in VR. Study participants rated their satisfaction with PA in VR very similarly and highly for exercise without an upper limb load (mean 6.19, SD 0.61 points) and with elastic resistance in the form of a green band (mean 6.17, SD 0.66 points). The differences between these scores were minimal and statistically insignificant. Study participants were by far the least satisfied with PA in VR with the silver band (mean 5.66, SD 0.94 points). Therefore, statistically significant differences were found between the results of the assessment of exercises without additional resistance and with the silver band (P<.001) and between the assessment of exercises with the green and silver bands *(P*=.002) (Figure 7).

Figure 7. Satisfaction of participants with physical activity in virtual reality versus upper limb load. PACES: Physical Activity Enjoyment Scale.



Discussion

Principal Findings

This study found that the use of elastic bands while playing the AVRG BoxVR had a significant effect on exercise intensity, as shown by objective measurements and manifested by an increase in the heart rate. Furthermore, PA accompanying boxing exercises in VR, which was classified as low-intensity exercise, became moderately intense with additional resistance, and therefore became an exercise program that is considered beneficial for health according to WHO recommendations [13]. The heart rates remained in the high zones for a longer period of time during the resistance band exercises. Therefore, this type of shoulder loading during PA in VR appears to be an effective solution for increasing the intensity of physical exercise based on arm work.

The method presented in our study to increase PA intensity in VR by using resistance bands represents a novel solution. To date, weights attached to the distal part of the limbs have been used for this purpose. A recent study showed that the use of 0.5-kilogram Velcro-fastened HHWs placed on the wrists increased the intensity of PA in VR [10]. The authors found that under such upper limb loading, the PA intensity while playing the popular AVRG Beat Saber increased from low to moderate, thus becoming a healthy exercise. Similar studies were carried out using 2-kilogram ankle weights while playing an AVRG based on locomotor movements and practiced on an Omni omnidirectional treadmill (Virtuix) [11]. In this case, the additional load on the lower limbs was found to significantly increase physical exertion during virtual active entertainment. Recently, manufacturers of VR accessories have begun to view additional limb loading as a way to improve the effectiveness of exercise in a virtual environment. More recently, special controller overlays in the form of small discs have been offered for sale for users of the Oculus Quest 2 headset, which, when placed on the controllers, cause an increase in their weight to increase the intensity of exercise while using the application. As the use of resistance bands promotes increased PA intensity in arm-based VR, the use of elastic resistance offers an alternative to HHW or controller-mounted weights. However, there is currently a lack of such dedicated solutions for VR users.

The effect of elastic arm resistance on the intensity of physical exertion of users while playing BoxVR was also evidenced by the participant-reported RPE scores, which increased significantly after the use of elastic bands. Comparison of RPE reported by the participants with the objective classification of PA intensity [23] revealed that the students rated the PA without external loading as light and that with the green elastic band as moderate, which was similar to the objective assessment based on the heart rate monitor. In contrast, physical exercise performed with the silver band was rated as vigorous by the participants, indicating that the students overestimated its intensity in relation to objective measurements. This overestimation may be confirmed by the correlation analysis between subjective and objective measures of exercise intensity, showing a statistically significant relationship between RPE and $\%HR_{max}$ for PA in VR without a load and with the green elastic band resistance, while no significant relationship was found for exercise with the silver band. The exaggerated level of the subjective rating of PA during exercise in VR is somewhat puzzling, as previous studies have demonstrated that being in a VR environment reduces the intensity of perception of various stimuli (eg, pain) because VR, by stimulating different senses, distracts the immersed person from the problem [25-28]. During exercise in VR, this phenomenon, known as cognitive distraction, can alleviate the discomfort associated with hard training. The few studies on this topic published to date suggest that VR may be useful in distracting from unpleasant bodily sensations occurring during aerobic PA in children with overweight and obesity [9] and in reducing negative sensations associated with the performance of isometric exercises [29].

According to the PACES survey, study participants highly rated their satisfaction with PA in a virtual environment while playing BoxVR. Scores exceeded 6 on a 7-point scale for two measurements. Although the ANOVA of the PACES questionnaire results revealed that additional elastic resistance significantly affects the participants' perceptions of the attractiveness of exercise in VR, study participants rated their satisfaction with PA in VR very similarly for exercise without upper limb loading and with green resistance bands; the differences found were minimal and statistically insignificant. This may indicate that the low external load on the arms does not bother users and does not cause discomfort that could reduce the enjoyment of the exercises performed in VR. This was also confirmed by the aforementioned studies using a 0.5-kilogram HHW and a 2-kilogram ankle weight [10,11]. However, our results suggest that as the elastic external load on the arms increases, there may be a reduction in user satisfaction with PA in VR. Participants in this study were the least satisfied with playing BoxVR while having to overcome the resistance of the silver band, although a score of 5.66 still seems to be relatively high. The attractiveness of PA in VR has also been assessed in other contexts [10-12,30-36], and most of these studies have indicated a high level of user satisfaction with such exercises. Because those studies assessed other forms of PA or the attractiveness of physical exercise was measured with different tools, it is difficult to compare their results with those obtained in our study. Due to the rapid development of AVRGs and training applications used in a virtual environment, further research is warranted to identify the determinants of satisfaction of people participating in PA in VR. This will help guide the further development of this new form of exercise. Notably, satisfaction is an important motivation for undertaking regular healthy PA, and how people feel when they exercise determines their future training engagement [37]. Therefore, identifying user preferences for different forms of PA in VR can increase the likelihood of the regular active use of modern technology, which should translate into health benefits.

Limitations and Prospects

Despite these promising results, the solution we have presented has some limitations. Namely, for the user to perform the exercise with the resistance band attached to the ground, they must be looking forward and cannot move freely. Consequently, the use of such a solution is only possible for certain AVRGs. However, there is a way to address this limitation. There are wearable resistance band (WRB) systems (eg, WearBands or MASS Suit), using specially designed belts, socks, gloves, and other items of clothing to anchor elastic resistance bands connecting two or more body segments. WRB training is a new entry to the field of resistance training. With WRBs, the exerciser can move freely while performing movements with elastic resistance [38]. Despite the lack of research on exercise with WRB, it appears that this type of training system may be useful for increasing the intensity and effectiveness of exercise in VR. Testing this assumption may provide objectives for further empirical research. Currently, elastic resistance is being used in AVGs in nonimmersive VR. An example is the original pointing device created by Nintendo called Ring Fit, which works with the Ring Fit Adventure app. This is an elastic ring that can be squeezed, stretched, and moved in space, which enable controlling the movements of the virtual avatar. The few studies conducted to date have shown that Ring Fit exercises have a beneficial effect on students' physical fitness [39], reduce lower back pain in adults [40], improve balance in older people [41], and may be a useful form of PA for children with overweight and obesity by increasing their daily energy expenditure [42].

Conclusions

The intensity of PA among students playing the BoxVR game is at a relatively low level. With the added resistance of elastic bands attached to the upper limbs of the participants, the intensity of the exercise increased to a moderate level, as recommended for obtaining health benefits. Participants in this study highly rated the attractiveness of PA in VR. The use of slight elastic resistance did not negatively affect the satisfaction of study participants with the BoxVR game, although the satisfaction declined with a higher load.

Due to the rapid development of VR, the great popularity of games and training programs in a virtual environment, and their attractiveness to users, it is expected that more and more people will enjoy active entertainment in VR. Therefore, research should be undertaken to assess user preferences and seek solutions to increase the usefulness and effectiveness of this newly developed form of PA so that regular users can improve their physical fitness and reap the health benefits. Our study may provide guidance to VR equipment manufacturers on how to make exercise more effective while playing AVRGs based on upper limb movements. However, the validity of the considerations outlined above should be confirmed in further research using applications that allow various forms of PA to be practiced in an immersive VR environment.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Langer A, Polechoński J, Polechoński P, Cholewa J. Ruler drop method in virtual reality as an accurate and reliable tool for evaluation of reaction time of mixed martial artists. Sustainability 2023 Jan;15(1):648. [doi: 10.3390/su15010648]
2. Petri K, Emmermacher P, Danneberg M, et al. Training using virtual reality improves response behavior in karate kumite. Sports Eng 2019 Mar;22(1):2. [doi: 10.1007/s12283-019-0299-0]
3. Polechoński J, Langer A. Assessment of the relevance and reliability of reaction time tests performed in immersive virtual reality by mixed martial arts fighters. Sensors 2022 Jun 24;22(13):35808260. [doi: 10.3390/s22134762] [Medline: 35808260]
4. Qian J, McDonough DJ, Gao Z. The effectiveness of virtual reality exercise on individual's physiological, psychological and rehabilitative outcomes: a systematic review. Int J Environ Res Public Health 2020 Jun 10;17(11):4133. [doi: 10.3390/ijerph17114133] [Medline: 32531906]
5. Todorov K, Manolova A, Chervendinev G. Immersion in virtual reality video games for improving physical performance measures: a review. Presented at: 27th National Conference with International Participation (TELECOM); Oct 30 to 31, 2019; Sofia, Bulgaria. [doi: 10.1109/TELECOM48729.2019.8994884]
6. Jochymczyk-Woźniak K, Nowakowska K, Polechoński J, Sładczyk S, Michnik R. Physiological gait versus gait in VR on multidirectional treadmill-comparative analysis. Medicina 2019 Aug 22;55(9):517. [doi: 10.3390/medicina55090517] [Medline: 31443382]
7. Wodarski P, Jurkojć J, Bieniek A, et al. The analysis of the influence of virtual reality on parameters of gait on a treadmill according to adjusted and non-adjusted pace of the visual scenery. In: Pietka E, Badura P, Kawa J, Wieclawek W, editors. Information Technology in Biomedicine ITIB 2019. Advances in Intelligent Systems and Computing, Vol 1011 2019:543-553. [doi: 10.1007/978-3-030-23762-2_48]
8. Lu AS, Pelarski V, Alon D, et al. The effect of narrative element incorporation on physical activity and game experience in active and sedentary virtual reality games. Virtual Real 2023 Jan 31;27:1607-1622. [doi: 10.1007/s10055-023-00754-7] [Medline: 36742343]
9. Baños RM, Escobar P, Cebolla A, et al. Using virtual reality to distract overweight children from bodily sensations during exercise. Cyberpsychol Behav Soc Netw 2016 Feb;19(2):115-119. [doi: 10.1089/cyber.2015.0283] [Medline: 26882326]
10. Polechoński J, Zwierzchowska A, Makioła Ł, Groffik D, Kostorz K. Handheld weights as an effective and comfortable way to increase exercise intensity of physical activity in virtual reality: empirical study. JMIR Serious Games 2022 Nov 23;10(4):e39932. [doi: 10.2196/39932] [Medline: 36416873]
11. Polechoński J, Kostorz K, Polechoński P. Using ankle weights as an effective way to increase the intensity of physical activity while playing immersive virtual reality games on an omnidirectional treadmill. J Multidiscip Appl Nat Sci 2023 Jan;13(20):11536. [doi: 10.3390/app132011536]
12. Polechoński J, Szczechowicz B, Ryśnik J, Tomik R. Recreational cycling provides greater satisfaction and flow in an immersive virtual environment than in real life. BMC Sports Sci Med Rehabil 2024 Jan 30;16(1):31. [doi: 10.1186/s13102-024-00818-4] [Medline: 38291484]
13. More Physical Activity (RUN) WHO Team. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2024-07-04]
14. Ilahi BR, Okilanda A, Raibowo S, et al. The effect of resistance bands rubber spring exercise on the front kick speed of adolescent Pencak Silat women. Int J Hum Mov Sports Sci 2023 Apr;11(2):418-423. [doi: 10.13189/saj.2023.110219]

15. Narita Devi S, Fauzi F, Sukamti ER, et al. The effect of 8 weeks of training with resistance band on limb power of Taekwondo athletes. Int J Multidiscip Res Anal 2022 Dec 24;05(12). [doi: 10.47191/ijmra/v5-i12-27]
16. Novak D, Loncar I, Sinkovic F, Barbaros P, Milanovic L. Effects of plyometric training with resistance bands on neuromuscular characteristics in junior tennis players. Int J Environ Res Public Health 2023 Jan 7;20(2):1085. [doi: 10.3390/ijerph20021085] [Medline: 36673841]
17. Wangi SP, Tomoliyus T, Prayoga HD, Wijayanti NPN, Prabowo TA. The effect of 8 weeks of punch resistance band and dumbbell training on the arm power of ‘youth’ male boxers. Int J Phys Educ Sports Health 2023 Sep 1;10(5):299-304. [doi: 10.22271/kheljournal.2023.v10.i5e.3120]
18. BoxVR. STEAM. URL: https://store.steampowered.com/app/641960/BOXVR/ [accessed 2024-07-08]
19. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol 2001 Jan;37(1):153-156. [doi: 10.1016/S0735-1097(00)01054-8] [Medline: 11153730]
20. Riebe D, Ehrman JK, Liguori G, Magal M. ACSM’S Guidelines for Exercise Testing and Prescription: Wolters Kluwer; 2018.
21. Borg G. Perceived exertion as an indicator of somatic stress. Scand J Rehabil Med 1970;2(2):92-98. [Medline: 5523831]
22. Scherr J, Wolfarth B, Christle JW, Pressler A, Wagenpfeil S, Halle M. Associations between Borg’s rating of perceived exertion and physiological measures of exercise intensity. Eur J Appl Physiol 2013 Jan;113(1):147-155. [doi: 10.1007/s00421-012-2421-x] [Medline: 22615009]
23. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European guidelines on cardiovascular disease prevention in clinical practice. Atherosclerosis 2016 Sep;252:207-274. [doi: 10.1016/j.atherosclerosis.2016.05.037] [Medline: 22615009]
24. Mullen SP, Olson EA, Phillips SM, et al. Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time. Int J Behav Nutr Phys Act 2011 Sep 27;8:103. [doi: 10.1186/1479-5868-8-103] [Medline: 21951520]
25. Hoffman HG, Boe DA, Rombokas E, et al. Virtual reality hand therapy: a new tool for nonopioid analgesia for acute procedural pain, hand rehabilitation, and VR embodiment therapy for phantom limb pain. J Hand Ther 2020 Apr;33(2):254-262. [doi: 10.1016/j.jht.2020.04.001] [Medline: 32482376]
26. Jin W, Choo A, Gromala D, Shaw C, Squire P. A virtual reality game for chronic pain management: A randomized, controlled clinical study. Stud Health Technol Inform 2016 Jan;220:154-160. [Medline: 27046570]
27. Jones T, Moore T, Choo J. The impact of virtual reality on chronic pain. PLoS One 2016 Dec;11(12):e0167523. [doi: 10.1371/journal.pone.0167523] [Medline: 27997539]
28. Tashjian VC, Mosadeghi S, Howard AR, et al. Virtual reality for management of pain in hospitalized patients: results of a controlled trial. JMIR Ment Health 2017 Mar 29;4(1):e9. [doi: 10.2196/mental.7387] [Medline: 28356241]
29. Matsangidou M, Ang CS, Mauger AR, Intarasirisawat J, Otkhmezuri B, Avraamides MN. Is your virtual self as sensational as your real? Virtual reality: the effect of body consciousness on the experience of exercise sensations. Psychol Sport Exerc 2019 Mar;41:218-224. [doi: 10.1016/j.psychsport.2018.07.004]
30. Feodoroff B, Konstantinidis I, Froböse I. Effects of full body exergaming in virtual reality on cardiovascular and muscular parameters: cross-sectional experiment. JMIR Serious Games 2019 Aug 28;7(3):e12324. [doi: 10.2196/12324] [Medline: 31464194]
31. Gomez D, Browne JD, Almalouhi A, et al. Muscle activity during immersive virtual reality exergaming incorporating an adaptive cable resistance system. Int J Exerc Sci 2022 Jan;15(7):261-275. [Medline: 36896022]
32. Høeg ER, Bruun-Pedersen JR, Cheary S, et al. Buddy biking: a user study on social collaboration in a virtual reality exergame for rehabilitation. Virtual Reality 2023 Mar;27(1):245-262. [doi: 10.1007/s10055-021-00544-z]
33. Jones L, Ekkekakis P. Affect and prefrontal hemodynamics during exercise under immersive audiovisual stimulation: improving the experience of exercise for overweight adults. J Sport Health Sci 2019 Jul;8(4):325-338. [doi: 10.1016/j.jshs.2019.03.003] [Medline: 31333885]
34. McDonough DJ, Pope ZC, Zeng N, Liu W, Gao Z. Comparison of college students’ blood pressure, perceived exertion, and psychosocial outcomes during virtual reality, exergaming, and traditional exercise: an exploratory study. Games Health J 2020 Aug 1;9(4):290-296. [doi: 10.1089/g4h.2019.0196]
35. Plante TG, Cage C, Clements S, Stover A. Psychological benefits of exercise paired with virtual reality: outdoor exercise energizes whereas indoor virtual exercise relaxes. Int J Stress Manag 2006 Feb;13(1):108-117. [doi: 10.1037/1072-5245.13.1.108]
36. Zeng N, Pope Z, Gao Z. Acute effect of virtual reality exercise bike games on college students’ physiological and psychological outcomes. Cyberpsychol Behav Soc Netw 2017 Jul;20(7):453-457. [doi: 10.1089/cyber.2017.0042]
37. Williams DM, Dunsiger S, Ciccolo JT, Lewis BA, Albrecht AE, Marcus BH. Acute affective response to a moderate-intensity exercise stimulus predicts physical activity participation 6 and 12 months later. Psychol Sport Exerc 2008 May;9(3):231-245. [doi: 10.1016/j.psychsport.2007.04.002] [Medline: 18496608]
38. Shavelson D. Resistance training (RT): a meta-analysis of the existing EBM. EC Orthop 2018 Feb;9:434-457 [FREE Full text]

39. Wu YS, Wang WY, Chan TC, et al. Effect of the Nintendo Ring Fit Adventure exergame on running completion time and psychological factors among university students engaging in distance learning during the COVID-19 pandemic. JMIR Serious Games 2022 Mar 22;10(1):e35040. [doi: 10.2196/35040] [Medline: 35315780]
40. Sato T, Shimizu K, Shiko Y, et al. Effects of Nintendo Ring Fit Adventure exergame on pain and psychological factors in patients with chronic low back pain. Games Health J 2021 Jun;10(3):158-164. [doi: 10.1089/g4h.2020.0180] [Medline: 33891508]
41. Chan WLS, Chan CWL, Lam FMH, et al. Feasibility, safety, and effects of a Nintendo Ring Fit Adventure balance and strengthening exercise program in community-dwelling older adults with a history of falls: a feasibility randomized controlled trial. Geriatr Gerontol Int 2024 Mar;24 Suppl 1:334-341. [doi: 10.1111/ggi.14771] [Medline: 38088479]
42. Comeras-Chueca C, Villalba-Heredia L, Pérez-Llera M, et al. Assessment of active video games' energy expenditure in children with overweight and obesity and differences by gender. Int J Environ Res Public Health 2020 Sep 15;17(18):32942663. [doi: 10.3390/ijerph17186714] [Medline: 32942663]

Abbreviations

%HR_{max}: percentage of maximal heart rate
AVG: active video game
AVRG: active virtual reality game
bpm: beats per minute
HHW: handheld weight
HR_{ave}: average heart rate
HR_{max}: maximum heart rate
PA: physical activity
PACES: Physical Activity Enjoyment Scale
RPE: Rating of Perceived Exertion
VR: virtual reality
WHO: World Health Organization
WRB: wearable resistance band

Edited by R Baranyi; submitted 14.03.24; peer-reviewed by B Szczechowicz, KE Naugle; revised version received 19.05.24; accepted 04.06.24; published 16.07.24.

Please cite as:
Polechoński J, Przepiórzyński A, Polechoński P, Tomik R
Effect of Elastic Resistance on Exercise Intensity and User Satisfaction While Playing the Active Video Game BoxVR in Immersive Virtual Reality: Empirical Study
JMIR Serious Games 2024;12:e58411
URL: https://games.jmir.org/2024/1/e58411
doi:10.2196/58411

The Effect of a Novel Video Game on Young Soccer Players' Sports Performance and Attention: Randomized Controlled Trial

Adrian Feria-Madueño[1], PhD; Germán Monterrubio-Fernández[2], BS; Jesus Mateo Cortes[3,*], BS; Angel Carnero-Diaz[1,*], BA

1

2

3

*these authors contributed equally

Corresponding Author:
Angel Carnero-Diaz, BA

Abstract

Background: Currently, the fusion of technology and sports is inevitable. The integration of various systems and devices has brought about significant transformations in established sports practices, impacting not only the rules but also physiological, biomechanical, and even psychological aspects.

Objective: The purpose of this study was to analyze the effect of an attention intervention through a video game on young soccer players.

Methods: Twelve young male soccer players (age: mean 8.5, SD 1 years) were divided into 2 groups: a control group (CG; n=10) and an experimental group (EG; n=10). During the 6-week training program, the EG received attention training through a video game twice a week for 15 minutes per session. Pre- and postintervention measurements included a specific decision-making soccer test and interviews with coaching staff. Additionally, success in the video game, muscular activity, and sweat levels were monitored.

Results: The EG demonstrated a significant improvement in video game success following the intervention program, as indicated by the achieved level (*P*<.001). However, no significant differences were found between groups regarding electromyographic (EMG) activity (*P*=.21) and sweating (*P*=.20). Prior to implementing the attention training program, both groups exhibited similar data for variables related to decision-making and execution mechanisms (≤10%). Only 2 decision-making variables exceeded 10% but remained below 15% (Shot_D=13.35%; Marking_with_Ball_D=−12.64%). Furthermore, changes in attacking action variables were more pronounced in execution-related variables, except for dribbling and fixing. Conversely, in defensive action variables, changes were greater in decision-related variables, except for marking with the ball and marking without the ball.

Conclusions: Our findings reveal that incorporating a specific attentional video game into a soccer training program enhances decision-making compared to a program without the video game. Therefore, it is advisable for practitioners to consider using this tool due to its high efficiency in terms of economic and temporal costs, particularly in improving a key psychological variable.

Trial Registration: ISRCTN Registry ISRCTN12742775; https://www.isrctn.com/ISRCTN12742775

(JMIR Serious Games 2024;12:e52275) doi:10.2196/52275

KEYWORDS

reaction time; serious games; executive function; decision making; game; games; gaming; sport; sports; soccer; football; athlete; athletes; athletic; training; performance; physiological; muscle; muscular; sweat; sweating; attention training; attentional; ball; exercise; physical activity; exergame; exergames; interview; interviews

Introduction

Currently, technology use is inseparable from sports. The integration of various systems and devices has brought about significant transformations in established sports practices, impacting not only the rules but also physiological, biomechanical, and even psychological aspects. This revolution is particularly evident in soccer, where the implementation of technology has primarily concentrated on the professional sphere. From a physiological perspective, technology has facilitated notable advancements in the understanding and control of physical demands during elite soccer matches. Tools such as the GPS [1,2], accelerometry, and specialized match analysis programs [2] have made it possible for coaches to design training tasks on the basis of data obtained from actual games. Optical tracking systems and GPS devices were compared in professional soccer, revealing no significant differences in analyzed variables such as total distance, distance per minute, average speed, and maximum speed [3]. However, limitations were identified, such as the ineffectiveness of the GPS in indoor sports and the inability of optical tracking systems to access internal variables. Nevertheless, researchers concluded that both technologies were suitable for monitoring the physiological demands of soccer players [3].

Technological applications in soccer have also shed light on performance mechanisms in areas such as jumping [4] and the minimization of lower limb injuries [5]. Usually, these technologies rely on 3D movement analysis, muscle activation assessment, and force production evaluation. Nonetheless, the high economic cost and specialized expertise required for their use pose significant challenges, making it difficult to implement them in the daily lives of athletes. Another way to harness the use of technology with a better cost-benefit relationship is through the use of video games; in particular, those with a clear aim to improve performance and learning, known as serious games [6]. Games have been used to introduce challenging concepts or develop skills in different areas such as surgery [7] and rehabilitation [8] or to enhance team coordination via cooperative training—the latter being called “small side games” [9,10]. All manner of play seeks to achieve improvements in skill acquisition in a more attractive way over and above traditional forms [11,12]; in particular, among pediatric populations with special attention and motivation characteristics [13,14]. Playful forms of intervention have improved affective responses in children compared to traditional teaching methodologies [15], having achieved benefits with the use of video games that optimize performance in variables as attention and executive control [16] and other psychological variables [17]. Serious games are emerging as valuable tools with positive impacts on cognitive aspects, providing specific value to their users. A pertinent study demonstrated that the implementation of a serious game based on chess, titled “The Secret Trail of Moon,” led to notable increases in emotional control levels and a reduction in attention deficits among participants [18]. In the sports domain, literature also highlights benefits, with a serious virtual reality game showing a significant improvement in players’ concentration during skiing tasks [13]. These findings underscore the potential positive impact of serious games on diverse cognitive and sports-related domains. Studies have increasingly focused on the psychological and cognitive aspects of athletes over the years [19,20], having highlighted attention, memory, and motivation as crucial psychological variables for athletes [21,22]. While the relationship among technology, games, and the psychological behavior of athletes has been explored, research specifically focusing on technology training in soccer remains scarce compared with other areas of knowledge. A notable study [23] revealed that young soccer players who trained with a computerized attention training system called Rejilla (version 1.0) [24] experienced significant improvements in their attention. However, these improvements were not directly linked to on-field performance variables. The literature has primarily centered around elite soccer, with limited attention given to youth soccer. Providing information to youth soccer coaches regarding how technology can impact the psychological and cognitive variables of their athletes will help optimize performance and achieve long-term success.

Methods

Participants

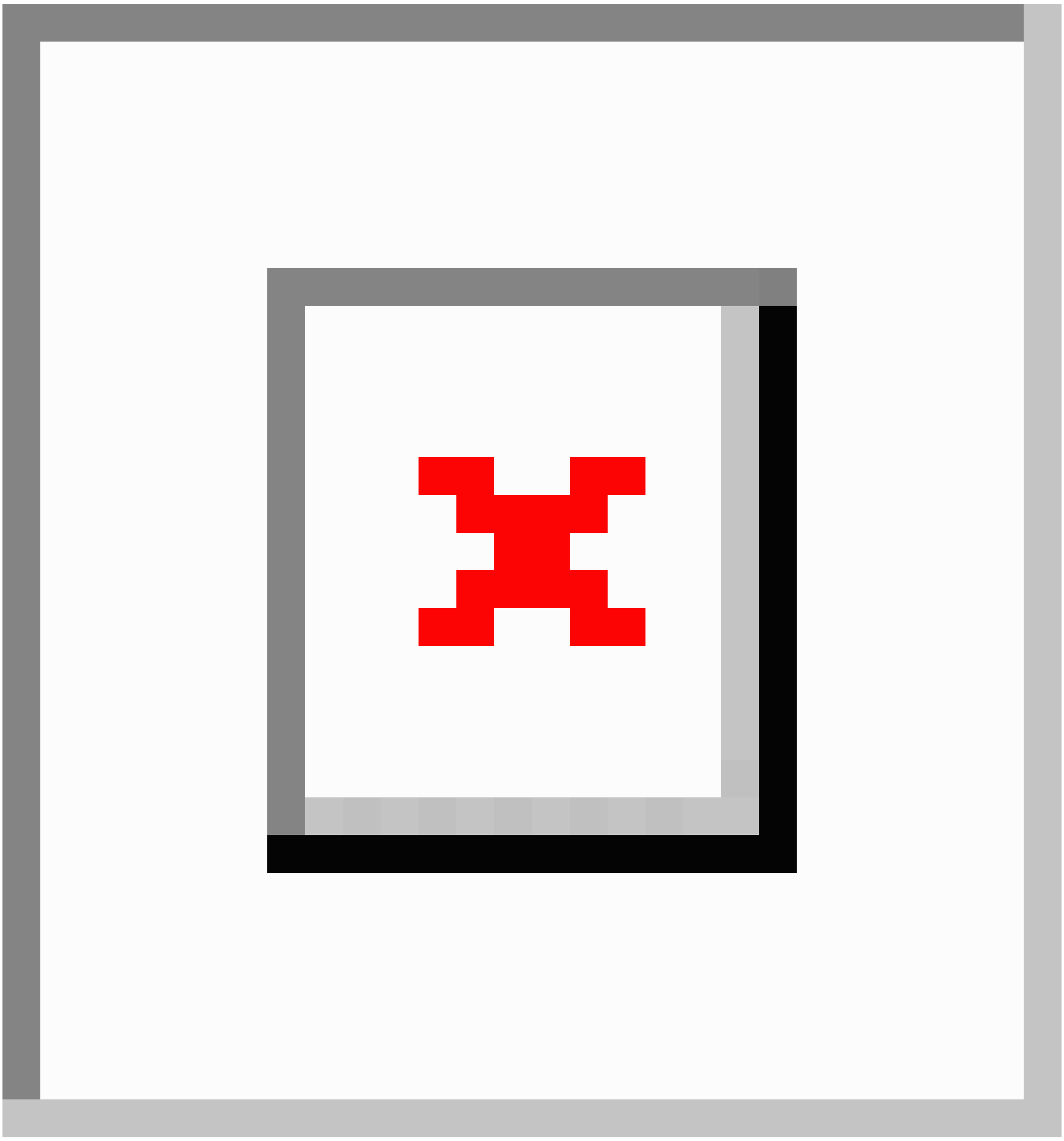
The study involved 12 young male soccer players aged between 8 and 9 years from a club in Seville, Spain. The participants underwent training sessions 3 days a week, each lasting approximately 120 minutes. The sample was divided into 2 groups: a control group (CG) and an experimental group (EG). Both groups were informed about their participation in a program aimed at enhancing their soccer skills. For the CG, the training involved watching videos of goals scored by the first team in previous seasons, followed by researchers posing questions to elicit responses. In search of a placebo effect, the EG was informed that the activity they were engaging in was attention training. This involved the visualization of videos showcasing goals scored by the first team in past seasons. After the video session, participants were required to answer questions related to the content (eg, “how many goals did the player with the number 9 score?”). Alternatively, the EG engaged in attention training using a video game.

Sample size calculation was convenience-based. For this, the study participants were players from a youth team belonging to a top-tier club in the Spanish Professional Football League. As an inclusion criterion, participants had to be 8 or 9 years old. Additionally, they should have been free from injuries that would have prevented them from participating in training or competitions. Their required training frequency was 3 times per week. During the 6 weeks, if any player reduced their training frequency to once a week, they were withdrawn from the study. Finally, a simple randomization method was chosen to randomize the groups, using the toss of a coin.

Procedure

The study spanned a duration of 6 weeks (Figure 1). Initially, all players underwent an evaluation of their decision-making skills in relation to soccer performance using the Game Performance Evaluation Tool (GPET) test [25]. The GPET is a tool to assess performance in invasion sports, specifically in soccer. Subsequently, both groups were assessed on the basis of their success in the video game, electromyographic (EMG) activity during the test, and sweat level. Throughout the study, all participants maintained their regular training routine, consisting of 3 sessions per week and a competitive game. The intervention varied, in that, on the one hand, the CG attended a room session twice a week where videos showcasing goals scored by the first team in past seasons were shown during 15 minutes, and after each video, the investigators posed a question to elicit a response; on the other hand, the EG attended a room session where attention training was conducted using a video game, with sessions lasting 15 minutes and taking place twice a week. Upon completion of the 6-week intervention, all players were reassessed regarding their decision-making abilities in relation to soccer performance using the GPET, as well as their performance in the video game, EMG activity, and sweat level during video game practice. Finally, an interview was conducted with the coaching staff to gauge the subjective attention levels of each soccer player, assigning them a score ranging from 1 (very low attentional level) to 3 (optimal attentional level) in relation to competitive situations.

Figure 1. Study design. Left to right: participants were randomized into 2 groups. Subsequently, all soccer players were assessed using the Game Performance Evaluation Tool (GPET), in addition to being evaluated for success in the video game, electromyographic (EMG) activity, and sweating levels. Afterward, the intervention proceeded for 6 weeks. Thereafter, both groups were evaluated again with the GPET, EMG activity, and sweating levels. Finally, the coaching staff was interviewed.



Instruments

The Video Game: BallApp

The video game used in the study involved the task of memorizing a ball with a distinct color among others displayed on the screen. After a period of 5 seconds, all the balls would change to the same color and move randomly across the screen. The participants were then required to identify the ball that initially had a different color. Accomplishments in the video game were measured in both groups, with the criterion being the ability to advance to the next level 3 consecutive times to mitigate the influence of luck. As the levels progressed, the speed of the balls, the number of selectable balls, and the presence of distractors such as noise increased. To replicate the environment of competitive soccer matches, specific background noise recordings from the soccer players' actual matches were provided.

The attention game for soccer players, BallApp, is designed to enhance their attention processes, aiming for a positive impact on attentional mechanisms when playing soccer. The game has been developed as an application using various technologies, including HTML for structural aspects, JavaScript for functionality, and Material Design for Bootstrap for design. The game is tailored for young soccer players and has been designed for individual play, although multiple players can participate on multiple screens.

The game involves selecting a ball of a different color than the others, and after a few seconds, all the balls turn into the same color and move randomly across the screen. The speed of movement, number of balls, depth of movement, and added noise vary as the difficulty levels progress. For example, in level 1, three balls appeared (one of them being of a different color), and the task was to choose 1 ball. All the balls moved without depth, and the speed was low. As levels increased, the difficulty level also increased on the basis of the number of balls on the screen. In subsequent levels, such as level 3, four balls appeared (one of them being of a different color), and the task was to choose 1 ball. The balls moved without depth, at a low speed, and without added noise. In total, the maximum number of programmed levels in the game was 8.

Level 4 onward, an ambient sound replicating those in real-world soccer matches was introduced. The aggressiveness of the noise increased with an increase in the difficulty level (Table 1). The aggressiveness of the noise was described on the basis of whether it was a murmur (low noise), cheering noise (medium noise), cheering noise with chants (moderate noise), noise with chants where specific phrases from a fan stood out (aggressive noise), or noise with chants and synchronized disturbances in the form of complaints or disgust toward plays (very aggressive noise).

Table . Overview of the characteristics of each level of the game.

Level	Number of balls on the display	Number of eligible balls	Depth of the movement	Speed of the balls	Sound
1	3	1	No depth	Low	Noiseless
2	3	1	With depth	Low	Noiseless
3	4	1	No depth	Low	Noiseless
4	5	1	No depth	Moderate	Slight noise
5	6	2	No depth	Moderate	Medium noise
6	6	2	With depth	Moderate	Moderate noise
7	7	2	No depth	High	Loud noise
8	8	2	With depth	High	Very loud noise

Regardless, to progress to a higher level, the player had to win the played level 3 times consecutively, eliminating luck as a factor for level progression. If the player correctly identified the balls of a different color in each level, they were awarded a score. This score depended on not only accuracy but also the time taken to make their choice.

The game has a configuration for individual use, although competitions can also be held with other players. Thus, it provides players with a leaderboard that compares scores among players, and this ranking changes on the basis of the following variables: the number of games played, the number of points achieved per game, and the number of levels surpassed by the player.

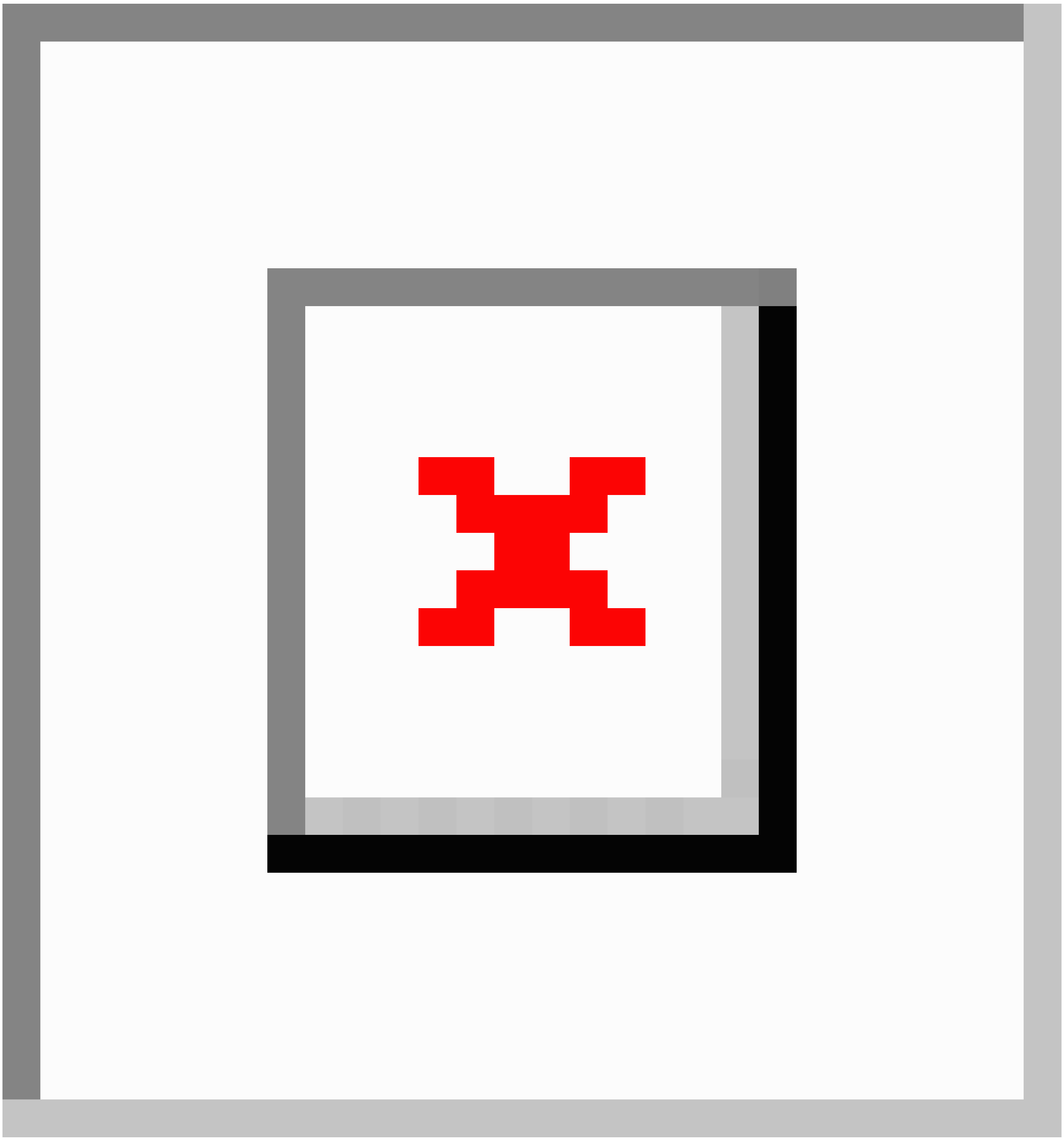
The selection of this particular video game for implementation is predicated on its consistent engagement of the attentional mechanism. This mechanism holds substantial implications in the developmental processes of young soccer players and, consequently, may yield positive effects on targeted soccer-related tasks.

EMG and Skin Conductance

EMG activity and electrodermal activity (EDA) were assessed on the basis of skin resistance and sweat production (bioPLUX). The EMG activity and EDA of the dominant hand's extensor muscle were evaluated in all participants. Electrodes were applied to the dominant hand, and these variables were analyzed before and after the intervention program during video game practice. These variables were examined from a somatic perspective, allowing for the monitoring of psychophysiological indicators during an attention-demanding task such as playing a video game (Figure 2).

The assessment of somatic variables of anxiety using biofeedback devices appears to be an excellent tool for controlling anxiety during different tests and among different populations [26], specifically in the analysis of EMG activity [27] and EDA variables [28], as specific control of somatic responses seems to be a good indicator of psychological states. Alterations in these variables could indicate somatic states and anxiogenic responses that would hinder the individuals' task performance [29]. Our study monitored these 2 responses while participants played the video game.

Figure 2. Evaluation of electromyographic activity (**A**) and electrodermal activity (**B**) during video game practice.



GPET

The GPET assessment tool was used to evaluate decision-making in relation to soccer performance. This instrument differentiates the cognitive-decisional aspect of performance from execution. The playing field has dimensions of 20×10 m^2, with delimited areas of 3×4 m^2 and goals measuring 0.95×0.70 m^2. The ball used is an A-7 soccer ball, with a circumference ranging between 0.635 and 0.66 m. The players engage in the task in teams consisting of 2 players each, with two 4-minute halves and a 3-minute break between them. The timer remains active throughout, and there are assistants designated to retrieve the ball. The variables captured by the test are assessed in both attacking and defensive situations, both with and without the ball (as outlined in Table 2). The test is recorded using a video camera for subsequent analysis. Using an observation sheet, successful actions are coded on the basis of the predefined roles into 4 options: appropriate decision (coded as 1), inappropriate decision (coded as 2), successful execution (coded as 1), and unsuccessful execution (coded as 2).

Table . Description of variables evaluated using the Game Performance Evaluation Tool.

	With the ball	Without the ball
Attack	• Control • Pass • Dribbling • Shot	• Losing one's defender • Fixing
Defense	• Marking with the ball • Defensive blocking • Tackle • Clearing with the ball	• Marking without the ball • Interception • Clearing without the ball

Outcome Measures

All variables evaluated using the GPET are described below and are classified in accordance with two criteria (Table 2): (1) whether the action is carried out with or without the ball and (2) if the actions are those wherein the decisional or execution mechanism predominates. For this purpose, all acronyms used have been "_with_ball" for variables with the ball, "_without_ball" for variables without the ball, "_D" for variables where the decisional mechanism predominates, and "_E" where the execution mechanism predominates. Thus, the variables are the following: (1) control: this refers to situations where the player keeps the ball under control; (2) pass: this is the action where the player successfully makes a pass to another teammate; (3) dribbling: this is an action where the player gets away from his opponent, eliminating any possibility of the ball being taken away from him; (4) shot: the player takes a shot at goal with the aim of scoring; (5) Losing_one's_defender: the player, by means of an attacking action, unmasks themself from their defender; (6) fixing: the player, in an attacking action, fixes the defender, forcing the latter to occupy a specific position on the field; (7) Marking_with_ball: in a defensive action, the player marks the player who has the ball; (8) Defensive_blockin: in a defensive action, the player tackles the player with the ball but does not fall to the floor; (9) tackle: in a defensive action, the player falls to the ground to snatch the ball from the opponent; (10) Clearing_with_Ball: in a defensive action, the player clears the player with the ball, using a free space; (11) Making_without_ball: in a defensive action, the player marks another player of the opposing team who is not in possession of the ball; (12) interception: in a defensive action, the player cuts off a play, preventing the ball from reaching its receiver; and (13) Clearing_without_ball: in a defensive action, the player makes a clearance to the player who does not have the ball, using a free space.

Interview

Traditionally, the coach's role as a key element in understanding the athlete has proven to be crucial [30]. Specifically, the athletes' psychological factors have been tested using various procedures, with the coach playing a pivotal role as a key to understanding the athlete's psychological status [31,32]. Thus, the qualitative insights obtained from interviews with coaches regarding the psychological variables of athletes can be a crucial element that significantly enhances the understanding of these psychological variables, aiding in decision-making about athletes [33].

An individual interview was conducted with the coach to subjectively assess the level of attention concerning decision-making and athletic performance for each soccer player. The coach was asked to rate, using a scale from 1 to 10, the level of attention exhibited by the soccer player in making optimal decisions that contribute to enhanced athletic performance. The coach provided a numerical rating, which was duly recorded by the evaluation team.

Statistical Analysis

A statistical analysis was conducted on the data obtained for each variable. Initially, a parametricity analysis was performed for each variable using the Shapiro-Wilk test. Subsequently, the mean and SD of each variable were assessed. To compare the means between groups, an independent samples *t* test was used, with a significance level set at 95% confidence ($P \leq .05$). Moreover, the effect size for the variables was determined using a 95% confidence limit. To qualitatively evaluate the potential quantitative changes observed after the program, the following categories were used [34]: highly improbable (<1%), very unlikely (1%-5%), unlikely (5%-25%), possible (25%-75%), likely (75%-95%), highly likely (95%-99%), and virtually certain (>99%).

Ethical Considerations

This study received ethical approval from the Ethics Committee of the Center for University Studies affiliated with the University of Seville (20190215) on February 5, 2019. This study complied with the ethical guarantees and requirements for experimentation on human beings and animals and the requirements established in Spanish legislation in the field of biomedical research, protection of personal data, and bioethics, and adhered to the fundamental principles of the Declaration of Helsinki and the European Convention on Human Rights. Informed consent was obtained from the responsible tutors of the players.

Results

The aim of this investigation is to assess the impact of an attention intervention among soccer players, using a video game. To assess the video game's impact on the athletic performance of these players, an analysis of various variables was conducted, categorized into 3 main groups: the first group focused on the video game itself, examining the success rate in the game, EMG activity, and EDA; the second group encompassed decision-making variables related to the athletic performance component in soccer players; and in the third group, a subjective evaluation of each soccer player's level of attention was obtained from their coach.

Regarding the variables analyzed in the video game, a significant improvement in the success rate of the soccer task was observed in the EG following the intervention program, as indicated by the level achieved. Both groups exhibited no differences in terms of success rate before undergoing training using the video game. However, after the 6-week duration, the EG demonstrated a substantial increase in the obtained score compared to the CG (Table 3).

Our findings indicate that there were no significant differences in EMG activity and assessed sweat between the 2 groups (Table 4).

Regarding the set of variables assessed using the GPET tool, both groups underwent evaluation before and after the implementation of the attention training program using the video game. The variables were categorized as successful or failed actions based on the prevailing decision or execution mechanism, and they were further classified as attacking or defensive actions. All data were analyzed and are presented as percentages. Furthermore, the change in percentages for each variable was calculated by subtracting the values to determine both inter- and intragroup differences, and these differences were evaluated. See Figures 3 and 4.

Regarding the mean score, a significant difference was observed between the EG and CG after the intervention period (pre-CG mean 46.2, SD 12.5; pre-EG mean 46.9, SD 13.8; P=.95; Cohen d=−0.01; post-CG mean 40.7, SD 13.7; post-EG mean 59.7, SD 17.3; P=.01; Cohen d=−1.73). As indicated, before the application of the attention training program using the video game, both groups showed relatively similar variables in which the decision mechanism predominated, as well as in those in which the execution mechanism predominated, the differences between both groups being less than 10%. Only 2 decision mechanism variables had values above 10% but never exceeding 15% (Shot_D=13/100, 13%; Marking_with_Ball_D=12/100, −12%).

Regarding the EG, following a 6-week intervention, all variables underwent some degree of modification, with minor changes below 15%. The variables that experienced the most substantial changes were Fixing_D (27/100, −27%), Fixing_E (37/100, −37%), Defensive_blocking_D (28/100, −28%), Marking_with_Ball_D (22/100, −22%), and Interception_D (18/100, −18%). The minus sign in these changes indicates that the percentage for each variable decreased after the intervention period. In the case of the CG, 9 variables showed changes exceeding 15%, with 5 of them surpassing 30%. On comparing the percentage changes between the 2 groups after the intervention, all variables had higher values in the EG than in the CG. The only variable that had higher values in the CG was the shot, on evaluation of both the decision mechanism (Shot_D=−1.43) and the execution mechanism (Shot_E=−1.69). However, the differences observed in both cases did not exceed 2 units.

Table . Comparison of success in the video game between groups before and after the intervention.

Variable	Before[a]		After[b]	
	CG[c]	EG[d]	CG	EG
Success, mean (SD)	1.66 (0.51)	1.5 (0.54)	1.83 (0.40)	4.16 (0.75)

[a]P=.61.
[b]P<.001.
[c]GC: control group.
[d]EG: experimental group.

Table . Comparison of EMG[a] activity and EDA[b] between groups before and after the intervention.

Variable	Before			After		
	CG[c]	EG[d]	P value[e]	CG	EG	P value
EMG activity (μV), RMS[f] (SD)	1.66 (0.10)	1.71 (0.09)	.50	1.78 (0.12)	1.72 (0.05)	.21
EDA (μΩ), RMS, (SD)	25.11 (0.07)	25.81 (0.88)	.08	25.51 (0.48)	25.01 (0.85)	.20

[a]EMG: electromyographic.
[b]EDA: electrodermal activity.
[c]CG: control group.
[d]EG: experimental group.
[e]P≤.05 was considered significant.
[f]RMS: root-mean-square.

Figure 3. Comparison between the groups before the training period. CG: control group; EG: experimental group; GPET: Game Performance Evaluation Tool.

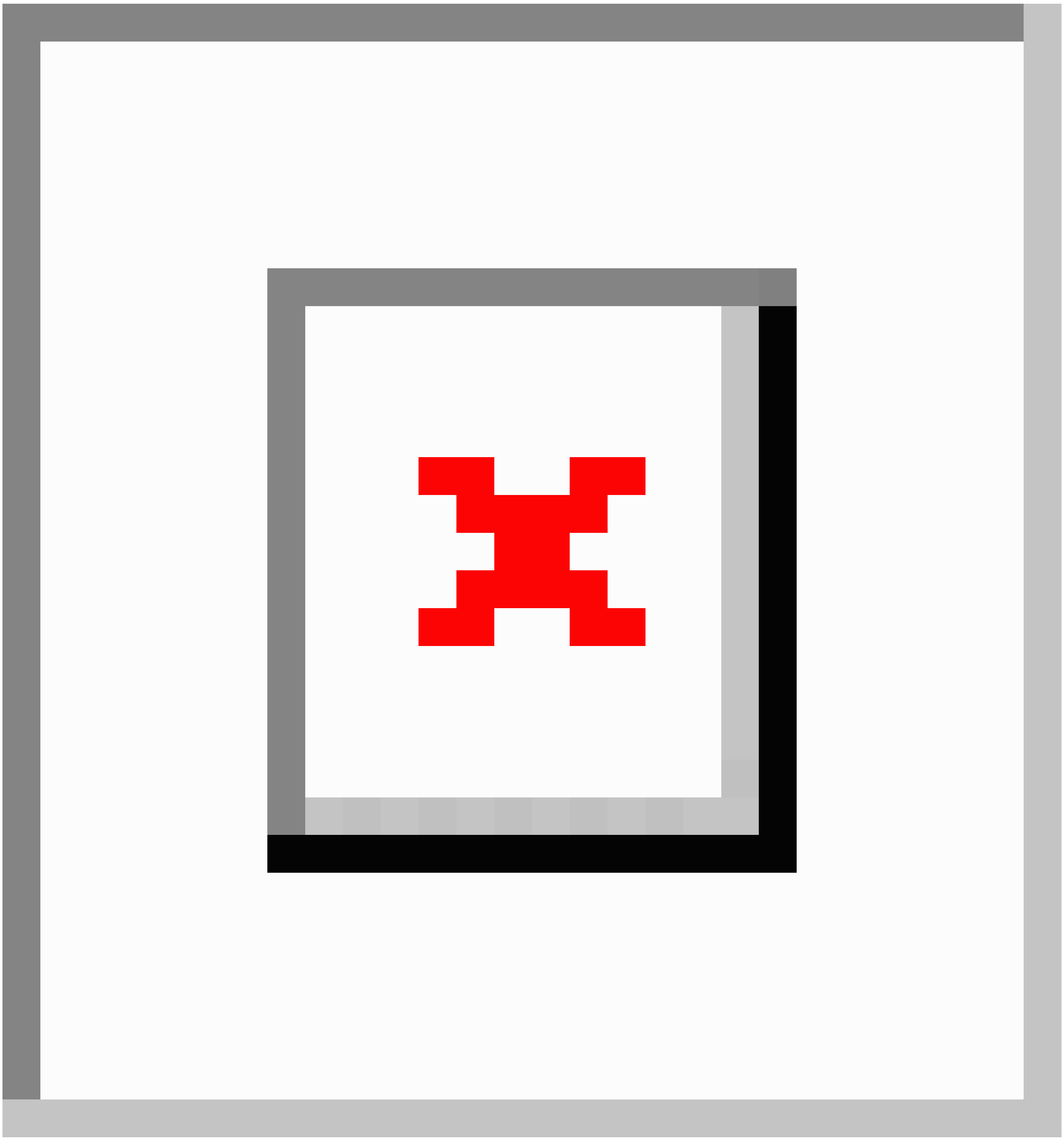
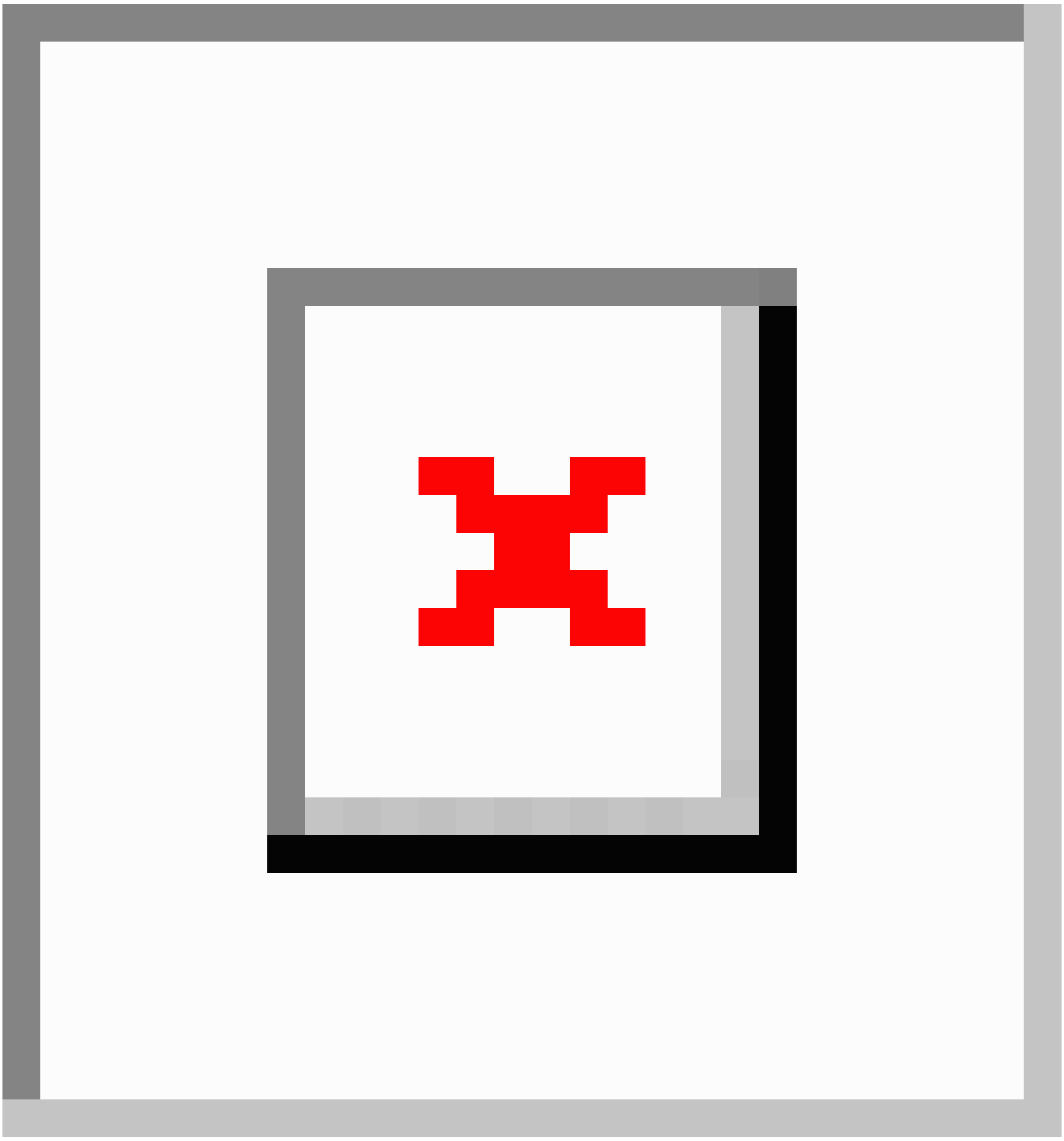


Figure 4. Comparison between groups after the training period. CG: control group; EG: experimental group; GPET: Game Performance Evaluation Tool.



Furthermore, it is evident that in the set of variables related to attack actions, the changes are more pronounced in variables associated with the execution mechanism than with those associated with the decision mechanism, except for dribbling and fixing. Conversely, in the set of variables related to defensive actions, the changes were more significant in the decision variables than in the execution variables, except for the variables of marking with the ball and marking without the ball.

Finally, the results obtained from the interview-based subjective evaluation of the soccer players' attention levels are presented in Table 5.

Table . Subjective levels of attention evaluated by the coach.

Variable	Control group, mean (SD)	Experimental group, mean (SD)	*P* value[a]
Subjective level of attention	4.83 (1.47)	5.83 (1.16)	.22

[a]*P*≤.05 was considered significant.

Discussion

Principal Findings

The primary aim of this study was to develop a video game targeted at enhancing the attentional skills of young soccer players. By providing coaches with a tool for training the psychological aspect of attention, positive effects on various performance variables in soccer players could be observed. Following the intervention program, the EG demonstrated improvements in video game performance, indicating significant progress in overcoming the challenges presented by the game. These findings are consistent with those of a previous study [23] that reported improvements in attention execution through software interventions and considered using video games as an interesting tool for improve cognitive variables [35] in this population [36]. While our study did not use a specific attention assessment such as the D2 Attention Test [37] used by Reigal et al [23]x, one of our key findings is that alongside attention training through the video game, we evaluated decision-making related to soccer performance.

To monitor somatic factors such as muscle activity and skin conductance, we used EMG activity and EDA measurements for all participants. Research suggests that cognitive control and management of the somatic aspects of anxiety could impact attention levels [38]. In our study, we specifically collected EMG activity data from the hand extensor muscles. Our evaluation methodology aligned with that developed by Palkowski and Redlarski [39], who assessed EMG activity during various hand gestures. Our results revealed similar EMG activity patterns during hand opening actions, comparable to the postures observed when our sample participants placed their nondominant hand on the table while using the dominant hand to operate the computer mouse.

However, our analysis of EDA did not reveal any significant differences between the groups. This lack of divergence in the somatic response to anxiety suggests potential heightened control over the outward display of anxiety during the test. Surprisingly, both the CG and the EG obtained similar results. One possible explanation for this finding could be the presence of the play component, as both EMG activity and EDA evaluations were conducted during the practice of the video game. These results are aligned with the findings of Pop-Jordanova and Pop-Jordanov [40], who observed similar skin resistance outcomes when assessing individuals using biofeedback in static and calm situations. However, it appears that the video game may not be an effective tool for regulating somatic anxiety.

The evaluation of performance-related variables in soccer players with the GPET confirms the perceptual improvements in decision-making and execution observed in the EG after engaging in the intervention program with the video game. This suggests that enhancing attention through video game practice has a positive impact on decision-making in relation to soccer performance. These findings are aligned with those of González-Víllora et al [41], who assessed decision-making and execution in soccer situations using the same tool and observed significant differences between groups. Thus, our results further support the notion that the video game can serve as a valuable tool for enhancing the psychological aspect of attention in young soccer players, yielding positive effects on various performance variables in soccer.

Regarding shooting decisions, the CG demonstrated higher results than the EG. One possible explanation for this is that the EG developed attentional skills through the video game intervention, leading to improved decision-making during soccer and consequently reducing the frequency of shooting attempts. While no significant differences were observed between the 2 groups in the coach's subjective evaluation of their players' attentional mechanisms, it is evident that the video game serves as an effective tool for decision-making training. Moreover, it shows promise as a tool for detecting attentional impairments or abnormalities, highlighting its potential beyond its training benefits.

Limitations

Despite our study reporting that this video game shows promise in improving attention in young soccer players, there are clear limitations that should be explained. Among the most important limitations of this study is that the sample size was small. This study comprised 12 soccer players, all members of a youth team of a first-level soccer team of the Spanish League. Although they comprised 100% of the team and effectively represented that team, our findings cannot be generalized; hence, it would be interesting to increase the sample size to verify our data.

The duration of the intervention was 6 weeks. Future studies could evaluate the effect of the intervention over a greater period, verifying if the effect on the improvement of some attentional processes assessed using the GPET was positive. In addition, the timing of the season where this attentional training program was integrated through the video game may also have been a limitation. Future studies could evaluate what happens at different times of the season.

Furthermore, somatic variables such as EDA and EMG activity were evaluated. Undoubtedly, the complexity underlying their evaluation is an important limitation in terms of reproducibility. Nevertheless, these variables were analyzed with the aim of determining whether there were any somatic issues that prevented the soccer players from being able to play the video game and obtain the best possible results.

Finally, there is a contextual limitation regarding the analysis of attention in real soccer situations. Although the ecological nature of the GPET has been proven, it is still a test that evaluates only the attention mechanism and the decision mechanism in soccer situations. This limitation could be overcome if future studies investigate what happens during a real match after attention training using the video game. Nevertheless, this evaluation was performed using the GPET to standardize the procedure so that all participants were evaluated in the same way and under the same conditions.

Acknowledgments

Sevilla FC (Football Club) deserves our heartfelt gratitude for generously granting us the opportunity to conduct this study. Their commitment to advancing knowledge and contributing to the sports community is truly commendable, and we are honored to have had their involvement in this endeavor.

We are very grateful to Enrique Arroyo for the fundamental work he has done in this study. His dedication is a symbol of effort and serves as a mirror for those of us who love sport.

Editorial Notice

This randomized study was only retrospectively registered. The editor granted an exception from ICMJE (International Committee of Medical Journal Editors) rules mandating prospective registration of randomized trials. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Conflicts of Interest

None declared.

Checklist 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (V 1.6.1).

[PDF File, 367 KB - games_v12i1e52275_app1.pdf]

References

1. Wehbe GM, Hartwig TB, Duncan CS. Movement analysis of Australian National League soccer players using global positioning system technology. J Strength Cond Res 2014 Mar;28(3):834-842. [doi: 10.1519/JSC.0b013e3182a35dd1] [Medline: 23897019]
2. Ballesta Castells C, García Romero J, Fernández García JC, Cruz A. Current methods of soccer match analysis. Rev Int Med Cienc Act Fis Deporte 2015;15(60):785-803.
3. Pons E, García-Calvo T, Resta R, et al. A comparison of a GPS device and a multi-camera video technology during official soccer matches: agreement between systems. PLoS One 2019;14(8):e0220729. [doi: 10.1371/journal.pone.0220729] [Medline: 31393932]
4. Jeras NMJ, Bovend'Eerdt TJH, McCrum C. Biomechanical mechanisms of jumping performance in youth elite female soccer players. J Sports Sci 2020;38(11-12):1335-1341. [doi: 10.1080/02640414.2019.1674526] [Medline: 31575323]
5. Landry SC, McKean KA, Hubley-Kozey CL, Stanish WD, Deluzio KJ. Neuromuscular and lower limb biomechanical differences exist between male and female elite adolescent soccer players during an unanticipated side-cut maneuver. Am J Sports Med 2007 Nov;35(11):1888-1900. [doi: 10.1177/0363546507300823] [Medline: 17921416]
6. Cosimini MJ, Watsjold B, Chan TM. Serious games without screens. Comment on "Involvement of End Users in the Development of Serious Games for Health Care Professions Education: Systematic Descriptive Review". JMIR Serious Games 2022 Feb 9;10(1):e34656. [doi: 10.2196/34656] [Medline: 35138255]
7. Olgers TJ, de Weg AAB, Ter Maaten JC. Serious games for improving technical skills in medicine: scoping review. JMIR Serious Games 2021 Jan 25;9(1):e24093. [doi: 10.2196/24093] [Medline: 33492234]
8. Clausen JD, Nahen N, Horstmann H, et al. Improving maximal strength in the initial postoperative phase after anterior cruciate ligament reconstruction surgery: randomized controlled trial of an app-based serious gaming approach. JMIR Serious Games 2020 Jan 24;8(1):e14282. [doi: 10.2196/14282] [Medline: 32012046]
9. Selmi O, Haddad M, Majed L, Ben Khalifa W, Hamza M, Chamari K. Soccer training: high-intensity interval training is mood disturbing while small sided games ensure mood balance. J Sports Med Phys Fitness 2018;58(7-8):1163-1170. [doi: 10.23736/S0022-4707.17.07292-9] [Medline: 28488826]
10. Selmi O, Ouergui I, Levitt DE, Nikolaidis PT, Knechtle B, Bouassida A. Small-sided games are more enjoyable than high-intensity interval training of similar exercise intensity in soccer. Open Access J Sports Med 2020;11:77-84. [doi: 10.2147/OAJSM.S244512] [Medline: 32210645]
11. Chang IC, Yang CY, Yen CE. The effects of a computer game (healthy rat king) on preschool children's nutritional knowledge and junk food intake behavior: nonrandomized controlled trial. JMIR Serious Games 2022 Jul 1;10(3):e33137. [doi: 10.2196/33137] [Medline: 35776502]
12. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, et al. Serious gaming and gamification education in health professions. J Med Internet Res 2019 Mar 28;21(3):e12994. [doi: 10.2196/12994] [Medline: 30920375]

13. Ko J, Jang SW, Lee HT, Yun HK, Kim YS. Effects of virtual reality and non–virtual reality exercises on the exercise capacity and concentration of users in a ski exergame: comparative study. JMIR Serious Games 2020 Oct 28;8(4):e16693. [doi: 10.2196/16693] [Medline: 33112240]
14. Ishak SA, Din R, Hasran UA. Defining digital game-based learning for science, technology, engineering, and mathematics: a new perspective on design and developmental research. J Med Internet Res 2021 Feb 19;23(2):e20537. [doi: 10.2196/20537] [Medline: 33605885]
15. Vanden Bempt F, Economou M, Dehairs W, et al. Feasibility, enjoyment, and language comprehension impact of a tablet- and gameflow-based story-listening game for kindergarteners: methodological and mixed methods study. JMIR Serious Games 2022 Mar 23;10(1):e34698. [doi: 10.2196/34698] [Medline: 35319480]
16. Boot WR, Kramer AF, Simons DJ, Fabiani M, Gratton G. The effects of video game playing on attention, memory, and executive control. Acta Psychol (Amst) 2008 Nov;129(3):387-398. [doi: 10.1016/j.actpsy.2008.09.005] [Medline: 18929349]
17. Concannon BJ, Esmail S, Roduta Roberts M. Immersive virtual reality for the reduction of state anxiety in clinical interview exams: prospective cohort study. JMIR Serious Games 2020 Jul 9;8(3):e18313. [doi: 10.2196/18313] [Medline: 32673223]
18. Rodrigo-Yanguas M, Martín-Moratinos M, González-Tardón C, et al. Effectiveness of a personalized, chess-based training serious video game in the treatment of adolescents and young adults with attention-deficit/hyperactivity disorder: randomized controlled trial. JMIR Serious Games 2023 Apr 24;11:e39874. [doi: 10.2196/39874] [Medline: 37093628]
19. Brown DJ, Fletcher D. Effects of psychological and psychosocial interventions on sport performance: a meta-analysis. Sports Med 2017 Jan;47(1):77-99. [doi: 10.1007/s40279-016-0552-7] [Medline: 27241124]
20. Kang JO, Kang KD, Lee JW, Nam JJ, Han DH. Comparison of psychological and cognitive characteristics between professional internet game players and professional baseball players. Int J Environ Res Public Health 2020 Jul 3;17(13):4797. [doi: 10.3390/ijerph17134797] [Medline: 32635282]
21. Ntoumanis N, Ng JYY, Prestwich A, et al. A meta-analysis of self-determination theory-informed intervention studies in the health domain: effects on motivation, health behavior, physical, and psychological health. Health Psychol Rev 2021 Jun;15(2):214-244. [doi: 10.1080/17437199.2020.1718529] [Medline: 31983293]
22. Vaughan RS, Laborde S. Attention, working-memory control, working-memory capacity, and sport performance: the moderating role of athletic expertise. Eur J Sport Sci 2021 Feb;21(2):240-249. [doi: 10.1080/17461391.2020.1739143] [Medline: 32129718]
23. Reigal RE, González-Guirval F, Morillo-Baro JP, Morales-Sánchez V, de Mier RJR, Hernández-Mendo A. Effects of a computerized training on attentional capacity of young soccer players. Front Psychol 2019;10:2279. [doi: 10.3389/fpsyg.2019.02279] [Medline: 31695637]
24. Hernández Mendo A, Martínez Jiménez M, Pastrana Brincones JL, Morales Sánchez V. Programa informático para evaluación y entrenamiento de la atención [Article in Spanish]. Rev Iberoam Psicol Ejerc Deporte 2012;7:339-357.
25. López LMG, Villora SG, Gutiérrez D, Serra J. Desarrollo y validación de la herramienta de evaluación del rendimiento de juego (HERJ) en fútbol [Article in Spanish]. SPORT TK-Revista EuroAmericana de Ciencias Del Deporte 2013;2:89-99. [doi: 10.6018/185791]
26. Markiewcz R. The use of EEG biofeedback/neurofeedback in psychiatric rehabilitation. Psychiatr Pol 2017 Dec 30;51(6):1095-1106. [doi: 10.12740/PP/68919] [Medline: 29432505]
27. Leonidou C, Pollatos O, Panayiotou G. Emotional responses to illness imagery in young adults: effects of attention to somatic sensations and levels of illness anxiety. Biol Psychol 2020 Jan;149:107812. [doi: 10.1016/j.biopsycho.2019.107812] [Medline: 31722237]
28. Tolin DF, Davies CD, Moskow DM, Hofmann SG. Biofeedback and neurofeedback for anxiety disorders: a quantitative and qualitative systematic review. Adv Exp Med Biol 2020;1191:265-289. [doi: 10.1007/978-981-32-9705-0_16] [Medline: 32002934]
29. Lang PJ, McTeague LM, Bradley MM. RDoC, DSM, and the reflex physiology of fear: a biodimensional analysis of the anxiety disorders spectrum. Psychophysiology 2016 Mar;53(3):336-347. [doi: 10.1111/psyp.12462] [Medline: 26877123]
30. Côté J, Saimela J, Trudel P, Baria A, Russell S. The coaching model: a grounded assessment of expert gymnastic coaches' knowledge. J Sport Exerc Psychol 1995 Mar;17(1):1-17. [doi: 10.1123/jsep.17.1.1]
31. Bloom G, Stevens D, Wickwire T. Expert coaches perceptions of team building. J of Appl Sport Psych 2003;15(2):129-143. [doi: 10.1080/10413200305397]
32. Cote J, Sedgwick WA. Effective behaviors of expert rowing coaches: a qualitative investigation of Canadian athletes and coaches. Int J Sports Sci 2003;7:62-77.
33. Escudero JT, Balagué G, García-Mas A. Estudio del conocimiento de variables psicológicas en entrenadores de baloncesto mediante una aproximación metodológica cuantitativa [Article in Spanish]. Revista de Psicología del Deporte 2002;11:111-122.
34. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc 2009 Jan;41(1):3-13. [doi: 10.1249/MSS.0b013e31818cb278] [Medline: 19092709]
35. Khaleghi A, Aghaei Z, Mahdavi MA. A gamification framework for cognitive assessment and cognitive training: qualitative study. JMIR Serious Games 2021 May 18;9(2):e21900. [doi: 10.2196/21900] [Medline: 33819164]

36. Halbrook YJ, O'Donnell AT, Msetfi RM. When and how video games can be good: a review of the positive effects of video games on well-being. Perspect Psychol Sci 2019 Nov;14(6):1096-1104. [doi: 10.1177/1745691619863807] [Medline: 31672105]
37. Brickenkamp R, Cubero NS. D2: Test de Atención [Book in Spanish]: TEA Ediciones; 2002. URL: https://selcap.cl/wp-content/uploads/2019/11/D2-Manual-Extracto-1.pdf [accessed 2024-04-19]
38. Eysenck MW, Derakshan N, Santos R, Calvo MG. Anxiety and cognitive performance: attentional control theory. Emotion 2007 May;7(2):336-353. [doi: 10.1037/1528-3542.7.2.336] [Medline: 17516812]
39. Palkowski A, Redlarski G. Basic hand gestures classification based on surface electromyography. Comput Math Methods Med 2016;2016:6481282. [doi: 10.1155/2016/6481282] [Medline: 27298630]
40. Pop-Jordanova N, Pop-Jordanov J. Electrodermal activity and stress assessment. Pril (Makedon Akad Nauk Umet Odd Med Nauki) 2020 Sep 1;41(2):5-15. [doi: 10.2478/prilozi-2020-0028] [Medline: 33011695]
41. González-Víllora S, García-López LM, Contreras-Jordán OR. Evolución de la toma de decisiones y la habilidad técnica en fútbol [Article in Spanish]. Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte 2015 Jan;59(2015):467-487. [doi: 10.15366/rimcafd2015.59.005]

Abbreviations

CG: control group
EDA: electrodermal activity
EG: experimental group
EMG: electromyographic
GPET: Game Performance Evaluation Tool

Edited by R Baranyi; submitted 29.08.23; peer-reviewed by G Caggianese, J Buele, M Aksoy; revised version received 08.03.24; accepted 09.03.24; published 27.05.24.

Please cite as:
Feria-Madueño A, Monterrubio-Fernández G, Mateo Cortes J, Carnero-Diaz A
The Effect of a Novel Video Game on Young Soccer Players' Sports Performance and Attention: Randomized Controlled Trial
JMIR Serious Games 2024;12:e52275
URL: https://games.jmir.org/2024/1/e52275
doi:10.2196/52275

Effects of a Virtual Reality Cycling Platform on Lower Limb Rehabilitation in Patients With Ataxia and Hemiparesis: Pilot Randomized Controlled Trial

Ana Rojo[1,2,*], PhD; Arantxa Castrillo Calvillo[3,*], BPT; Cristina López[3,*], BPT; Rafael Raya[1,*], PhD; Juan C Moreno[2,*], PhD

1

2

3

*all authors contributed equally

Corresponding Author:
Ana Rojo, PhD

Abstract

Background: New interventions based on motor learning principles and neural plasticity have been tested among patients with ataxia and hemiparesis. Therapies of pedaling exercises have also shown their potential to induce improvements in muscle activity, strength, and balance. Virtual reality (VR) has been demonstrated as an effective tool for improving the adherence to physical therapy, but it is still undetermined if it promotes greater improvements than conventional therapy.

Objective: Our objective was to compare the effect on lower limb range of motion (ROM) when using VR technology for cycling exercise versus not using VR technology.

Methods: A randomized controlled trial with 20 patients with ataxia and hemiparesis was carried out. The participants were divided into 2 groups: the experimental group (n=10, 50%) performed pedaling exercises using the VR system and the control group (n=10, 50%) performed pedaling exercises without using VR. Measurements of the active and passive ROM of the hip and knee joint were taken before and after a cycling intervention, which consisted of 3 sessions of the same duration but with progressively increasing speeds (4, 5, and 6 km/h). Repeated measures ANOVAs were conducted to compare the preintervention (T_i) and postintervention (T_e) assessments within each group. Additionally, the improvement effect of using the VR system was analyzed by comparing the variation coefficient ($\Delta = 1 - [T_e / T_i]$) between the preintervention and postintervention assessments for each group. Group comparisons were made using independent 1-tailed *t* tests.

Results: Significant improvements were shown in active left hip flexion (*P*=.03) over time, but there was no group-time interaction effect (*P*=.67). Passive left hip flexion (*P*=.93) did not show significant improvements, and similar results were observed for active and passive right hip flexion (*P*=.39 and *P*=.83, respectively). Neither assessments of knee flexion (active left: *P*=.06; passive left: *P*=.76; active right: *P*=.34; passive right: *P*=.06) nor knee extension showed significant changes (active left: *P*=.66; passive left: *P*=.92; active right: *P*=.12; passive right: *P*=.38). However, passive right knee extension (*P*=.04) showed a significant improvement over time. Overall, although active and passive ROM of the knee and hip joints showed a general improvement, no statistically significant differences were found between the groups.

Conclusions: In this study, participants who underwent the cycling intervention using the VR system showed similar improvement in lower limb ROM to the participants who underwent conventional training. Ultimately, the VR system can be used to engage participants in physical activity.

Trial Registration: ClinicalTrials.gov NCT05162040; https://www.clinicaltrials.gov/study/NCT05162040

(JMIR Serious Games 2024;12:e39286) doi:10.2196/39286

KEYWORDS

ataxia; cycling; hemiparesis; lower limb; neuropathology; rehabilitation; virtual reality; limb; intervention; neural; neural plasticity; therapy; muscle; strength; balance; tool; exercise; physical activity; neuroplasticity

Introduction

Background

Ataxia is an umbrella term for describing deficits in limb movement coordination such as dysmetria, dyspraxia, and dyssynergia [1]. The persistence of these deficits affects an individual's functional ability and poses a health challenge for both patients and clinicians.

Current scientific evidence indicates that the most effective treatment for ataxia should combine balance and coordination retraining and constraint-induced functional movement therapy [2]. However, the scientific literature still lacks a consensus on the details of these interventions and the timing of their implementation to enhance the recovery of the functionality of motor deficits in an individual [3].

On the other hand, in the field of neurophysiology, it is well known that to induce changes in neuroplasticity to achieve the functional recovery of motor deficits, the application of therapies based on the repetition of movements is required [4]. Some studies point out that the principles of motor learning are directly related to the regeneration of structures and the reorganization of neuronal function [5,6]. Moreover, the amount of practice is a key factor in motor learning, as well as the feedback provided during practice [7]. In fact, physical therapists must consider both the error feedback and activity guidance as 2 fundamental components of patient interaction during therapy to promote neuromotor learning [8]. Thus, interventions that promote normal function rather than the compensation of deficits are more recommended and should be applied to generate a physical activity plan based on the principles of motor learning and neural plasticity for patients with ataxic hemiparesis.

Prior Work

The scientific literature in the field of neurorehabilitation shows that pedaling exercises have the potential to induce improvements in muscle activity, strength, and balance [9]. This is mainly due to the fact that pedaling exercises based on the use of a cycloergometer provide a high number of flexion and extension repetitions [10] in the lower extremities for considerable periods of time. Because pedaling and walking are cyclical locomotor tasks that require the lower limb to alternate between flexion and extension [11,12], both share similar locomotor patterns of alternating muscle activation of antagonists [10,13]. Thus, cycling exercises are found to be useful for strengthening the lower limb muscles while acting as a pseudowalking task-oriented exercise. Some studies eluded that those biomechanical functions may be altered by the muscle groups involved in the pedaling tasks [14-16]. In fact, it was found that the degradation of pedaling performance in adults with hemiparesis was related to abnormalities in the execution of specific biomechanical functions [15]. Subsequently, it has been proven that human walking and cycling shared similar muscle synergies [16]. This evidence is the basis for rehabilitation treatments based on pedaling movements with potential positive outcomes for walking [16].

The ergometer is an equipment designed to perform cardiovascular work based on the alternative circular movement of the lower limb. Its use is advantageous for a muscle coordination study because balance is not an applicable factor in this kinematically constrained task [13]. In fact, applying an ergometer-based cycling routine could be useful because it requires no balance. Moreover, the exercise intensity of the ergometer-based cycling can be adapted to the user by adjusting the resistance of the pedal or the target speed. The ability to personalize the intensity of the exercise is a relevant factor for the patient's rehabilitation process. For these reasons, regular ergometer-based cycling is found to be a safer unsupervised exercise that is recommended for lower limb rehabilitation. Nevertheless, cycling exercise is also a static and repetitive form of exercise that leads to boredom and listlessness in patients. To deal with this discouragement factor, emerging technologies have been applied to elicit intrinsic motivation for rehabilitation patients [17]. Several studies pointed out the usefulness of gaming elements and virtual environments as assistive technology [18,19] and their potential effectiveness in physical therapies as opposed to conventional therapies [20].

Quite a few studies have focused on the analysis of functional metrics in virtual pedaling. A recent study evaluated the functionality of a virtual reality (VR) cycling training program that was applied to 10 patients with stroke [21]. It assessed the improvement of the bilateral asymmetry between the experimental group and the control group after the VR cycling intervention program. To evaluate this index, they equipped the ergometer pedals with force plates to determine the effect of the VR cycling training on each limb. The improvement of bilateral strength and standing balance was significantly different between VR cycling training and traditional physical training. Similarly, a previous study compared the effects of a cycling training program with extrinsic biofeedback and a nonimmersive interface versus traditional physical training on lower limb functional recovery in patients with stroke [22]. The results showed that improvements in walking endurance, walking speed, and muscle spasticity of the group using VR were significantly better than the group who underwent traditional physical training.

Objectives

The main objective of this study was to evaluate 2 different interventions: pedaling with VR and pedaling without VR. This study focused on comparing the improvements in lower limb range of motion (ROM) in pedaling activity between the group using VR and the group not using VR. To this end, a randomized controlled trial was carried out with patients with ataxia and hemiparesis. Hip and knee ROMs were measured before and after the cycling intervention. The overall aim of these analyses was to determine the effects of the 2 different interventions on short-term improvement of lower limb function and ROM.

Methods

VR System

The VR system implements extrinsic feedback strategies, gamification by levels, and personalization of the sessions with the aim of achieving greater adherence to pedaling exercise sessions. Its immersive nature means an increase in the sense of "presence," promoting the active involvement of the user.

The VR system is based on the transmission of the cycling kinematic data captured by the inertial sensors to the Oculus Quest 2 (Meta) head-mounted display (HMD) via Bluetooth. Therefore, the virtual application estimates the pedaling cycles, cadence, and distance during the exercise activity. The VR scenarios generated for this therapy consist of mapping the cycling cadence to the vehicle speed. Thus, the patient is placed inside a vehicle and visualizes the session data on the control panel while moving at the speed of the pedaling motion.

The design of the VR experience has been technically validated computationally to ensure low latency in motion analysis and visual representation of motion [23], thus preserving the embodiment effect and the sense of presence. Subsequently, the platform has also been validated from the point of view of satisfaction and ease of use of the system [24]. Additionally, considering that it is a stationary experience with an HMD that simulates a displacement, we evaluated to which extent the VR experience generates the type of motion sickness that causes fatigue, nausea, disorientation, postural instability, or visual fatigue [25]. Indeed, we verified that the platform does not generate adverse effects due to cybersickness [24].

Recruitment

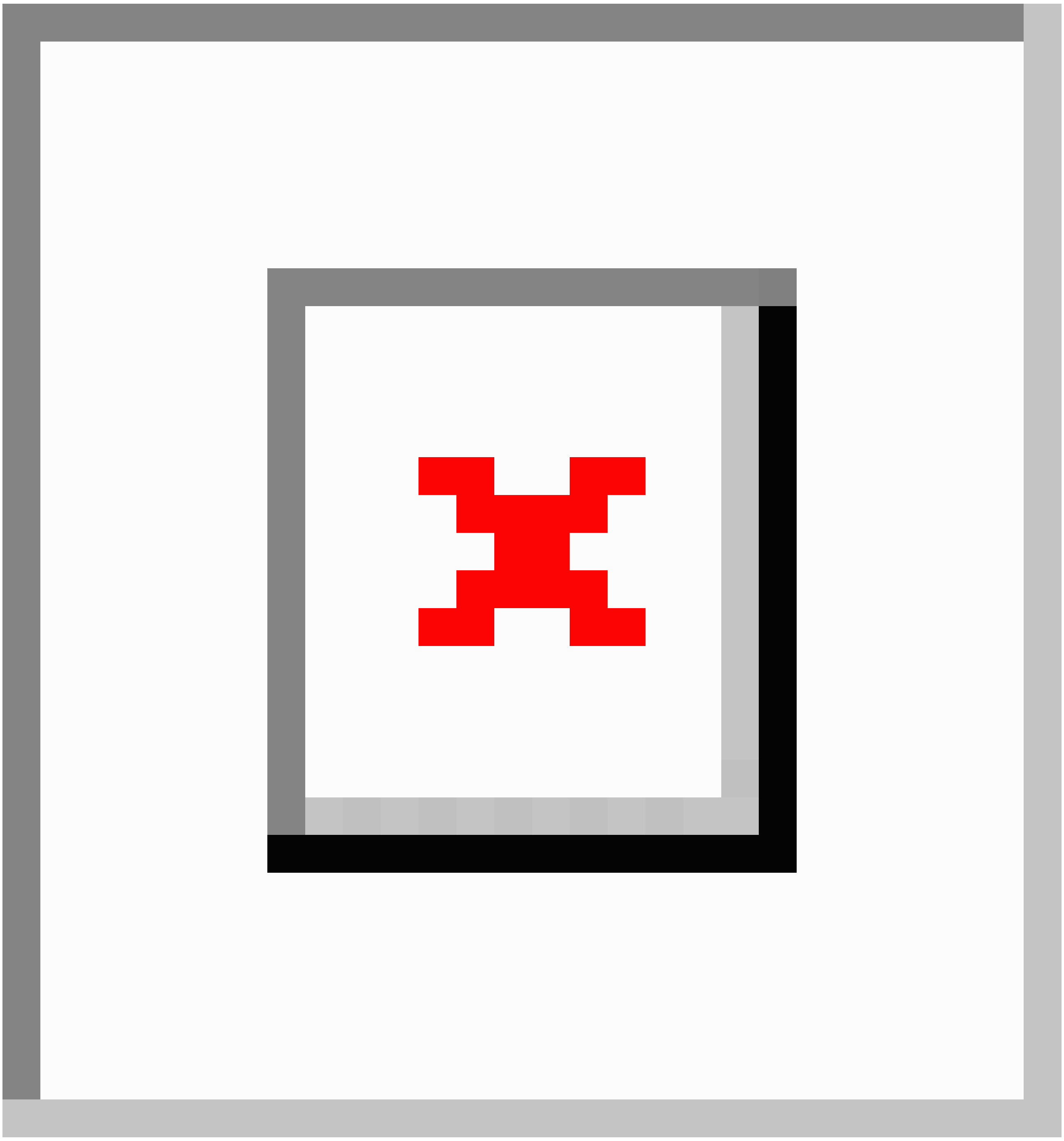
The participants were patients of both sexes between 18 and 90 years of age, recruited at the Lescer Clinic applying the inclusion and exclusion criteria. Inclusion criteria were as follows: individuals were eligible if they (1) had been prescribed pedaling exercise as treatment for lower limb rehabilitation and (2) were able to perform a pedaling session with VR technology. Exclusion criteria were as follows: (1) an insufficient cognitive state, (2) an unbound bone fracture, (3) severe disorders of vision or audition (inability to perceive visual or auditory information coming from VR), and (4) any incompatibility with the use of a VR system according to the clinical record. A sample of 22 participants (n=13, 59% male and n=7, 32% female; mean age 59.90, SD 13.56 y) volunteered to participate in this pilot randomized controlled trial (Table 1). Of this 22-person cohort, 1 participant dropped out of the study and 1 participant did not complete the study (Figure 1). The cohort was randomly divided into the experimental group (EG; 9/10, 90% male and 1/10, 10% female; mean age 60.80, SD 12.26 y) with VR cycling exercises or the control group (CG; 4/10, 40% male and 6/10, 60% female; mean age 59.00, SD 14.69 y) with traditional cycling exercises.

Table . Clinical and epidemiological features of the experimental group (EG) and control group (CG) participants.

Group and participant number	Sex	Age (y)	Etiology	Condition
EG				
1	Male	57	Ischemic stroke	Hemiparesis
2	Male	71	Hemorrhagic stroke	Ataxia
3	Male	53	Hemorrhagic stroke	Ataxia
4	Male	72	MCA[a] stroke	Hemiparesis
5	Male	53	MCA stroke	Hemiparesis
6	Male	62	Ischemic stroke	Hemiparesis
7	Male	59	Hemorrhagic stroke	Ataxia
8	Male	56	Progressive multifocal leukoencephalopathy	Ataxia
9	Female	86	Hemorrhagic stroke	Hemiparesis
10	Male	39	Ischemic stroke	Hemiparesis
CG				
1	Male	45	MCA stroke	Hemiparesis
2	Female	64	Hemorrhagic stroke	Ataxia
3	Male	58	Guillain-Barré syndrome	Hemiparesis
4	Female	41	Hemorrhagic stroke	Ataxia
5	Female	49	Ischemic stroke	Ataxia
6	Male	83	Ischemic stroke	Ataxia
7	Female	80	Hemorrhagic stroke	Hemiparesis
8	Female	72	Traumatic brain injury	Hemiparesis
9	Male	57	Ischemic stroke	Hemiparesis
10	Female	41	Guillain-Barré syndrome	Ataxia

[a]MCA: middle cerebral artery.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram.



Ethical Considerations

Ethical approval was obtained from the Research Ethics Committee of the San Pablo CEU University (550/21/51). This study has been registered at ClinicalTrials.gov (NCT05162040). All the participants were given written information in accordance with the Research Ethics Committee. The informed consent and the ability for participants to opt out was provided. Additionally, participants were informed that the data collected in this study can only be used for this study, not for secondary studies. The approval of the Research Ethics Committee of San Pablo CEU University only covers this study and does not cover a secondary analysis without additional consent. However, no additional analysis had been carried out.

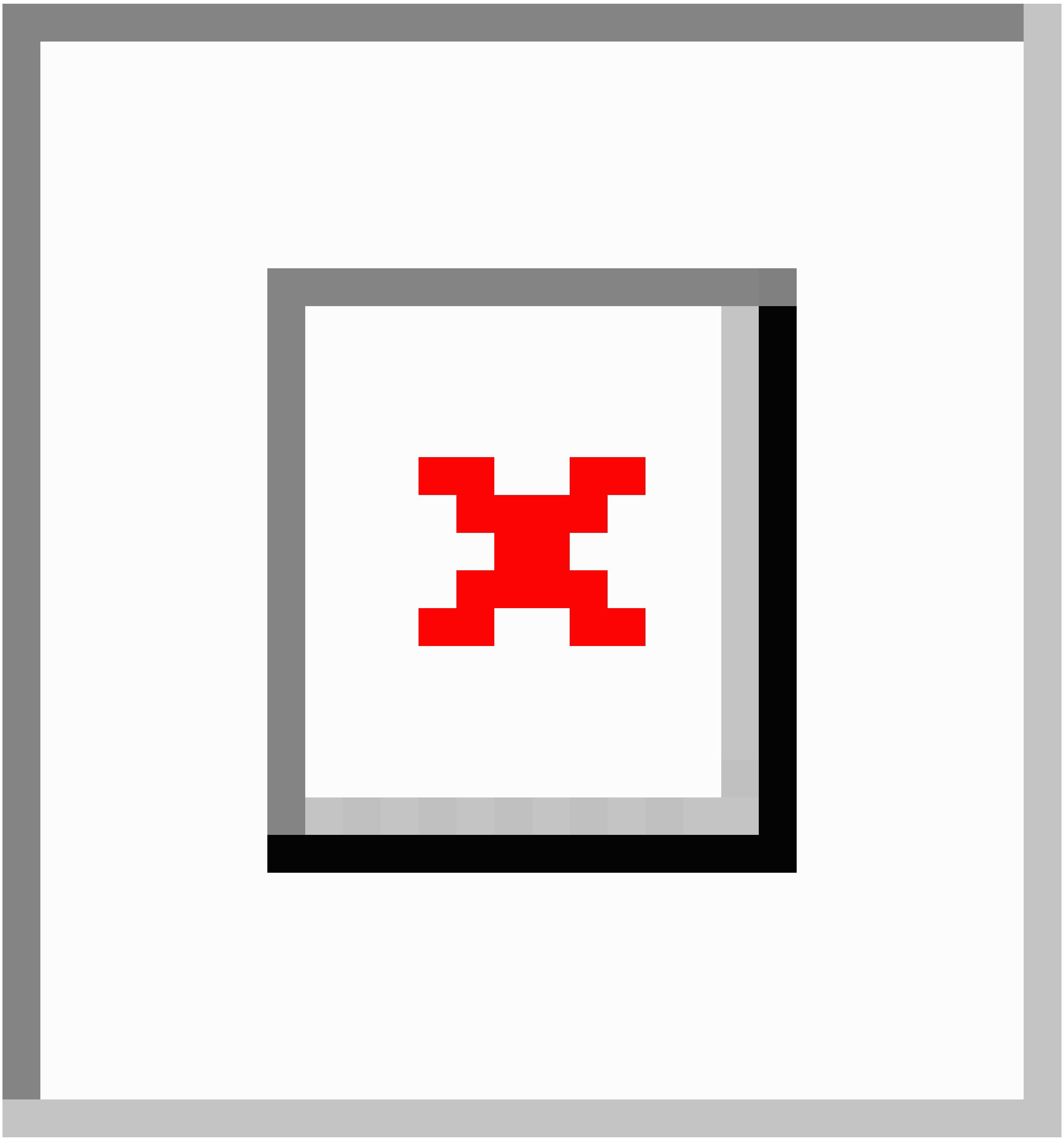
To ensure privacy and confidentiality, data are collected by employees of the agencies participating in the study. Each participant is assigned a unique code along with personal sociodemographic data and informed consent. These files remain in the custody of the principal investigator in charge of the project, while the assigned number is the one that identifies the anonymized data that was later analyzed. Finally, the participation in this study is completely voluntary; no compensation of any nature is offered to the human participants.

Intervention

This study was designed as a randomized controlled trial with 20 participants divided into 2 groups, following a block randomization method. The participants of the EG (n=10) performed pedaling exercises while using the VR system, whereas the participants of the CG (n=10) performed pedaling exercises without using the VR system. Before and after completing the exercise program, measurements of gait function metrics and joint ranges were performed to assess the effect of using VR stimulus during the cycling exercises.

The participants completed the cycling intervention simultaneously with their rehabilitation sessions. Afterward, for each participant, 3 cycling sessions were scheduled over 1 week with a maximum of 48 hours between sessions. Each session consisted of 2 sets of a 5-minute pedaling exercise spaced with a 2-minute break (to rest). Similar studies [19,26] have tested robotic unicycles in pedaling sessions at a cadence of 60 revolutions per minute. In our case, the pedaling speed of 1 cycle per second is equivalent to a target speed of 6 km/h. For this reason, it was decided to set this speed as the maximum speed and to start the first session with a slightly more comfortable speed (4 km/h) and increase it progressively (Figure 2). The participants of both groups performed the exercise following a set pedaling speed so that they received visual feedback according to the set target speed of 4-6 km/h for each session. The EG participants received visual feedback through the immersive VR application, whereas the CG participants received visual feedback on the ergometer display. All participants were instructed to maintain a constant pedaling speed throughout the session at the target cadence.

Figure 2. Summary of the intervention program for experimental and control group participants. VR: virtual reality.



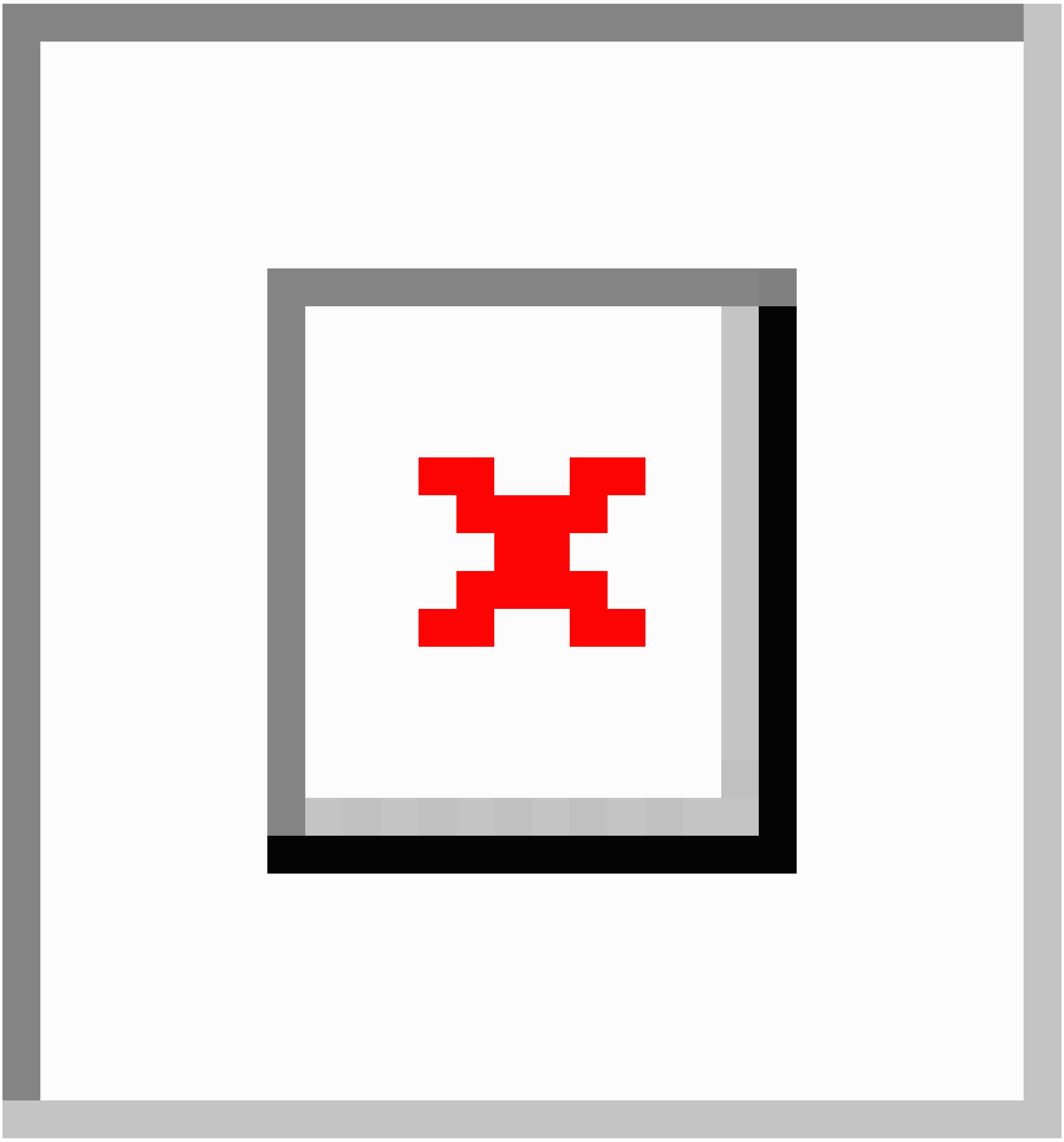
Physical Assessment

For the assessment of active and passive ROM of the hip and knee joint, a specifical ROM assessment tool was used. Measurements were extracted from biomechanical analysis using an inertial motion capture system (Werium; Werium Solutions) consisting of 2 inertial sensors: 1 placed in the distal part of the extremity (moving sensor) and the other in the proximal part (fixed sensor). Both sensors send their measurements via Bluetooth to a PC that runs the data acquisition software, Pro Motion Capture (Werium Solutions). This software computes the relative angle from both angle measurements (avoiding compensations) with an accuracy of 1 degree.

Protocol

The cycling sessions for both groups consisted of the use of a leg ergometer that allows training of the lower limb. Additionally, the EG used an inertial sensor placed on the right thigh and the Oculus Quest 2 HMD (Figure 3).

Figure 3. Cycling session of a participant in the experimental group using the virtual cycling platform.



The EG underwent the following procedure each session:

- The clinician connected the inertial sensor to the Oculus Quest 2 HMD.
- The patient was seated in a nonmovable chair (with no armrests) in front of the pedaling station during the entire session. The inertial sensor was placed on the right thigh

of the patient by adjusting an elastic band, and the sensor was turned on.

- The clinician fitted the Oculus Quest 2 HMD comfortably on the patient and guided him or her through the selection of the game scene. Once the game environment was entered, the clinician indicated the number of minutes of exercise and the target speed of the session so that the patient could configure these parameters on the interactive settings panel.
- Finally, the user performed 2 sets of a 5-minute cycling exercise with a 2-minute break between the sets.

Similarly, the CG underwent the following procedure each session:

- The patient was seated in a nonmovable chair (with no armrests) in front of the pedaling station during the entire session.
- The clinician turned on the ergometer's display and entered the number of minutes of exercise and the target speed of the session.
- Finally, the user performed 2 sets of a 5-minute cycling exercise with a 2-minute break between the sets.

Statistical Analysis

The data analysis model is the repeated measures model between 2 groups and the analysis of the longitudinal effect in increments of the measurements. Multifactor ANOVA analysis (with $P<.05$) were computed with SPSS Statistics (version 27.0; IBM Corp). The sample size was calculated using the software tool G*Power (version 3.1.9.7; Heinrich Heine Universität Düsseldorf). Ideally, assuming an effect size of 0.7, a minimum sample of 20 participants was required for the study to provide consistent statistical results. Since the effect size shows the strength of the relationships, it represents a minimum clinically meaningful difference. Of the many different types of effect sizes, the G*Power software uses Cohen *d* to characterize effect size by relating the mean difference to variability. Therefore, his study standardized the effect size to 0.7 for sample size calculation and power analysis.

Results

To identify the underlying differences between the preintervention (T_i) and postintervention (T_e) assessments in each group, repeated measures ANOVAs were conducted with time ($T_i – T_e$) as the dependent variable and group as the main within-subjects factor. When the ANOVA was significant, the Bonferroni post hoc test was used. To ensure that the error variance of the dependent variables is equal across groups, the Levene test was applied beforehand for all the metrics.

In addition, to identify the improvement effect due to the use or nonuse of the VR system, the variation coefficient between the preintervention and postintervention assessments was analyzed for each group as follows: $\Delta = 1 – (T_e / T_i)$. The variation coefficient outcomes were compared between groups by the independent 1-tailed *t* test. The mean and SD of the ROM outcomes for the hip and knee of each group are shown in Table 2. The mean increase Δ for each measurement is shown in Figures 4 and 5.

Table . Hip and knee range-of-motion outcomes.

Outcome	Experimental group, mean (SD)			Control group, mean (SD)		
	Preintervention (°)	Postintervention (°)	Variation coefficient (%)	Preintervention (°)	Postintervention (°)	Variation coefficient (%)
ALHF[a]	81.25 (36.09)	94.23 (32.26)	26.30 (33.52)	92.84 (21.40)	94.37 (25.83)	1.21 (14.20)
PLHF[b]	106.07 (21.16)	107.94 (17.63)	2.61 (5.81)	112.92 (17.76)	110.70 (16.83)	–1.43 (9.40)
ARHF[c]	97.55 (20.94)	97.13 (21.26)	0.28 (10.94)	97.11 (28.05)	101.79 (27.35)	5.60 (10.50)
PRHF[d]	106.63 (17.06)	109.82 (14.99)	3.69 (8.72)	119.74 (14.73)	117.71 (13.42)	–1.13 (8.72)
ALKF[e]	46.07 (14.62)	45.97 (11.47)	4.27 (26.31)	37.47 (12.03)	35.65 (8.47)	1.63 (30.86)
PLKF[f]	58.82 (9.84)	55.96 (9.79)	–3.48 (17.40)	57.14 (13.92)	54.58 (12.15)	–1.66 (19.96)
ARKF[g]	39.13 (16.54)	37.81 (10.68)	8.98 (35.88)	43.03 (10.00)	44.58 (13.32)	5.36 (30.29)
PRKF[h]	50.57 (10.02)	49.81 (10.31)	–0.65 (15.15)	63.35 (12.28)	57.28 (13.95)	–9.35 (15.17)
ALKE[i]	61.72 (14.86)	62.92 (13.11)	3.28 (9.74)	55.57 (17.13)	63.41 (11.77)	26.70 (51.68)
PLKE[j]	66.46 (11.74)	69.95 (15.09)	4.94 (15.92)	64.75 (11.94)	72.30 (12.46)	14.91 (24.40)
ARKE[k]	64.00 (10.11)	68.02 (10.14)	8.33 (20.67)	57.49 (14.91)	57.19 (14.76)	2.22 (23.52)
PRKE[l]	66.67 (11.53)	67.18 (10.93)	1.58 (14.10)	57.65 (11.21)	68.78 (6.67)	25.29 (34.70)

[a]ALHF: active left hip flexion.
[b]PLHF: passive left hip flexion.
[c]ARHF: active right hip flexion.
[d]PRHF: passive right hip flexion.
[e]ALKF: active left knee flexion.
[f]PLKF: passive left knee flexion.
[g]ARKF: active right knee flexion.
[h]PRKF: passive right knee flexion.
[i]ALKE: active left knee extension.
[j]PLKE: passive left knee extension.
[k]ARKE: active right knee extension.
[l]PRKE: passive right knee extension.

Figure 4. Summary of increments in active and passive hip ROM parameters with SD bars. The vertical axis represents the percentage of postintervention increase or decrease of each hip ROM parameter. ALHF: active left hip flexion; ARHF: active right hip flexion; PLHF: passive left hip flexion; PRHF: passive right hip flexion; ROM: range of motion.

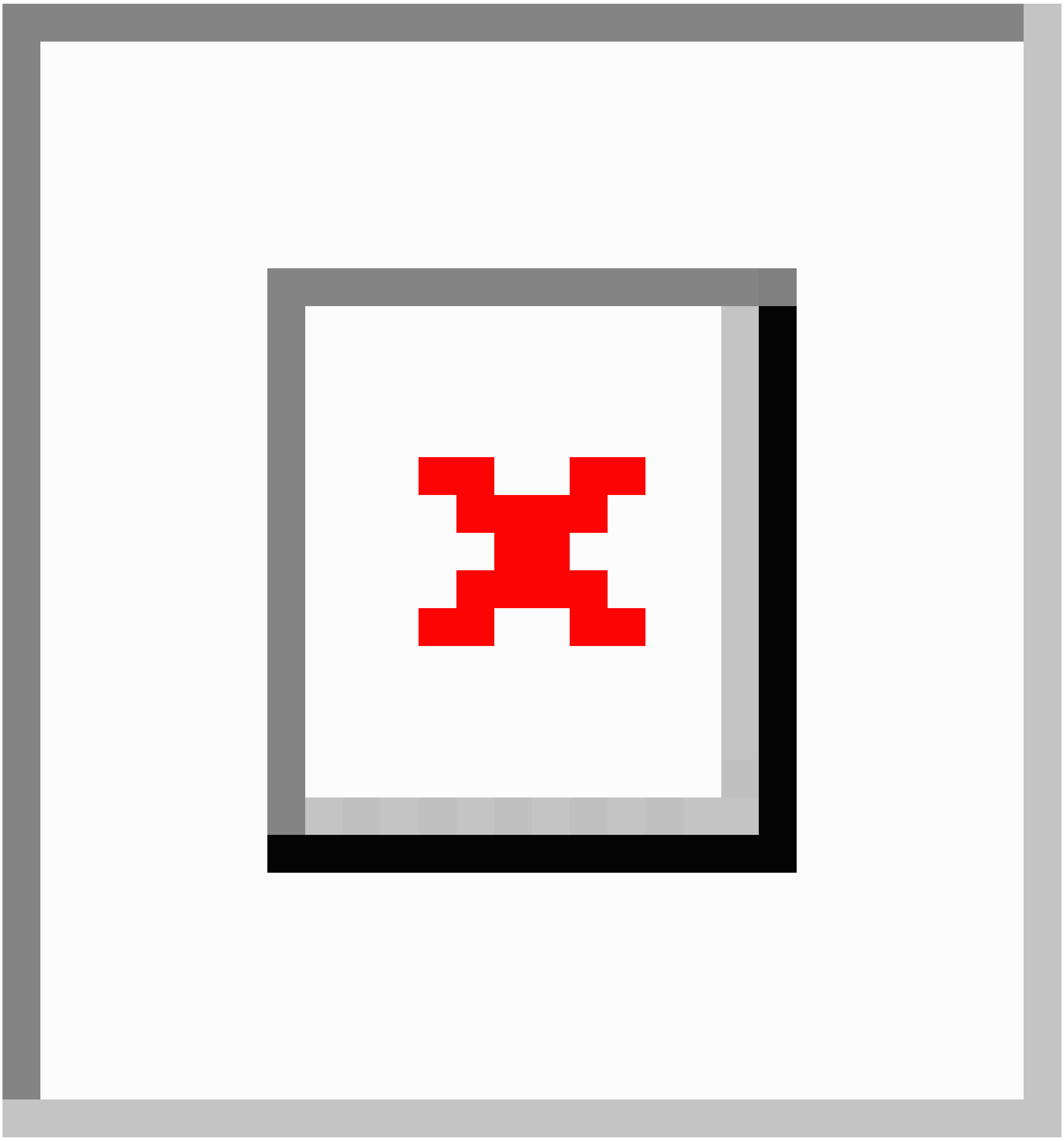
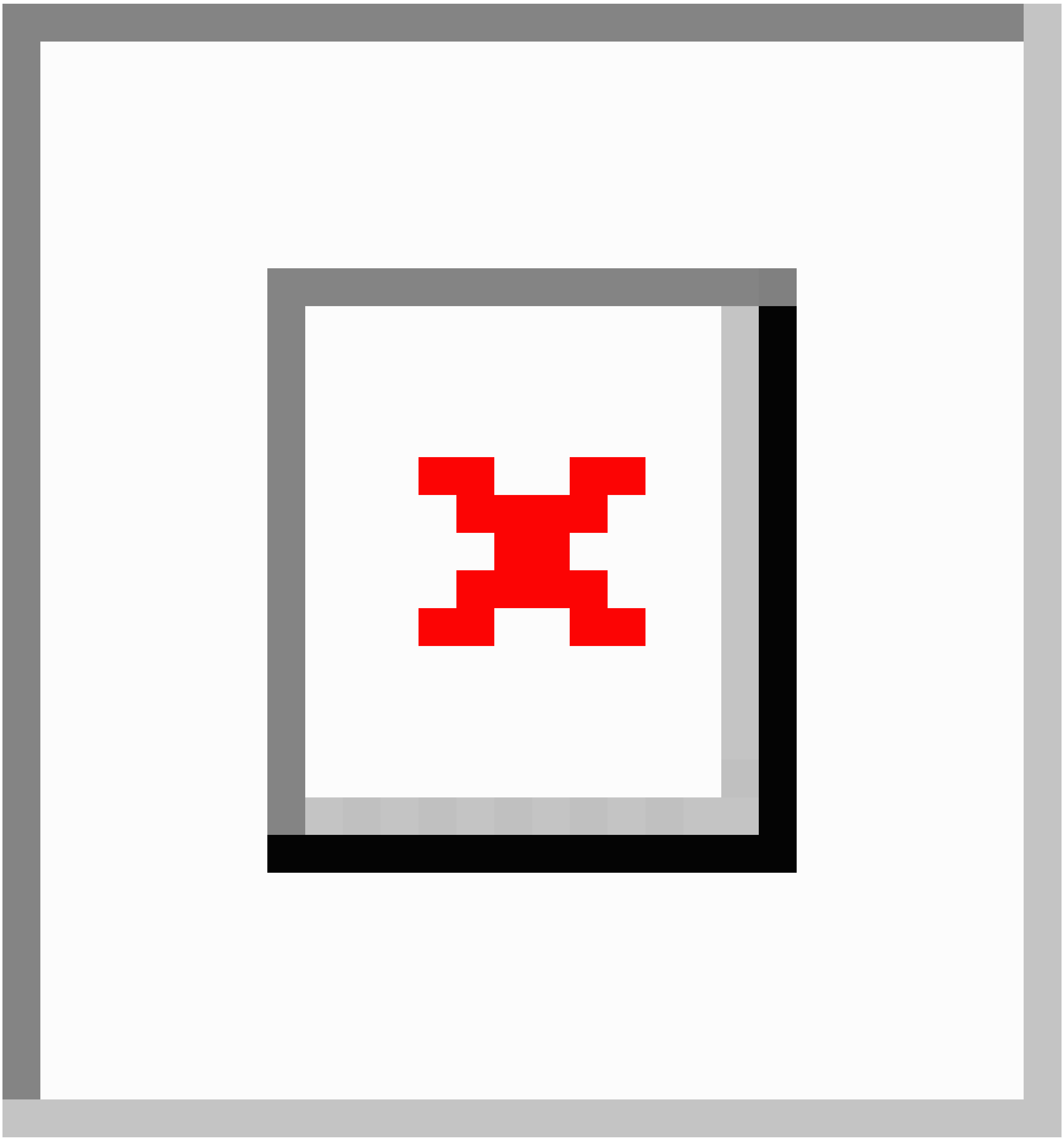


Figure 5. Summary of increments in active and passive knee ROM parameters with SD bars. The vertical axis represents the percentage of postintervention increase or decrease of each knee ROM parameter. ALKE: active left knee extension; ALKF: active left knee flexion; ARKE: active right knee extension; ARKF: active right knee flexion; PLKE: passive left knee extension; PLKF: passive left knee flexion; PRKE: passive right knee extension; PRKF: passive right knee flexion; ROM: range of motion.



With regard to the hip flexion outcomes, the active left hip flexion results were significant by ANOVA (*P*=.03), with no significance observed for the between-subjects effects test (*P*=.67). However, the within-subjects effects test was significant for the time factor (*P*=.03), but no significant group-time interaction effect was found (*P*=.08). Despite the opposing results showing passive left hip flexion improvements for each group, there was no significance difference by ANOVA (*P*=.93) and no statistically significant result was obtained by the between-subjects effects test. Passive left hip flexion was statistically significant in the within-subjects effects test for the time factor (*P*=.008). The active and passive right hip flexion results were not significant by ANOVA (*P*=.39 and *P*=.83, respectively). In both cases, no significant results were obtained for the between- and within-subjects effects tests.

For the knee ROM measurements, when analyzing the left knee assessments, the active and passive left knee flexion outcomes were not significant by ANOVA (*P*=.06 and *P*=.76, respectively). No statistically significant results were obtained by the between- and within-subjects effects tests in both cases. Similar results were obtained for the active left knee extension outcomes. Although reasonable differences in the active and passive left knee extension increases between groups can be observed in Figure 5, neither active nor passive left knee extension were significant by ANOVA (*P*=.66 and *P*=.92, respectively). No statistically significant results were obtained by the between- and within-subjects effects tests in both cases.

Regarding the right knee assessments outcomes, all outcomes were not significant by ANOVA (active flexion: *P*=.34; passive flexion: *P*=.06; active extension: *P*=.12; passive extension: *P*=.38). No statistically significant results were obtained by the between- and within-subjects effects tests for all cases, except for passive right knee extension, which was statistically significant for the time factor (*P*=.04) by the within-subjects effects test.

Discussion

Principal Findings

The aim of this study was to test the short-term effects of 2 different interventions on short-term improvement of lower limb function and ROM. For this purpose, a randomized controlled trial was carried out with participants with ataxia and hemiparesis.

In this study, the improvement outcomes of active and passive knee and hip joint ROMs due to the use of VR technology were inconclusive. Likewise, no statistically significant differences in the results between groups can be indicated. Even so, all the active ROMs measured—that is, performed by the patients—showed an increase with respect to the initial values. A greater disparity was observed in the passive measurements, although this may be attributed to the different passive mobilizations performed at each time by different physiotherapists. In this case, the active measurement is of special relevance in clinical terms because it indicates a ROM that the patient is able to achieve autonomously. On the other hand, large SDs in outcome variables clearly indicate that the improvements in the functional gait outcomes are not entirely consistent or represent a group effect. We observe that no significant effect can be attributed to VR intervention based on the statistical analysis of the immediate effects on gait function and joint ROM.

However, considering this similarity between groups, it can be pointed out that the use of VR has similar positive effects as the use of the conventional pedaling treatment. Thus, this immediate observation of effects leads us to conclude that the use of VR during pedaling exercise has similar effects to non-VR exercise training. Therefore, given that the use of VR technology does not worsen the improvement of lower limb ROM, and in line with the scientific literature [17-20], it may be advantageous to use it to maintain the patient's motivation.

Strengths and Limitations

A limitation of this study is the short-term nature of the intervention program. It is arguable that a longer intervention program would have shown more notable effects on functional improvement. However, assuming that it is precisely the treatment time that is one of the main causes of progress in physical improvement, the motivational impact of VR technology over time would need to be assessed. Therefore, further studies on the motivational impact of VR cycling versus conventional cycling on long-term physical activity remain to be addressed. Regarding these future studies, we suggest that cohort studies should be conducted among a population with more homogeneous neurological conditions. This recommendation is based on the limitations encountered in this study, where the difficulty of drawing conclusions about group changes or improvements with such wide SDs is presumably a reflection of the heterogeneity of the group.

Another factor to consider is that different physiotherapists were involved in taking the ROM measurements of the participants, although the measurement system was the same. This fact could be considered in future studies to evaluate interrater effects.

Future Directions

We consider it relevant to analyze, in future studies, whether these improvements in active and passive ROM are accompanied by greater muscle activation, in particular, the hamstrings, rectus femoris, gastrocnemius, and tibialis anterior muscles, as suggested by scientific literature [27].

Conclusions

The results of this trial demonstrate that pedaling exercises coordinated with VR technology works as successfully as conventional training for patients with lower limb disorders such as ataxia and hemiparesis. In this study, it was found that participants who performed the pedaling exercise program using the VR system showed similar results to the participants who performed the exercise activity without using VR technology. Overall, VR technologies can be a useful tool to help patients with ataxia and hemiparesis engage in lower limb exercise therapies.

Acknowledgments

The authors would like to thank all the participants who collaborated in this study, as well as the therapists and health care professionals from Centro Lescer for their participation in this study. The financial support for the industrial doctorate project "Desarrollo y estudio de una plataforma interactiva y un sistema electrónico de pedaleo para rehabilitación funcional de personas mayores" of the Autonomous Community of Madrid (IND2019/TIC17090) toward this research is hereby acknowledged. Grant PID2021-127096OB-I00 funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe."

The funding sponsors have no role in the design of the study; the collection, analyses, or interpretation of data; the writing of the manuscript; and the decision to publish the result

Data Availability

The data sets generated or analyzed during this study are available on the GitHub repository [28].

Authors' Contributions

AR contributed to software, data curation, formal analysis, and writing–original draft. ACC contributed to data curation and methodology. CL contributed to methodology, resources, and supervision. RR contributed to funding acquisition, supervision, and writing–review and editing. JCM contributed to funding acquisition, supervision, and writing–review and editing.

Conflicts of Interest

RR is the chief executive officer of Werium Solutions, and AR is a software developer at Werium Solutions. The other authors declare no conflicts of interest.

Checklist 1
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (V 1.6.1).
[PDF File, 1274 KB - games_v12i1e39286_app1.pdf]

References

1. Blumenfeld H. Neuroanatomy Through Clinical Cases: Oxford University Press; 2010.
2. Crum EO, Baltz MJ, Krause DA. The use of motor learning and neural plasticity in rehabilitation for ataxic hemiparesis: a case report. Physiother Theory Pract 2020 Nov;36(11):1256-1265. [doi: 10.1080/09593985.2019.1566941] [Medline: 30686101]
3. Januário F, Campos I, Amaral C. Rehabilitation of postural stability in ataxic/hemiplegic patients after stroke. Disabil Rehabil 2010;32(21):1775-1779. [doi: 10.3109/09638281003734433] [Medline: 20350209]
4. Lohse KR, Hilderman CGE, Cheung KL, Tatla S, van der Loos HFM. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. PLoS One 2014 Mar 28;9(3):e93318. [doi: 10.1371/journal.pone.0093318] [Medline: 24681826]
5. Katz-Leurer M, Sender I, Keren O, Dvir Z. The influence of early cycling training on balance in stroke patients at the subacute stage. Results of a preliminary trial. Clin Rehabil 2006 May;20(5):398-405. [doi: 10.1191/0269215505cr960oa] [Medline: 16774090]
6. Lucareli PRG, Greve JMD. Alteration of the load-response mechanism of the knee joint during hemiparetic gait following stroke analyzed by 3-dimensional kinematic. Clinics (Sao Paulo) 2006 Aug;61(4):295-300. [doi: 10.1590/s1807-59322006000400004] [Medline: 16924319]
7. Krakauer JW. Motor learning: its relevance to stroke recovery and neurorehabilitation. Curr Opin Neurol 2006 Feb;19(1):84-90. [doi: 10.1097/01.wco.0000200544.29915.cc] [Medline: 16415682]
8. Winstein C, Lewthwaite R, Blanton SR, Wolf LB, Wishart L. Infusing motor learning research into neurorehabilitation practice: a historical perspective with case exemplar from the accelerated skill acquisition program. J Neurol Phys Ther 2014 Jul;38(3):190-200. [doi: 10.1097/NPT.0000000000000046] [Medline: 24828523]
9. Peng CW, Chen SC, Lai CH, et al. Review: clinical benefits of functional electrical stimulation cycling exercise for subjects with central neurological impairments. J Med Biol Eng 2011;31(1):1-11 [FREE Full text] [doi: 10.5405/jmbe.718]
10. Ting LH, Kautz SA, Brown DA, Zajac FE. Phase reversal of biomechanical functions and muscle activity in backward pedaling. J Neurophysiol 1999 Feb;81(2):544-551. [doi: 10.1152/jn.1999.81.2.544] [Medline: 10036258]
11. Coast JR, Welch HG. Linear increase in optimal pedal rate with increased power output in cycle ergometry. Eur J Appl Physiol Occup Physiol 1985;53(4):339-342. [doi: 10.1007/BF00422850] [Medline: 4039261]
12. Winter DA. Biomechanical motor patterns in normal walking. J Mot Behav 1983 Dec;15(4):302-330. [doi: 10.1080/00222895.1983.10735302] [Medline: 15151864]
13. Raasch CC, Zajac FE. Locomotor strategy for pedaling: muscle groups and biomechanical functions. J Neurophysiol 1999 Aug;82(2):515-525. [doi: 10.1152/jn.1999.82.2.515] [Medline: 10444651]
14. Neptune RR, Kautz SA, Hull ML. The effect of pedaling rate on coordination in cycling. J Biomech 1997 Oct;30(10):1051-1058. [doi: 10.1016/s0021-9290(97)00071-7] [Medline: 9391872]
15. Brown DA, Kautz SA. Increased workload enhances force output during pedaling exercise in persons with poststroke hemiplegia. Stroke 1998 Mar;29(3):598-606. [doi: 10.1161/01.str.29.3.598] [Medline: 9506599]
16. Barroso FO, Torricelli D, Moreno JC, et al. Shared muscle synergies in human walking and cycling. J Neurophysiol 2014 Oct 15;112(8):1984-1998. [doi: 10.1152/jn.00220.2014] [Medline: 25057144]
17. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior: Springer; 2013. [doi: 10.1007/978-1-4899-2271-7]

18. de Bruin ED, Schoene D, Pichierri G, Smith ST. Use of virtual reality technique for the training of motor control in the elderly. Z Gerontol Geriatr 2010 Aug;43(4):229-234. [doi: 10.1007/s00391-010-0124-7] [Medline: 20814798]
19. Bruun-Pedersen JR, Pedersen KS, Serafin S, Kofoed LB. Augmented exercise biking with virtual environments for elderly users: a preliminary study for retirement home physical therapy. Presented at: 2014 2nd Workshop on Virtual and Augmented Assistive Technology (VAAT); Mar 30, 2014; Minneapolis, MN p. 23-27. [doi: 10.1109/VAAT.2014.6799464]
20. Laver K, Ratcliffe J, George S, Burgess L, Crotty M. Is the Nintendo Wii Fit really acceptable to older people? A discrete choice experiment. BMC Geriatr 2011 Oct 20;11:64. [doi: 10.1186/1471-2318-11-64] [Medline: 22011360]
21. Yin C, Hsueh YH, Yeh CY, Lo HC, Lan YT. A virtual reality-cycling training system for lower limb balance improvement. Biomed Res Int 2016;2016:9276508. [doi: 10.1155/2016/9276508] [Medline: 27034953]
22. Yang HC, Lee CL, Lin R, et al. Effect of biofeedback cycling training on functional recovery and walking ability of lower extremity in patients with stroke. Kaohsiung J Med Sci 2014 Jan;30(1):35-42. [doi: 10.1016/j.kjms.2013.07.006] [Medline: 24388057]
23. Rojo A, Raya R, Moreno JC. Virtual reality application for real-time pedalling cadence estimation based on hip ROM tracking with inertial sensors: a pilot study. Virtual Real 2023 Mar;27:3-17. [doi: 10.1007/s10055-022-00668-w]
24. Rojo A, Castrillo A, López C, et al. PedaleoVR: usability study of a virtual reality application for cycling exercise in patients with lower limb disorders and elderly people. PLoS One 2023 Feb 22;18(2):e0280743. [doi: 10.1371/journal.pone.0280743] [Medline: 36812248]
25. LaViola JJ. A discussion of cybersickness in virtual environments. ACM SIGCHI Bulletin 2000 Jan;32(1):47-56. [doi: 10.1145/333329.333344]
26. Costa-Felix R, Pomer-Escher A, Longo BB, et al. Neurorehabilitation platform based on EEG, sEMG and virtual reality using Robotic Monocycle. In: Costa-Felix R, Machado JC, Alvarenga AV, editors. XXVI Brazilian Congress on Biomedical Engineering: Springer; 2019:315-321. [doi: 10.1007/978-981-13-2119-1_48]
27. Kuo AD, Zajac FE. A biomechanical analysis of muscle strength as a limiting factor in standing posture. J Biomech 1993;26 Suppl 1:137-150. [doi: 10.1016/0021-9290(93)90085-s] [Medline: 8505348]
28. DataAvailability_JMIR. GitHub. URL: https://github.com/anarojoagusti/DataAvailability_JMIR [accessed 2023-12-20]

Abbreviations

CG: control group
EG: experimental group
HMD: head-mounted display
ROM: range of motion
VR: virtual reality

Edited by T Leung; submitted 05.05.22; peer-reviewed by I Cikajlo, P Postigo-Martin; revised version received 01.11.23; accepted 07.11.23; published 04.01.24.

Please cite as:
Rojo A, Castrillo Calvillo A, López C, Raya R, Moreno JC
Effects of a Virtual Reality Cycling Platform on Lower Limb Rehabilitation in Patients With Ataxia and Hemiparesis: Pilot Randomized Controlled Trial
JMIR Serious Games 2024;12:e39286
URL: https://games.jmir.org/2023/1/e39286
doi:10.2196/39286

Original Paper

Exploring the Use of a Learning-Based Exergame to Enhance Physical Literacy, Soft Skills, and Academic Learning in School-Age Children: Pilot Interventional Study

Aurelie Goncalves[1], PhD; Florence Lespiau[1], PhD; Gaëtan Briet[1], PhD; Eugénie Vaillant-Coindard[1], MSc; Angèle Palermo[1], MSc; Elsa Decobert[1], BSc; Nathan Allegret-Bourdon[1], BSc; Elodie Charbonnier[1], PhD

APSY-v, University of Nîmes, Nîmes, France

Corresponding Author:
Aurelie Goncalves, PhD
APSY-v
University of Nîmes
Rue du docteur Georges Salan
CS13019
Nîmes, F-30021
France
Phone: 33 466364518
Email: aurelie.goncalves@unimes.fr

Abstract

Background: There is ample evidence that most children do not perform enough physical activity (PA). To address this major public health problem, the French government implemented 30 minutes of daily PA (DPA) at schools but did not provide any supplemental resources or concrete guidance. Considering both children's interest in video games and the need for teachers to complete their curriculum, the use of a learning-based exergame that combines PA and learning appears particularly relevant.

Objective: The first objective of this study was to evaluate the feasibility of implementing 30 minutes of DPA through exergaming among school-age children. The second objective was to examine the effects of an exergaming program on physical literacy, academic learning, and soft skills (motivation, self-efficacy, and concentration).

Methods: This interventional study had a pre-post design and used the Play LÜ exergame platform. The study included 79 children with a mean age of 8.9 (SD 1.2) years from grade 2 (7 years old) to grade 5 (11 years old). Play LÜ requires players to throw balls against a wall to reach a target or to activate an object and provides an interactive game area for educational activities linked to specific learning themes. After a 4-session familiarization phase during which the teachers chose to prioritize mathematics learning in 30-minute DPA sessions, students took part in DPA sessions over a period of 3 weeks with Play LÜ and a motor skills circuit behind the LÜ setup to keep them continuously active. All sessions were carried out by PA specialists. Each session started with a warm-up using the Grööve application, continued with main activities promoting mathematics learning adapted to each grade level, and ended with a 3-minute meditation for returning to a calm and serene state using the Gaïa application. Before (T0) and after (T1) the program, students completed a self-evaluation booklet to assess their levels of physical literacy, academic performance, and soft skills.

Results: The implementation of this exergaming program was welcomed by the school's administration, teaching staff, and parents. After the program, we observed increased scores for physical literacy (difference +2.6, percentage change +3.6%; W=933.0; P=.002; r_{rb}=−0.39, 95% CI −0.58 to −0.16) and motivation in mathematics (+0.7, +9.8%; W=381.5; P=.005; r_{rb}=−0.44, 95% CI −0.66 to −0.16). In addition, it is important to note that some measures progressed differently across learning levels and age groups.

Conclusions: The study results indicate positive impacts of learning-based exergaming on physical literacy and motivation in mathematics among school-age children.

(JMIR Serious Games 2024;12:e53072) doi:10.2196/53072

KEYWORDS

learning support; exergaming; physics playground; educational games; primary school; children

Introduction

Background

Regular exercise and physical activity (PA) have been shown to benefit children's physical and social health, as well as their academic performance [1-3]. For children, the World Health Organization (WHO) recommends a minimum PA practice of 60 minutes of daily moderate-to-vigorous PA [4]. Yet, in France, only 41.8% of children reach these recommendations [5]. In other words, nearly 6 out of 10 children are physically inactive. In this context, the French national health strategy [6] has set as a major objective the implementation of a comprehensive policy of prevention and health promotion. As a result, since September 2022, primary schools have been required to provide 30 minutes of daily PA (DPA) to promote PA and encourage the development of children's motor skills and physical abilities. Distinct from the teaching of physical education (PE), the 30 minutes of DPA can take a variety of forms, adapted to the context of each school. They can be split up and combined over the various school and extra-curricular periods. This was extended to all elementary schools after 2 years of testing in 11,000 volunteer schools [7].

Given the significant amount of time children spend in school throughout their childhood, schools represent an ideal setting to achieve maximum impact with regard to improving PA levels [8,9]. Furthermore, several studies have suggested that combining PA with academic activities can improve children's health and cognitive functioning, which could subsequently lead to an improvement in children's academic performance [10-12]. Attention, in particular, which is a prerequisite for learning, is often targeted during classroom-based PA [13]. However, other variables, such as motivation and self-efficacy (referring to the child's perception of his or her capabilities), are well known to influence children's school performance and acquisition [14-17] and could be influenced by a more entertaining approach to learning via exergaming.

Despite all the benefits of implementing regular PA at school, in a systematic review, Nathan et al [18] highlighted several barriers, such as environmental context and resources with "a lack of time in the curriculum;" goals with "competing curriculum demands of other subjects" or "physical activity considered a lower priority than other subjects;" and beliefs about capabilities, such as a lack of teacher expertise and confidence in delivering PA, and intentions with "a lack of teacher motivation to implement PA" [18]. By contrast, the authors also mentioned several facilitators. Among them, the knowledge domain was indicated to play a facilitating role, for example, "sufficient knowledge about PA and health to effectively conduct PA" [18]. This dimension could be explored through the notion of physical literacy (PL), which corresponds to "the motivation, confidence, physical competence, knowledge, and understanding to value and take responsibility for engagement in physical activities for life" [19]. Indeed, PL is particularly important in early childhood, a crucial period for the development of fundamental movement skills [20] and the adoption of PA habits. Physically literate individuals are more physically active, spend more time playing sports, and are less sedentary. PL is a multi-level concept that is increasingly taken into consideration in the field of public health as it is a key determinant of PA habits across the lifespan [21].

One of the factors behind children's low levels of PA and high levels of sedentary behaviors is screen time use. Indeed, 71.7% of French boys and 58.5% of French girls aged 6 to 10 years have more than 2 hours a day of screen time [5]. Children aged between 8 and 12 years have 1.5 hours of daily screen time attributable to video games [22]. Over the last decades, video games have emerged as one of the most popular forms of global entertainment. Given children's keen interest in video games, it seems particularly appropriate to use gamification to encourage PA. For staying active while enjoying the pleasures of video games, a worthwhile alternative is exergaming. Indeed, exergaming or active video gaming requires bodily movements to play the game and encourages PA, with a focus on children's interest in the game's dynamics and stimulation. Our approach to exergaming takes into account a health dimension and can be associated with the conceptualization proposed by Oh and Yang [23], defining an exergame as "a video game that promotes (either via using or requiring) players' physical movements (exertion) that is generally more than sedentary and includes strength, balance, and flexibility activities."

Given that children spend most of their time at school, that they have a particular appeal for video games, and that exergaming seems to have beneficial effects on school learning [24], the use of exergaming at school appears to be an ideal solution for promoting PA, PL [25-27], and learning [28]. Furthermore, it appears that the use of a technology-based learning environment at school can increase soft skills, such as motivation and concentration on academic tasks [29]. Similarly, it has been shown that incorporating technology into an instructional intervention can improve students' sense of self-efficacy [30], which is a key variable for academic learning. Exergames, in particular, have been found to promote cognitive functions, motor skill training, enjoyment, and motivation to play among school-age children [31], and improve self-efficacy over traditional exercises [32]. Supporting this idea, it has been shown that exergaming (eg, Nintendo Wii Games [33]) incorporated into PE classes combined with health messages has a higher potential to enhance PA-related attitudes and behaviors than regular PE classes, especially in elementary school children [27]. An interesting exergaming tool for reconciling learning and DPA is the Play LÜ exergame platform (LÜ Interactive Playground) [34]. This technology can be used to change the traditional sports-school atmosphere into an interactive learning environment through interactive wall projection and a synchronized sound system. LÜ Playground activities are designed to improve the learning of children and adolescents by allowing them to respond to questions in specific fields (eg, mathematics, history, and natural sciences) by throwing balls against an interactive wall. This tool would therefore allow the practice of PA within non-PE curricula and thus ensure the 30 minutes of DPA among primary school children. Moreover, given the associations among cognitive functioning, soft skills, learning, and PA demonstrated in the literature, it appears essential to assess whether an exergaming program can improve these different variables.

Objectives, Research Questions, and Hypotheses

The first objective was to study the feasibility of implementing 30 minutes of DPA through exergaming. Given that exergaming combines the interests of children (for video games) and teachers (for learning and respecting the curriculum) while promoting PA, we hypothesized that it will enable effective implementation of the 30 minutes of DPA in schools.

The second objective of this study was to evaluate the effects of an exergaming program on PL, academic learning, and soft skills (motivation, self-efficacy, and concentration). We hypothesized that implementing a DPA program involving exergaming on a specific academic course combined with information on health-promoting behaviors daily could increase children's PL (hypothesis 1) and increase academic performance (hypothesis 2). Indirectly, allowing students to work on an academic subject more entertainingly through exergaming could improve students' motivation in the academic discipline (hypothesis 3), their sense of self-efficacy in the subject (hypothesis 4), and their concentration in class and the academic subject (hypothesis 5).

Methods

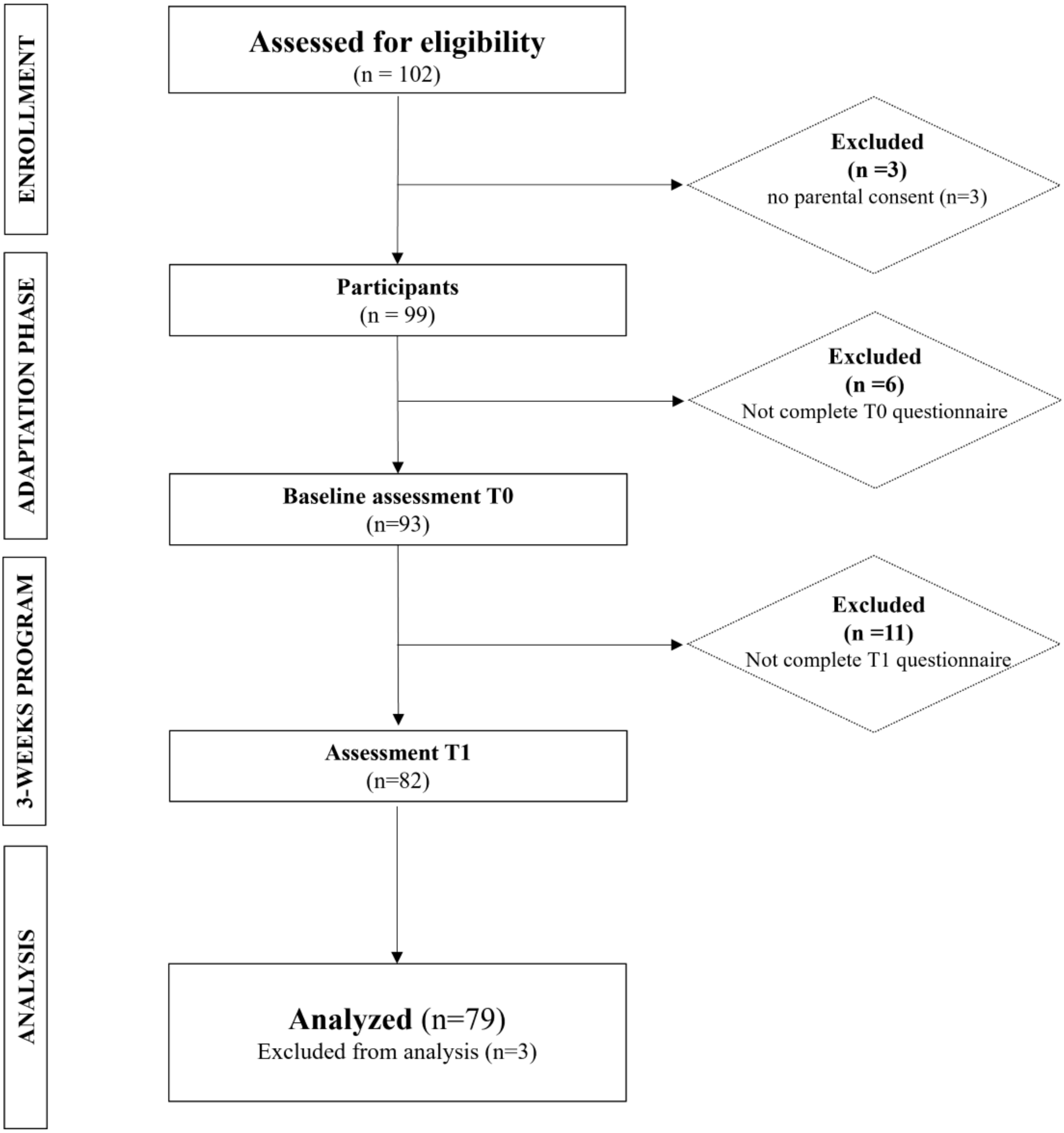
Population

This study was conducted with children aged 7 to 11 years as part of the implementation of the 30 minutes of DPA policy. The study was designed as an interventional study with a pre-post design. It included children from grade 2 (7 years old) to grade 5 (11 years old) in a mid-sized city school in the southern part of France, who had never benefited from any intervention in the field of exergaming. Before the project, the study and objectives of this research were presented to the school administration and then to the teachers. This pilot study took place in a small school with 1 class per level and 1 teacher per grade, with each of them (n=4) having no experience of exergaming and volunteering to take part in the research protocol. This school was selected for its pre-existing collaboration with the research team and middle-school students (8th grade), as well as for the availability of a space that could be used to install the LÜ mobile setup over a period of several weeks.

Subsequently, the parents of the children in the classes concerned were informed that their children would be part of a research protocol on 30 minutes of DPA during school time. A request for parental consent was sent via the school administration to each parent. In the event of parental refusal (only 3 parents refused), the children's data were not analyzed. A habituation phase was then proposed, and the teachers were able to learn about the various potentialities of the LÜ tool, as well as the implementation of the 30 minutes of DPA by the project team. The intervention then began and lasted 3 weeks, and preintervention (T0) and postintervention (T1) assessments were conducted.

During the enrollment period, 102 children were eligible (Figure 1). However, owing to the absence of parental consent (n=3) or the absence of children at evaluation time 0 (n=6) or time 1 (n=11), the analyses were carried out on 79 children. This final sample was made up of 34 girls and 45 boys, with a mean age of 8.9 (SD 1.2) years.

Figure 1. Study flow chart.



Class Measures

All the teachers expressed the wish to work on mathematics (geometry and arithmetic). A planning schedule was drawn up with the classes concerned so that the sessions could be scheduled during mathematics lessons.

Teacher Measures

At the end of the program, teachers were asked the following questions: On a scale from 0 to 10, how would you rate (1) your students' motivation for mathematics before the program? (2) your students' concentration for mathematics before the program? (3) your students' motivation for physical activity before the program? (4) your students' motivation for mathematics today? (5) your students' concentration for mathematics today? (6) your students' motivation for physical activity today?

Child Measures

At T0 and T1, students completed a questionnaire consisting mainly of analog visualization scales or checkboxes on different variables of interest (PL, motivation, self-efficacy, and concentration), which are described in the following sections. In addition, exercises adapted according to grade level were proposed in the target subject (mathematics) and a control subject (French).

Physical Literacy

PL was assessed using the Physical Literacy Assessment for Youth Self (PLAYself), designed for children aged 7 years or older, to explore children's perceptions of their PL [35]. PLAYself demonstrated robust psychometric properties, with good fit statistics, internal reliability, and a lack of item bias and problematic local dependency [36]. For a better understanding of the different dimensions of PL in the PLAYself questionnaire, the forms are available in English [37] and French [38] versions. The adaptation of this form within the evaluation booklet of this pilot project is available in Multimedia Appendix 1.

PLAYself consists of 22 questions divided into the following four subsections: (1) *Fitness*, which involves children's perceived fitness level with "disagree" and "agree" response categories for a single item; (2) *Environment*, which involves measures of 6 different environments in which children can do sports and activities (eg, "How good are you at doing sports and activities in the gym?") on a 5-point Likert scale ranging from 1 ("never tried") to 5 ("excellent"); (3) *Physical literacy self-description*, which involves 12 statements about doing sports and activities based on cognitive and affective factors (eg, "It doesn't take me long to learn new skills, sports, or activities"), where the children are asked to rank how well they agree on a 4-point Likert scale ranging from 1 ("not true at all") to 4 ("very true"); and (4) *Relative ranking of literacies*, which involves children's ranking of the importance of literacies in school, at home with family, and with friends (eg, "Math and numbers are very important in school") on a 4-point Likert scale ranging from 1 ("strongly disagree") to 4 ("strongly agree").

The first section is informative, while in the other 3 subsections, a separate score can be calculated and a total score can be obtained for PLAYself. The total PLAYself score is the average across the scores of each subsection, excluding the fitness question. A higher score (range 0-100) indicates a higher self-perceived PL.

Academic Achievement

To measure academic achievement, exercises in French and mathematics were retrieved from the national program by level following teacher school year progression. The test evaluated students' academic knowledge and skills related to specific subject areas, including French and mathematics. The test was grade-specific; did not contain any bias regarding age, gender, or ethnicity; and was scored as a percentage of achievement in French on one side and mathematics on the other.

Motivation, Self-efficacy, and Concentration

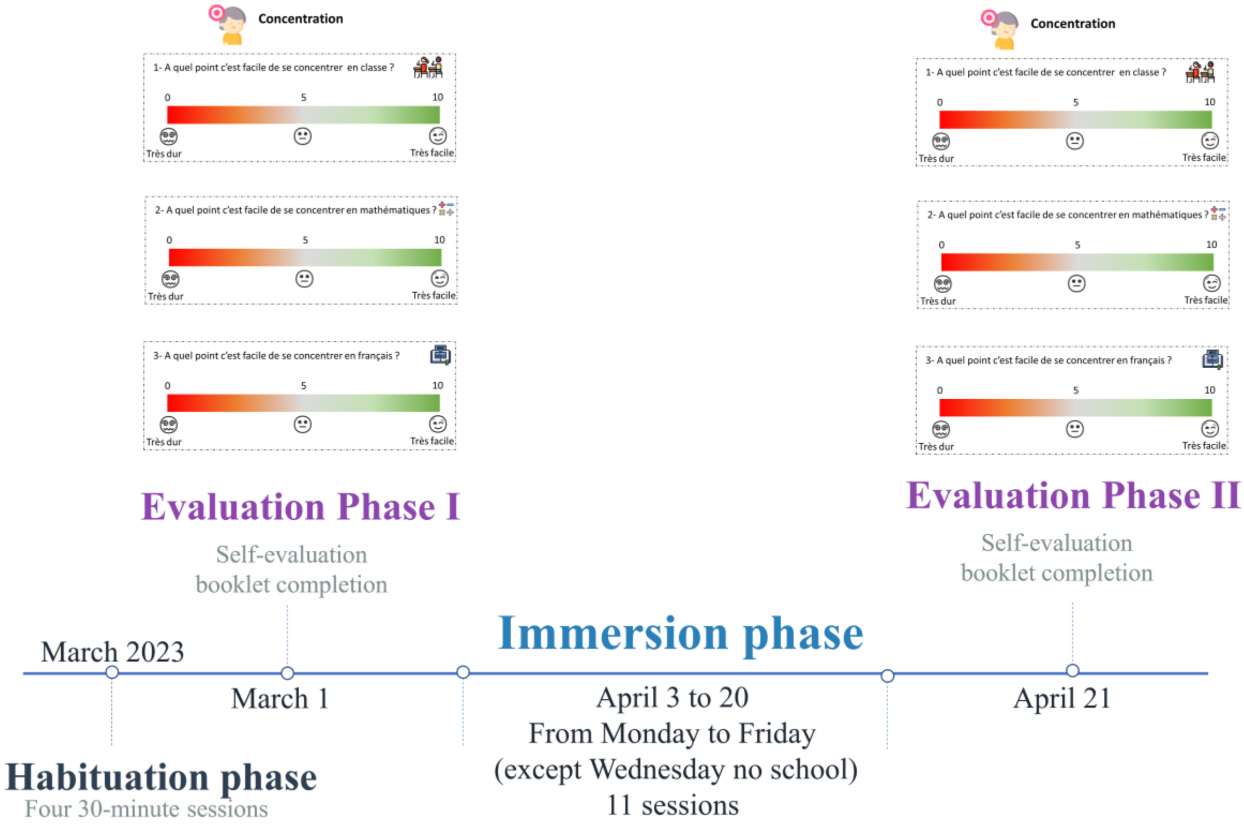
Motivation and self-efficacy were assessed by 2 items each, one for mathematics and the other for French. For motivation, children were asked: "How much do you enjoy doing [mathematics/French] exercises?" For self-efficacy, children were asked: "How well do you think you did on the [mathematics/French] exercises?" Concentration was assessed by 3 items, one for mathematics, one for French, and a more general one targeting concentration in class. For this variable, children were asked: "How easy is it to concentrate in [class/mathematics/French]?" We used a simple question per variable to reduce the time needed to complete the entire protocol. The items were formulated as clearly as possible to be adapted to the children's age and to ensure that they measure the core component of each variable. For all items, children were asked to respond using a 10-cm–long visual analog scale representing their feelings and marked by extreme labels at 0 cm (eg, very hard) and 10 cm (eg, very easy), which appeared as reliable response options in children's questionnaires [39].

Procedure

Habituation Phase

On Thursdays in March 2023, students had 4 LÜ 30-minute habituation sessions, spaced 1 week apart, enabling them to familiarize themselves with the interactive gymnasium. Activities linked to the academic development of the LÜ catalog were proposed, targeting language (ie, Minewörd), mathematics (ie, Wäk, Newton, Constello, and SphYnX), science and technology (ie, Brüsh and Grüb), history (ie, Störia), and arts (ie, Pixël). With mathematics accounting for one-fifth of the school program in each grade and the LÜ catalog offering more mathematics-related applications (except PE, which was not at the center of the project), the 4 teachers wanted to work on mathematics during the 3-week DPA immersion phase (Figure 2).

Figure 2. Timeline of the project.



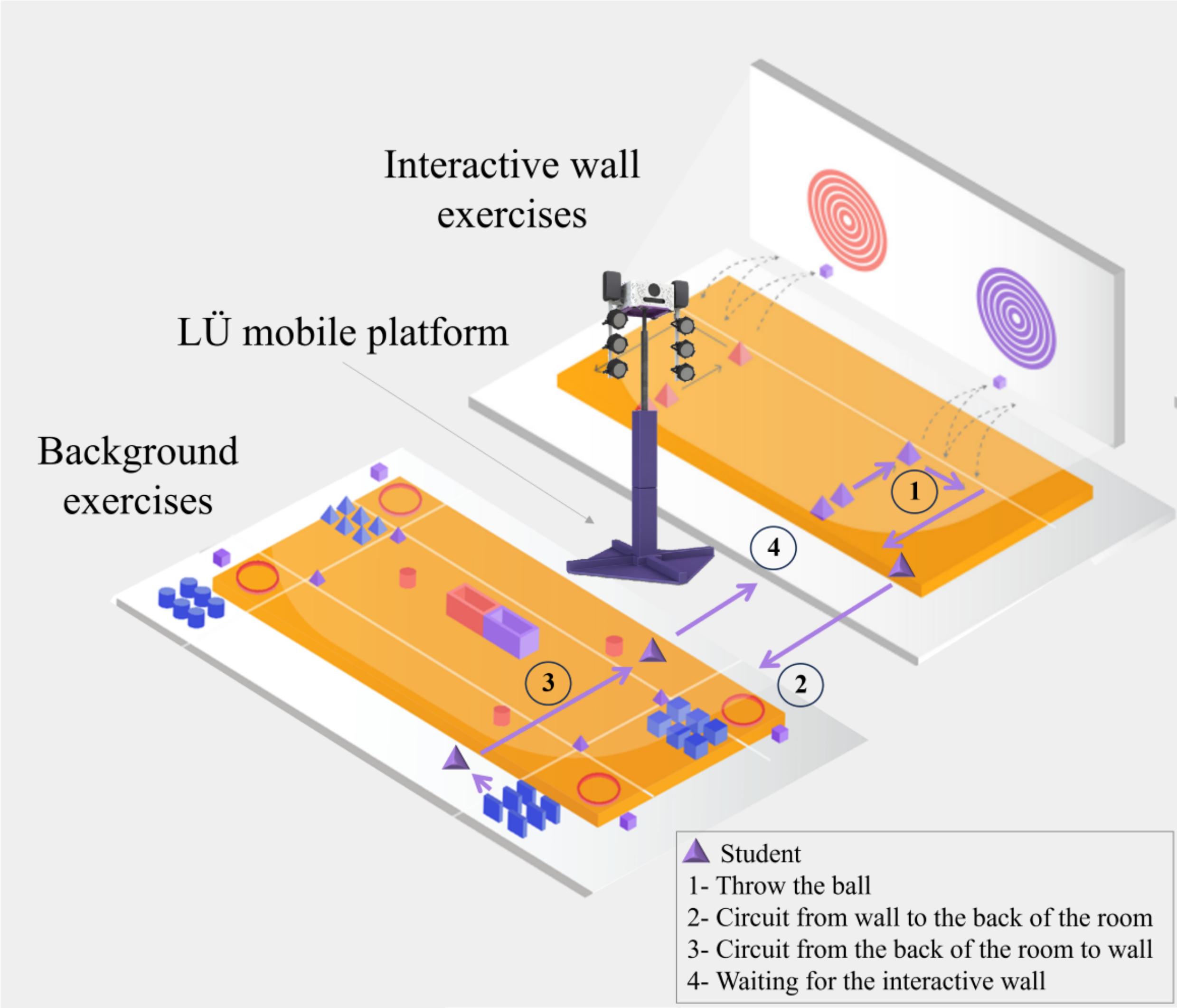
Immersion Phase

After the habituation phase, students took part in 30-minute DPA sessions using the Play LÜ exergame platform and worked on a single subject selected by the teacher (Figure 2).

Exergame Setting

For this research protocol, the LÜ mobile equipment owned by the research team was made available to the school for this pilot project and was installed in a designated space for the duration of the project. The Play LÜ exergame platform (LÜ Interactive Playground) has the potential to overcome the limitations of a physical room. With Play LÜ, the participants are immersed in the games displayed on a giant projection wall (6×3 m). The principal mechanism of Play LÜ requires the players to throw balls against the wall (eg, to reach a target or to activate an object; Step 1 in Figure 3). In addition, this mobile platform offers an interactive game area for educational activities linked to specific learning themes (calculations, puzzles, etc). For this research protocol, a work area with a daily changing activity circuit was implemented behind the LÜ mobile setup (without the interactive wall). With class compositions ranging from 26 to 30 students, this "with" and "without" interactive wall configuration was essential to keep students active during the 30-minute session. Once the ball was sent to the interactive wall (Step 1), the student was required to go behind the LÜ mobile setup toward the back of the room (Step 2) to carry out various exercises to promote different motor skills (eg, jumping, throwing, and balancing) and perform other exercises on the way back (Step 3). At the end of the circuit behind the LÜ setup (Steps 2 and 3), the student waited for his or her turn in front of the interactive wall (Step 4).

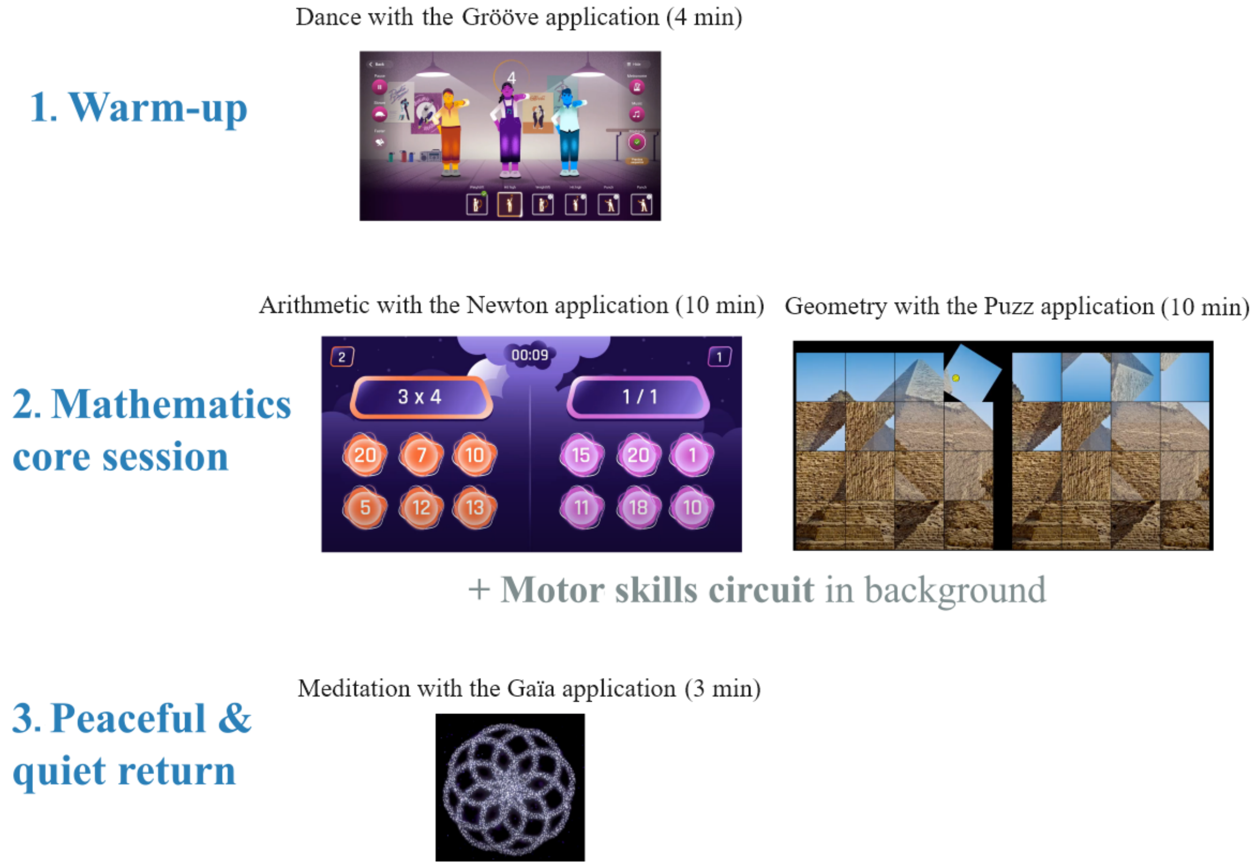
Figure 3. Practice environment implementation.



Daily Session Exercises With Play LÜ

During the 3-week immersion phase, sessions were structured in the same way, with a warm-up using the Grööve application at the beginning (Figure 4), which can be assimilated with the active video game Just Dance (Nintendo) [40], and then a core session promoting mathematics learning adapted to each grade, involving a section in front of the interactive wall with the Newton application for arithmetic and Puzz application for geometry (ie, with picture geometric forms or rules), and a section without the interactive wall consisting of a motor skills circuit (eg, throwing, jumping, and balancing) that enabled the child to be as active as possible (Figure 4). During break times (mainly while waiting for their turn on the interactive wall), the children had access to posters presenting active health behaviors with their favorite heroes according to age (ie, The Minions, Miraculous, and a successful French singer or Youtuber, depending on student age). The session ended with a 3-minute meditation for a return to peace and quiet, using the Gaïa application (Figure 4).

Figure 4. Daily physical activity core session details.



Focus on LÜ Applications

The Play LÜ exergame platform allows the use of applications (with or without customization) that can be used to meet general or specific learning objectives. The applications used during the immersion phase and their pedagogical benefits are summarized in Table 1.

Table 1. Play LÜ applications.

Application	Duration (min)	Description	Learning objectives		
			Academic development	Physical and motor development	Sociocognitive development
Grööve	4	This is a perfect warm-up and allows the development of gross motor skills.	• Movement • Dance	• Coordination • Cardiovascular endurance	• Cognitive flexibility • Inhibition
Newton	10	Newton is a fun way to combine physical activity and mathematics with the customization of equations by grade.	• Arithmetic	• Manipulation (throw the ball and aim at the target)	• Inhibition
Puzz	10	Throw the ball at a piece of the puzzle to rotate it and allow it to create an active knowledge competition.	• Puzzle created with geometry-related images • Spatial orientation	• Manipulation (throw the ball and aim at the target)	• Task engagement
Gaïa	3-5	This application helps students cool down after they have been active.	• Health and healthy habits	• Proprioception	• Emotional regulation

The customization of pre-existing games in LÜ was carried out on the Newton and Puzz applications for mathematics learning. For the Newton application, the difficulty of the operations proposed depended on each grade (eg, addition and subtraction calculations for grades 2 and 3, and multiplication calculations for grades 4 and 5). For the Puzz application, the images to be assembled were linked to the geometry program (eg, polygons and nonpolygons for grades 2 and 3, and complex polygons for grades 4 and 5).

Evaluation Phase

Before (T0) and after (T1) the 3 weeks of DPA (Figure 2), students completed a self-evaluation booklet to assess their levels of PL, academic performance, and soft skills that could be impacted by the program (eg, motivation, concentration, and self-efficacy).

At the end of the 3-week immersion phase, the students retook the same questionnaire, with only minor changes to the exercises (eg, 12+18 replaced by 14+15), but with the same instructions and level of difficulty. Following this, a short interview was conducted with the teachers, asking them to assess the changes observed in their classrooms.

All DPA sessions were carried out by sports science students specializing in adapted PA and health, under the supervision of a qualified teacher of adapted PA and health.

Statistical Analysis

Power analyses were conducted using G*Power (Heinrich-Heine-Universität Düsseldorf). For a pre-post comparison, with a medium effect size (0.50), an alpha error probability of .05, and a power of 0.95, we obtained a total sample size of 45. We then adjusted according to the number of participants available in the school, which allowed us to reach the sample size of 45.

As our data did not follow a normal distribution and given the characteristics of our sample, the intervention impact was tested with nonparametric within-group comparisons (T0 vs T1; Wilcoxon test, bilateral *P* values) for all participants and then by school-grade grouping. Effect sizes were expressed as the rank biserial correlation (r_{rb}) and its 95% CI. We also provided the score differences between T1 and T0, and expressed them as a percentage of improvement. Data were analyzed using JASP software (version 0.17.2.1; JASP Team).

Ethical Considerations

This pilot study involved an experiment in human and social sciences in the field of health. As mentioned in article R1121-1 section II subsection D of the French Public Health Code, this type of experimentation in human and social sciences does not require the authorization of the Committee for the Protection of Persons. Before the start of the study, a favorable opinion was obtained from the president of the University of Nîmes ethics committee. This individual verified that the study was conducted in accordance with institutional and national ethical standards, as well as the Declaration of Helsinki (2008). Moreover, this study was integrated into the school's activities and projects, and the protocol was validated by the school administration.

Consent to Participate

Concerning consent and information, the study was first presented to the school's teaching staff and administration who gave their approval to take part. Next, an online information notice and online informed consent form with the names and university affiliations of the experimenters were provided to the parents of all children in the classes involved in the study before initiation. Finally, the information and informed consent of the children and their teachers were collected face-to-face. Recruitment was based on voluntary participation, with no compensation for participants. Participants were informed that they could withdraw their consent at any time, whether at the request of the child, parent, or teacher.

Specific Measures Taken

To assure safety and security, all activities took place during class hours, under the supervision of the teacher, and the exercises were led by an associate professor specializing in public health and PA and two 3rd-year students in adapted PA from the University of Nimes. The expertise of the 3 animators enabled them to adapt the PA to the children's abilities in order to prevent any risk of injury. Moreover, the number of animators made it possible to provide individual support when needed. Finally, to guarantee the security of the data, they were stored on a secure university computer, and the printed versions of the data were kept in a secure cupboard in a university office.

Results

Feasibility of DPA and Exergaming

With regard to our first objective, which was to study the feasibility of implementing 30 minutes of DPA through exergaming, our results showed that implementing exergaming during school time is entirely feasible. First, regarding the parents, all but 3 were in favor of their children taking part in the project. Second, regarding the teachers, all agreed to take part in the project. Third, all scheduled sessions (n=11) were carried out, with no sessions canceled. External constraints, such as educational visits or other activities, could have led to cancellations, but the teachers expressed a desire not to miss any sessions and agreed to exchange schedules with other classes when constraints arose, demonstrating their interest.

Effects of the Exergaming Program

Effects on the Entire Cohort

With regard to our second objective for the whole cohort, the 3-week intervention of DPA led to increased scores in PL (in accordance with hypothesis 1) and motivation in mathematics, which was the subject covered in the intervention (in accordance with hypothesis 3). There was a general improvement regarding concentration in class, and we expected (hypothesis 5) this increase to be observed for mathematics as well (Table 2). In addition, contrary to our assumptions, we did not observe any changes in academic performance (hypothesis 2) or feelings of self-efficacy (hypothesis 4) in mathematics. Surprisingly, the intervention also favored French learning, which was not covered in the intervention, with academic performance, concentration, and self-efficacy in French being higher after the intervention.

Table 2. Descriptive analyses of variables of interest before (T0) vs after (T1) the 3 weeks of daily physical activity in the whole cohort (N=79).

Variable	Score before the program (T0), mean (SD)	Score after the program (T1), mean (SD)	Difference (T1-T0), value (% change)	*W*	*P* value	r_{rb}[a], value (95% CI)
Physical literacy	70.8 (13.6)	73.4 (13.3)	+2.6 (+3.6%)	933.0	.002[b]	−0.39 (−0.58 to −0.16)
Academic achievement						
Mathematics	63.6 (32.5)	60.8 (31.0)	−2.8 (−4.4%)	1440.0	.35	0.12 (−0.13 to 0.37)
French	59.3 (26.2)	65.0 (27.9)	+5.7 (+9.6%)	858.0	.05[b]	−0.26 (−0.50 to 0.00)
Motivation						
Mathematics	7.1 (3.6)	7.8 (3.0)	+0.7 (+9.8%)	381.5	.005[b]	−0.44 (−0.66 to −0.16)
French	5.3 (3.2)	5.6 (2.9)	+0.3 (+5.6%)	1222.0	.48	−0.09 (−0.34 to 0.16)
Concentration						
Classroom (general)	5.5 (3.0)	6.0 (2.7)	+0.5 (+9.0%)	934.5	.07	−0.24 (−0.48 to 0.01)
Mathematics	7.0 (3.2)	7.2 (2.8)	+0.2 (+2.8%)	580.0	.16	−0.21 (−0.48 to 0.08)
French	5.6 (2.7)	6.4 (2.8)	+0.8 (+14.2%)	838.0	.01[b]	−0.34 (−0.55 to −0.09)
Self-efficacy						
Mathematics	6.8 (2.8)	7.2 (3.1)	+0.4 (+5.8%)	887.5	.30	−0.14 (−0.40 to 0.13)
French	6.1 (3.1)	6.7 (2.8)	+0.6 (+9.8%)	728.5	.02[b]	−0.32 (−0.54 to −0.05)

[a]Rank biserial correlation.

[b]Statistically significant.

Effects on Grade 5 Participants

Focusing specifically on each grade (see Multimedia Appendix 2 for full details), participants in grade 5 were those most affected by the intervention.

After the intervention, grade 5 participants showed an increase in PL (difference +5.6, percentage change +7.9%; *W*=9.0; *P*<.001; r_{rb}=−0.91, 95% CI −0.96 to −0.78; hypothesis 1), mathematics motivation (+1.1, +20.0%; *W*=15.5; *P*=.007; r_{rb}=−0.77, 95% CI −0.91 to −0.43; hypothesis 3), and mathematics concentration (+0.8, +12.7%; *W*=21.5; *P*=.01; r_{rb}=−0.68, 95% CI −0.88 to −0.27; hypothesis 5) scores. Regarding French classes (even though they were not targeted by the intervention), grade 5 participants demonstrated an increase in motivation (+0.8, +17.7%; *W*=60.0; *P*=.05; r_{rb}=−0.48, 95% CI −0.76 to −0.03) and self-efficacy (+1.5, +29.4%; *W*=45.0; *P*=.02; r_{rb}=−0.57, 95% CI −0.81 to −0.14) scores after the intervention. They also tended to show an increase in academic achievement scores in French after the intervention (+7.3, +11.4%; *W*=53.0; *P*=.09; r_{rb}=−0.44, 95% CI −0.75 to −0.03). Contrary to hypothesis 1, grade 5 participants showed a decrease in academic achievement scores in mathematics after the intervention (−10.7, −13.6%; *W*=192.0; *P*=.006; r_{rb}=0.66, 95% CI 0.29 to 0.85).

Effects on Grade 2 to 4 Participants

The other grades also benefited from the intervention but to a lesser extent. After the intervention, grade 2 participants showed an increase in PL scores (difference +5.2, percentage change +7.1%; *W*=28.0; *P*=.03; r_{rb}=−0.58, 95% CI −0.84 to −0.11; hypothesis 1) and a marginal increase in academic performance in mathematics (+15.1, +33.7%; *W*=18.5; *P*=.06; r_{rb}=−0.59, 95% CI −0.86 to −0.06; hypothesis 2). Grade 3 participants showed an increase in mathematics concentration (+0.5, +7.5%; *W*=50.5; *P*=.04; r_{rb}=−0.51, 95% CI −0.79 to −0.07; hypothesis 5) after the intervention. Regarding French classes, grade 2 participants showed a marginal increase in self-efficacy scores (+1.0, +14.4%; *W*=24.0; *P*=.07; r_{rb}=−0.54, 95% CI −0.83 to −0.01), grade 3 participants showed a significant increase in academic performance (+16.8, +40.1%; *W*=68.0; *P*=.01; r_{rb}=−0.58, 95% CI −0.80 to −0.21), and grade 4 participants showed a marginal increase in concentration (+1.1, +20.7%; *W*=25.0; *P*=.09; r_{rb}=−0.52, 95% CI −0.82 to 0.01) after the intervention. Grade 4 participants showed a decrease in academic performance in French after the intervention (−15.0, −19.8%; *W*=103.5; *P*=.01; r_{rb}=0.72, 95% CI 0.33 to 0.90).

Effects on Teachers' Perceptions

Concerning the results obtained from teachers, 3 out of 4 teachers observed an improvement in their students' motivation for mathematics after the program (grade 2, T0=5.5 vs T1=7.5; grade 3, T0=7.0 vs T1=7.5; grade 5, T0=8.0 vs T1=8.5), and 1 teacher observed no change (grade 4, T0=6.0 vs T1=6.0). In addition, 2 out of 4 teachers observed an improvement in their students' concentration for mathematics after the program (grade 2, T0=6.0 vs T1=6.5; grade 4, T0=3.0 vs T1=5.0), and the other 2 teachers observed no change (grade 3, T0=6.0 vs T1=6.0; grade 5, T0=6.5 vs T1=6.5). Finally, regarding students' motivation for PA, only 1 teacher observed an increase (grade 2, T0=8.0 vs T1=9.0). It should be noted that 2 of the teachers who observed no change (grades 4 and 5) had already identified the maximum motivation score before the start of the program. The last teacher rated the students' motivation at 9.0 before and after the program.

Discussion

Principal Findings

Since September 2022, primary schools have been required to provide 30 minutes of DPA. In this pilot study, the implementation of an exergaming program as part of the 30 minutes of DPA policy was welcomed by the school's administration, parents, and teachers, with an increase in perceived motivation for mathematics.

After the program, we observed that children showed increased scores in PL and motivation in mathematics following 11 learning-based exergaming sessions.

Exergaming Implementation

Although this PA reform is recent in France, it has already been introduced in other countries several years ago. Indeed, the DPA school policy has been implemented since 2005 in Canada to promote active lifestyles for children in school settings [41], and in the province of Ontario, all elementary school children perform DPA during instructional time [42]. Yet, 10 years later (in 2015), it was revealed that only half of Ontario teachers were meeting this expectation [43], and this number dropped to 23% 5 years later in the report by Martyn et al [44]. The Canadian experience underscores the need to explore effective and sustainable methods for implementing the 30 minutes of DPA in schools. Consequently, this pilot study shows that exergaming can be used as a valuable tool in the deployment of DPA at schools.

Efficiency and Usefulness of Exergaming

Our second objective was to find out whether an exergaming intervention could be effective and useful. We hypothesized that implementing a DPA program involving exergaming on a specific academic course could have an impact on different aspects of a child's experience. First, in line with hypothesis 1, our results showed a significant increase in the PL of the entire cohort, with significant increases of over 7% for both grades 2 and 5. This result is all the more important as it has been highlighted that elevated PL leads to greater PA participation, resulting in positive physiological, social, and psychosocial adaptations, and thus improved physical, mental, and social health [45]. In other words, PL could play a role across the lifespan in promoting positive health. Therefore, exergaming seems to be an effective and useful instrument to promote PL. This observation is in line with the review by Sun [46], which highlighted that active video gaming could contribute to enhancing children's PL, in particular on the motivational aspect of exergaming, making it possible to provide a variety of opportunities to develop or reinforce basic motor skills among children.

Second, concerning motivational aspects, in line with hypothesis 3, our results showed an overall positive effect on students' motivation toward the discipline. Indeed, we found an increase in motivation for mathematics (target subject), with a significant increase of almost 10% for the total cohort, while motivation for French (control subject) was not impacted. It seems that allowing students to work on mathematics more entertainingly (ie, by throwing balls onto calculation operations) helps to increase their appeal in this course.

Third, contrary to hypotheses 2, 4, and 5, our results showed no increase in academic performance, motivation, and sense of efficacy in mathematics, but they showed an increase in these variables in French, even though this subject was not directly targeted in the sessions. Although the interpretation is limited without a control group that did not benefit from the intervention, it is conceivable that the participants benefited from additional motivational resources provided by the 30 minutes of DPA toward learning at school, in accordance with the results of Vazou et al [16] regarding motivation and self-efficacy. An argument in support of this explanation may be the marginal increase in general concentration in class for the total cohort, as has been observed in the review by Taras [47], which noted an immediate increase in concentration in students after PA. This overall concentration may have benefited all subjects, especially those frequently considered less difficult than mathematics (ie, French).

Finally, it is important to note that the positive effects of the intervention were found in all school grades, even if a greater benefit was observed in grade 5. As the ability to apply skills or knowledge learned during one activity to another activity is evidence of a transfer process, older children are likely to be more sensitive to it [48]. Indeed, in this study's intervention, the children were learning with different tasks and objectives (in DPA exergaming and their normal lessons). The transfer of skills from one to the other was therefore not obvious (even if the "mathematics" cue was common to both) and remains a particularly demanding cognitive process for which the children need to be motivated. Decreases in academic performance in French and mathematics (grades 4 and 5) may be explained by constraints in the classrooms, as the teachers were rotated during the semester and the last data collection took place the day before the vacation (the participants were less involved overall in the academic exercises required). However, the marginal increase in mathematics performance in grade 2 and the increase in French performance (overall cohort and grade 3) demonstrate the importance of continuing to test this intervention.

Practical Implications in the Educational Context

Learning-based exergames can be powerful allies in the implementation of the DPA policy at schools. For schools and educational teams, the first obstacle could be the associated cost. In France, the Ministry of Education has launched a call for projects entitled, "Pour un socle numérique dans les écoles élémentaires" ("for a digital base in elementary schools") [49] to equip the schools of tomorrow. In this context, it is necessary to create links between the worlds of research and education. Researchers need to present teachers with the advantages (ie, academic performance, self-efficacy, motivation, PL, PA, and sedentary behavior) and constraints (ie, update, group management, and security) of this type of practice to make the teachers as efficient as possible in different teaching situations. Indeed, as part of the 30 minutes of DPA policy in elementary schools, one of the major difficulties is sustaining the actions and motivation of teachers, as presented in the Canadian study [43,44]. Once the equipment has been acquired and installed in a fixed position (ceiling-mounted model), one of the solutions for maintaining motivation among teaching teams would be to integrate PA professionals into the internship framework. This option enables teachers to not only benefit from the specific skills of the trainees but also position themselves as observers of the class, to be able to work on specific notions during PE teaching [18].

Limitations

In the context of this pilot study, which focused mainly on implementation feasibility and learning, it would have been interesting to consider the children's physical fitness (ie, muscular strength, agility, and cardiorespiratory fitness) and general state of health. Indeed, a French longitudinal study with a 3-year follow-up of children aged 7.7 years at the start of the study showed that the physical fitness of French youth decreased between childhood and early adolescence [50].

It would also have been interesting to compare the effects of this program with a control group. For example, it would have been worthwhile to compare the scores of the experimental group involving exergaming and targeted school exercises to 2 control groups: the first one with no PA and no school exercises, and the second one with no PA but with school exercises identical to the experimental group (eg, on a tablet computer). To verify the validity of the results, it would also have been necessary to vary the targeted school exercises (eg, mathematics and French; randomizing their inclusion in the intervention to ensure that the most difficult material is not the only one tested). Moreover, this study was carried out in a single school with a single class per level. It would be worthwhile to increase the size of the cohort by increasing the number of classes per level in different schools. Furthermore, our program had a limited duration (3 weeks), and a longer program (at least 1 trimester) with more DPA sessions would undoubtedly have increased the effects we observed and allowed additional benefits to be observed. In future studies, it would be interesting to compare the effects of a short program like ours with those of a longer program, and this could shed light on the duration of effects through time (in particular, following longer interventions).

Perspectives

Future studies could explore the possible diffusion effects of enhanced DPA interventions with or without exergaming on various PA indicators (eg, physical fitness, increased mobility by accelerometry, sedentary time and breaks, and increased implication in PE curricula at school). It would also be useful to conduct a longitudinal study to measure the impact of exergaming on not only PA and fitness levels but also the evolution of overweight and obesity in children.

Furthermore, in future studies, it could be relevant to assess intervention effects on students' academic performance, motivation, and self-efficacy in specific academic courses during interventions, or general attitudes and performances in different courses. Finally, specifying various student profiles concerning these measures (eg, depending on the initial levels of PL and PA, and depending on age or grade) could provide information on the subgroups of children benefiting the most from such exergaming interventions. In addition to student characteristics, it might be useful to consider teacher characteristics (eg, attitudes toward exergaming) to better understand the individual and environmental factors likely to moderate the effects of such interventions.

Conclusions

As part of the 30-minute implementation of DPA, the use of learning-based exergaming showed very interesting results in increasing PL as well as student motivation toward mathematics. Furthermore, supporting pedagogical teams with qualified teachers in PA has been proven to be beneficial for both students and staff.

With this encouraging pilot study, it is necessary to continue investigations by increasing the number of students per grade and to carry out research over a longer school period with a control group to confirm these results regarding the use of learning-based exergaming with Play LÜ within the framework of DPA.

Acknowledgments

We would like to thank the school (administration and teaching staff) for allowing us to carry out this pilot study.

Authors' Contributions

AG conceptualized and supervised the project, and contributed to the data collection and processing, project administration, and writing of the original draft. FL, GB, EVC, and EC conceptualized and supervised the project and were involved in reviewing and editing the manuscript. FL contributed to the statistical analysis. AP was involved in daily physical activity implementation supervision and physical literacy during the session. ED and NAB were students in adapted physical activity who carried out the daily physical activity sessions as part of their end-of-year internship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PlaySELF adaptation questionnaire.

[PDF File (Adobe PDF File), 255 KB - games_v12i1e53072_app1.pdf]

Multimedia Appendix 2
Effects of the exergaming program according to grade.
[DOCX File , 39 KB - games_v12i1e53072_app2.docx]

Multimedia Appendix 3
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1325 KB - games_v12i1e53072_app3.pdf]

References

1. de Greeff JW, Bosker RJ, Oosterlaan J, Visscher C, Hartman E. Effects of physical activity on executive functions, attention and academic performance in preadolescent children: a meta-analysis. J Sci Med Sport 2018 May;21(5):501-507. [doi: 10.1016/j.jsams.2017.09.595] [Medline: 29054748]
2. Neil-Sztramko SE, Caldwell H, Dobbins M. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. Cochrane Database Syst Rev 2021 Sep 23;9(9):CD007651 [FREE Full text] [doi: 10.1002/14651858.CD007651.pub3] [Medline: 34555181]
3. Rasberry CN, Lee SM, Robin L, Laris BA, Russell LA, Coyle KK, et al. The association between school-based physical activity, including physical education, and academic performance: a systematic review of the literature. Prev Med 2011 Jun;52 Suppl 1:S10-S20. [doi: 10.1016/j.ypmed.2011.01.027] [Medline: 21291905]
4. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
5. Verdot C, Salanave B, Deschamps V. Physical activity and sedentary behaviour in French population. Situation in 2014-2016 and evolution since 2006-2007 [Activité physique et sédentarité dans la population française. Situation en 2014-2016 et évolution depuis 2006-2007]. Bull Epidémiol Hebd 2020(15):296-304.
6. La stratégie nationale de santé 2018-2022. Ministère du Travail, de la Santé et des Solidarités. 2017. URL: https://sante.gouv.fr/systeme-de-sante/strategie-nationale-de-sante/article/la-strategie-nationale-de-sante-2018-2022 [accessed 2023-10-31]
7. 30' d'activité physique quotidienne. Génération 2024. URL: https://generation.paris2024.org/30-dactivite-physique-quotidienne [accessed 2023-08-30]
8. Messing S, Rütten A, Abu-Omar K, Ungerer-Röhrich U, Goodwin L, Burlacu I, et al. How Can Physical Activity Be Promoted Among Children and Adolescents? A Systematic Review of Reviews Across Settings. Front Public Health 2019 Mar 19;7:55 [FREE Full text] [doi: 10.3389/fpubh.2019.00055] [Medline: 30941342]
9. Promoting physical activity through schools: a toolkit. World Health Organization. 2021. URL: https://apps.who.int/iris/handle/10665/350836 [accessed 2024-02-04]
10. Mavilidi M, Okely AD, Chandler P, Cliff DP, Paas F. Effects of Integrated Physical Exercises and Gestures on Preschool Children's Foreign Language Vocabulary Learning. Educ Psychol Rev 2015 Aug 5;27(3):413-426. [doi: 10.1007/s10648-015-9337-z]
11. Donnelly JE, Hillman CH, Castelli D, Etnier JL, Lee S, Tomporowski P, et al. Physical Activity, Fitness, Cognitive Function, and Academic Achievement in Children: A Systematic Review. Med Sci Sports Exerc 2016 Jun;48(6):1197-1222 [FREE Full text] [doi: 10.1249/MSS.0000000000000901] [Medline: 27182986]
12. Donnelly J, Hillman C, Greene J, Hansen D, Gibson C, Sullivan D, et al. Physical activity and academic achievement across the curriculum: Results from a 3-year cluster-randomized trial. Prev Med 2017 Jun;99:140-145 [FREE Full text] [doi: 10.1016/j.ypmed.2017.02.006] [Medline: 28193490]
13. Schmidt M, Benzing V, Wallman-Jones A, Mavilidi M, Lubans DR, Paas F. Embodied learning in the classroom: Effects on primary school children's attention and foreign language vocabulary learning. Psychology of Sport and Exercise 2019 Jul;43:45-54. [doi: 10.1016/j.psychsport.2018.12.017]
14. Hwang M, Choi H, Lee A, Culver J, Hutchison B. The Relationship Between Self-Efficacy and Academic Achievement: A 5-Year Panel Analysis. Asia-Pacific Edu Res 2015 May 7;25(1):89-98. [doi: 10.1007/s40299-015-0236-3]
15. Quílez-Robres A, Moyano N, Cortés-Pascual A. Motivational, Emotional, and Social Factors Explain Academic Achievement in Children Aged 6–12 Years: A Meta-Analysis. Education Sciences 2021 Sep 07;11(9):513. [doi: 10.3390/educsci11090513]
16. Vazou S, Gavrilou P, Mamalaki E, Papanastasiou A, Sioumala N. Does integrating physical activity in the elementary school classroom influence academic motivation? International Journal of Sport and Exercise Psychology 2012 Dec;10(4):251-263. [doi: 10.1080/1612197x.2012.682368]
17. Dishman RK, Motl RW, Saunders R, Felton G, Ward DS, Dowda M, et al. Self-efficacy partially mediates the effect of a school-based physical-activity intervention among adolescent girls. Prev Med 2004 May;38(5):628-636. [doi: 10.1016/j.ypmed.2003.12.007] [Medline: 15066366]

18. Nathan N, Elton B, Babic M, McCarthy N, Sutherland R, Presseau J, et al. Barriers and facilitators to the implementation of physical activity policies in schools: A systematic review. Prev Med 2018 Feb;107:45-53. [doi: 10.1016/j.ypmed.2017.11.012] [Medline: 29155228]
19. International Physical Literacy Association. URL: https://www.physical-literacy.org.uk/ [accessed 2023-07-19]
20. Hulteen RM, Morgan PJ, Barnett LM, Stodden DF, Lubans DR. Development of Foundational Movement Skills: A Conceptual Model for Physical Activity Across the Lifespan. Sports Med 2018 Jul;48(7):1533-1540. [doi: 10.1007/s40279-018-0892-6] [Medline: 29524160]
21. Shearer C, Goss HR, Boddy LM, Knowles ZR, Durden-Myers EJ, Foweather L. Assessments Related to the Physical, Affective and Cognitive Domains of Physical Literacy Amongst Children Aged 7-11.9 Years: A Systematic Review. Sports Med Open 2021 May 27;7(1):37 [FREE Full text] [doi: 10.1186/s40798-021-00324-8] [Medline: 34046703]
22. Alanko D. The Health Effects of Video Games in Children and Adolescents. Pediatr Rev 2023 Jan 01;44(1):23-32. [doi: 10.1542/pir.2022-005666] [Medline: 36587018]
23. Oh Y, Yang S. Meaningful Play. URL: https://meaningfulplay.msu.edu/proceedings2010/mp2010_paper_63.pdf [accessed 2024-02-05]
24. Gao Z, Hannan P, Xiang P, Stodden DF, Valdez VE. Video game-based exercise, Latino children's physical health, and academic achievement. Am J Prev Med 2013 Mar;44(3 Suppl 3):S240-S246 [FREE Full text] [doi: 10.1016/j.amepre.2012.11.023] [Medline: 23415189]
25. Merino-Campos C, Del Castillo Fernández H. The Benefits of Active Video Games for Educational and Physical Activity Approaches: A Systematic Review. Journal of New Approaches in Educational Research 2016 Jul 15;5(2):115-122. [doi: 10.7821/naer.2016.7.164]
26. Moller AC, Sousa CV, Lee KJ, Alon D, Lu AS. Active Video Game Interventions Targeting Physical Activity Behaviors: Systematic Review and Meta-analysis. J Med Internet Res 2023 May 16;25:e45243 [FREE Full text] [doi: 10.2196/45243] [Medline: 37191992]
27. Lwin MO, Malik S. Can exergames impart health messages? Game play, framing, and drivers of physical activity among children. J Health Commun 2014;19(2):136-151. [doi: 10.1080/10810730.2013.798372] [Medline: 24191779]
28. Rüth M, Kaspar K. Exergames in formal school teaching: A pre-post longitudinal field study on the effects of a dance game on motor learning, physical enjoyment, and learning motivation. Entertainment Computing 2020 Aug;35:100372. [doi: 10.1016/j.entcom.2020.100372]
29. Jamshidifarsani H, Garbaya S, Lim T, Blazevic P, Ritchie JM. Technology-based reading intervention programs for elementary grades: An analytical review. Computers & Education 2019 Jan;128:427-451. [doi: 10.1016/j.compedu.2018.10.003]
30. Liu M, Hsieh P, Cho Y, Schallert D. Middle School Students' Self-Efficacy, Attitudes, and Achievement in a Computer-Enhanced Problem-Based Learning Environment. Journal of Interactive Learning Research 2006;17(3):225-242 [FREE Full text]
31. Best JR. Exergaming in Youth: Effects on Physical and Cognitive Health. Z Psychol 2013 Apr 01;221(2):72-78 [FREE Full text] [doi: 10.1027/2151-2604/a000137] [Medline: 25097828]
32. Gao Z, Zhang T, Stodden D. Children's physical activity levels and psychological correlates in interactive dance versus aerobic dance. Journal of Sport and Health Science 2013 Sep;2(3):146-151. [doi: 10.1016/j.jshs.2013.01.005]
33. Wii Fit Plus. Nintendo. URL: https://www.nintendo.fr/Jeux/Wii/Wii-Fit-Plus-283905.html [accessed 2023-11-02]
34. Lü Interactive Playground. URL: https://play-lu.com/ [accessed 2023-11-02]
35. PLAYself - Physical Literacy Assessment for Youth. Sport for Life Society. URL: https://play.physicalliteracy.ca/playtools/playself/ [accessed 2023-07-18]
36. Jefferies P, Bremer E, Kozera T, Cairney J, Kriellaars D. Psychometric properties and construct validity of PLAYself: a self-reported measure of physical literacy for children and youth. Appl Physiol Nutr Metab 2021 Jun;46(6):579-588. [doi: 10.1139/apnm-2020-0410] [Medline: 33315524]
37. PLAY Tools. Physical Literacy. URL: https://physicalliteracy.ca/play-tools/ [accessed 2023-11-03]
38. Les outils PLAY. Littératie Physique. URL: https://litteratiephysique.ca/les-outils-play/ [accessed 2023-11-03]
39. van Laerhoven H, van der Zaag-Loonen H, Derkx B. A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. Acta Paediatr 2004 Jun 1;93(6):830-835. [doi: 10.1080/08035250410026572] [Medline: 15244235]
40. Just Dance - Nintendo Switch. Nintendo. URL: https://www.nintendo.com/fr-ca/store/products/just-dance-switch/ [accessed 2023-11-06]
41. PlaySport. Ophea. URL: https://ophea.net/playsport [accessed 2024-02-08]
42. Policy/Program Memorandum 138. Ministry of Education, Ontario. URL: https://www.ontario.ca/document/education-ontario-policy-and-program-direction/policyprogram-memorandum-138 [accessed 2023-08-30]
43. Allison K, Philipneri A, Vu-Nguyen K, Manson H, Dwyer J, Hobin E, et al. School and classroom effects on Daily Physical Activity (DPA) policy implementation fidelity in Ontario classrooms: a multi-level analysis. BMC Public Health 2018 Jun 27;18(1):802 [FREE Full text] [doi: 10.1186/s12889-018-5720-2] [Medline: 29945575]

44. Martyn L, Bigelow H, Graham JD, Ogrodnik M, Chiodo D, Fenesi B. A mixed method investigation of teacher-identified barriers, facilitators and recommendations to implementing daily physical activity in Ontario elementary schools. BMC Public Health 2022 Oct 31;22(1):1986 [FREE Full text] [doi: 10.1186/s12889-022-14359-3] [Medline: 36316654]
45. Cairney J, Dudley D, Kwan M, Bulten R, Kriellaars D. Physical Literacy, Physical Activity and Health: Toward an Evidence-Informed Conceptual Model. Sports Med 2019 Mar;49(3):371-383. [doi: 10.1007/s40279-019-01063-3] [Medline: 30747375]
46. Sun H. Operationalizing physical literacy: The potential of active video games. Journal of Sport and Health Science 2015 Jun;4(2):145-149. [doi: 10.1016/j.jshs.2015.03.006]
47. Taras H. Physical activity and student performance at school. J Sch Health 2005 Aug;75(6):214-218. [doi: 10.1111/j.1746-1561.2005.00026.x] [Medline: 16014127]
48. Fyfe ER, McNeil NM, Borjas S. Benefits of “concreteness fading” for children's mathematics understanding. Learning and Instruction 2015 Feb;35:104-120. [doi: 10.1016/j.learninstruc.2014.10.004]
49. Appel à projets pour un socle numérique dans les écoles élémentaires. Ministre de l'Education nationale, de la Jeunesse, des Sports et des Jeux Olympiques et Paralympiques. URL: https://www.education.gouv.fr/bo/21/Hebdo2/MENN2100919X.htm [accessed 2023-11-06]
50. Vanhelst J, Ternynck C, Ovigneur H, Deschamps T. [Tracking of physical fitness during chilhood: Longitudinal findings from the Diagnoform program]. Rev Epidemiol Sante Publique 2020 Jun;68(3):163-169 [FREE Full text] [doi: 10.1016/j.respe.2020.04.052] [Medline: 32417152]

Abbreviations

DPA: daily physical activity
PA: physical activity
PE: physical education
PL: physical literacy

Edited by R Baranyi; submitted 29.09.23; peer-reviewed by W Cools, C Merino Campos; comments to author 30.10.23; revised version received 14.11.23; accepted 28.11.23; published 23.02.24.

Please cite as:
Goncalves A, Lespiau F, Briet G, Vaillant-Coindard E, Palermo A, Decobert E, Allegret-Bourdon N, Charbonnier E
Exploring the Use of a Learning-Based Exergame to Enhance Physical Literacy, Soft Skills, and Academic Learning in School-Age Children: Pilot Interventional Study
JMIR Serious Games 2024;12:e53072
URL: https://games.jmir.org/2024/1/e53072
doi:10.2196/53072
PMID:38393767

Original Paper

Understanding Senior Adults' Needs, Preferences, and Experiences of Commercial Exergames for Health: Usability Study

Yu-Han Wang[1], BSc, PhD

Department of Multimedia Design, National Taichung University of Science and Technology, Taichung City, Taiwan

Corresponding Author:
Yu-Han Wang, BSc, PhD
Department of Multimedia Design
National Taichung University of Science and Technology
No 129, Section 3, Sanmin Road, North District
Taichung City, 404336
Taiwan
Phone: 886 422196699
Email: han@gm.nutc.edu.tw

Abstract

Background: Many senior adults are at risk of mental and physical disorders due to a lack of sufficient exercise. Therefore, adherent exercise should be urgently promoted to improve senior adults' muscle strength, preventing falls and conditions caused by physical and cognitive decline. However, off-the-shelf exercise games, so-called exergames, are mainly targeted at the younger generation or children, while senior adults are neglected, when this age group strongly needs exercise. Exergames could serve as a health intervention for promoting exercise.

Objective: This study aimed to investigate senior adults' experience, perceptions, and acceptance of game technology to promote exercise in order to suggest game design guidelines.

Methods: In this usability study, participants engaged in playing Nintendo Switch and Xbox Kinect games, after which semistructured interviews were conducted. Before the gameplay, the participants provided their background information, exercise habits, and use of technology products. Next, all participants completed a workshop including 3 activities (brief instructions on how to play the games: 20 minutes; playing the selected exergames: 80 minutes; semistructured interviews: 20 minutes) for 2 hours a day for 3 days each. The participants played the latest Nintendo Switch games (eg, Just Dance, Boxing, Ring Fit Adventure) and Xbox Kinect games (eg, Kinect Adventures!, Mini Games). Just Dance, Zumba, and Boxing were played in activity 1; Ring Fit Adventure and Mini Games in activity 2; and Kinect Adventures! in activity 3. Reflexive thematic analysis was applied to identify the relative themes generated from the interviews.

Results: In total, 22 participants (mean age 70.4, SD 6.1 years) were enrolled in the workshop in May 2021. The results of the generated themes included incomprehension of game instructions, psychological perception of game technology, and game art preferences. The subthemes generated from game art preferences included favorite game genres, characters, and scenes.

Conclusions: There is a significant need for customized game tutorials considering senior adults' cognitive and physical aging. Furthermore, the adventure game genre is preferable to other games. Humanlike game characters are preferable, especially those with a fit and healthy body shape. Nature scenes are more enjoyable than indoor stages or rooms. Furthermore, the game intensity design and playing time should be carefully planned to meet the World Health Organization's criteria for physical activity in older adults. Intelligent recommendation systems might be helpful to support older adults with various health conditions. The guidelines suggested in this study might be beneficial for game design, exercise training, and game technology adoption of exergames for older adults to improve health.

(JMIR Serious Games 2024;12:e36154) doi:10.2196/36154

KEYWORDS

exergame; senior user experience; senior technology acceptance; game technology; psychological perception; serious games; exercise; aging; older adults; physical activity

Introduction

Background

The rapid growth of the aging population has drastically increased the need for medical and health care services. The life expectancy is increasing; however, the average life in a state of good health, the so-called healthy life expectancy, is almost 8-10 years less than the life expectancy [1,2]. It is important to increase the healthy life expectancy for active aging and reduce the financial burden on health systems caused by incapacitation, bedridden patients, and chronic physical or mental health conditions. The World Health Organization (WHO) [3] has suggested that adults 65 years and older perform 150-300 minutes of moderate-intensity or 75-150 minutes of strenuous-intensity aerobic physical activity in a week in order to improve physical and functional health and lower the risk of noninfectious chronic illness and mental health conditions. However, it has been reported that in Taiwan, 33.8% of older adults in the age group of 65-74 years and about 36% of 55-64-year-old "graying" adults do not exercise at all [4]. Many "graying" adults in Taiwan are at risk of mental and physical disorders due to a lack of sufficient exercise. Therefore, providing preventive health care and promoting adherent exercise are urgent issues to be considered in order to improve senior adults' muscle strength and physical fitness to prevent falls and conditions caused by physical and cognitive decline.

Studies on Exergames for Health

Several studies have highlighted the benefits and positive impact of exercise games, so-called exergames, on overall health. The term "exergame" is most frequently used by health-related researchers to refer to video games that use strength training, balance, or flexibility [5]. Exergaming has gained public and commercial interest due to its combination of fun and fitness together and is particularly popular for obesity-related interventions [6]. Furthermore, it is regarded as the future of fitness to maintain regular activity as it is a healthy, appealing alternative to other physical activities, as seen especially during the COVID-19 quarantine [7]. In addition to increasing physical exercise, exergaming could also be beneficial for psychological, cognitive, or psychosocial issues [8]. Exergames have been gradually applied in psychological treatment or physical rehabilitation for senior adults. For example, some studies have adopted exergames in psychological treatment for cognition [9], dementia [10], and depression [11]. Exergames have also been adapted for physical treatment to improve physical skills or fitness [12,13], balance [14,15], fall prevention [16], and rehabilitation [17-19]. Although many studies have applied off-the-shelf exergames for mental and physical health, with proof of their benefits and effectiveness, older adults' perceptions, experiences, and expectations are rarely discussed. Therefore, it is important to investigate older adults' experiences and expectations of playing exergames as a preventive approach to promote exercise habits.

Exergame Technology

Exergames require players to interact with the virtual gameplay by means of their physical movements and gestures in the real world [20]. Exergames generally use motion-sensing technology, which tracks and monitors players' movements while they perform game tasks using sensory devices, such as Nintendo's Joy-Con or Microsoft's Xbox Kinect. Motion-sensing technology differs between Nintendo and Xbox consoles. The Joy-Con contains an accelerometer and a gyroscope built into the device to track a player's motion. Additionally, the Joy-Con also contains 3D touch technology, which provides subtle vibration feedback to the player. In contrast, Xbox Kinect, which is a camera-based and infrared depth-scanning-approach device, enables a player to interact with the game with body movements. Various types of game consoles with different underlying body motion technologies offer different user experiences, but the amount of physical effort and energy required to play exergames remain similar [21]. Accessibility and entertainability are the 2 distinguishing factors of exergames in promoting exercise and activity [22]. This study explored which interactive technology is preferred by older adults.

Technology Acceptance of Older Adults

For decades, many studies have researched how people adopt technology [23-26]. Davis [23] first developed the technology acceptance model to explain user behavior and the intention to use technology. Chen and Chan [26] further modified the technology acceptance model, calling their modified version the senior technology acceptance model, to learn about senior users' adoption of technology. The term "older adults" generally refers to people whose chronological age is 60 years or older [27]. Researchers suggest that older adults' attitude toward using and their intention to use technology are particularly affected by their self-efficacy, anxiety, health conditions, cognitive ability, social relationships, attitude toward life and satisfaction, physical functioning, and the support of technology use they receive around them. The senior technology acceptance model has been applied to understand older adults' behavior with regard to technology in Hong Kong [26], Sweden [28], and Taiwan [29].

According to a statistic survey by the American Association of Retired Persons (AARP), in 2019, in the United States, 44% of older adults enjoyed playing video games compared to 38% in 2016 [30]. The growing population of older adults in the game market shows that their needs should be considered in game design. However, many older adults may not have had much experience playing video games during their childhood, as computer-based video games became popular around the 1980s [31]. The target users of these exergames are mostly younger adults or children. The aging population is neglected in the market, when this age group strongly needs exercise. Therefore, this study investigated older adults' experience, perceptions, and acceptance of game technology to promote routine exercise.

Methods

Study Design and Recruitment

The workshop was conducted in a spacious classroom that was transformed to accommodate exergame testing. The classroom was large enough to ensure there were no obstacles to the participants performing their movements and gestures while playing the games. Four Sanlux 55-inch 4K Ultra HD televisions with Nintendo consoles and Xbox Kinect consoles were set up in the classroom, and each television equipped with the consoles was shared by 2 participants while playing the selected exergames. The participants had sufficient space to move around and had full range of motion with the whole body. The classroom accommodated 6-8 participants in each workshop, and 3 workshops were conducted in total.

The inclusion criteria for the workshop and interview were participants aged 60 years and older who could exercise and understand the instructions provided. A total of 22 senior adults (≥60 years old) were enrolled in this study.

Gameplay

Each of the 3 workshops included 3 activities (Figure 1). Every activity was conducted for 2 hours a day for 3 days. Before playing the games, a questionnaire was administered, in which the participants provided their background information, exercise habits, and use of technology products. The following games were selected for this study: the latest Nintendo Switch games *Just Dance*, *Boxing*, *Mini-Games*, and *Ring Fit Adventure* and the Xbox Kinect game *KinectAdventures!* Before each activity, the participants were provided with brief instructions for 20 minutes on how to play the games. Next, they played the selected exergames for 80 minutes, and finally, a 20-minute semistructured interview was conducted. The participants played *Just Dance* and *Boxing* in activity 1, *Ring Fit Adventure* and *Mini-Games* in activity 2, and *Kinect Adventures!* in activity 3.

Figure 1. Participants played Boxing (left) and Ring Fit Adventure (right).

Data Analysis

Data were collected from the semistructured interviews using audio recordings and transcribed verbatim. Qualitative data were analyzed using thematic analysis, a qualitative data analytics method proposed by Braun and Clarke [32,33] and further clarified and defined as reflexive thematic analysis. Reflexive thematic analysis is a common approach to identify, analyze, and report themes from qualitative data sets in content analysis and grounded theory [34]. Braun and Clarke [35] have provided a 6-phase process of thematic analysis for researchers to conduct qualitative studies: (1) familiarizing with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing potential themes, (5) defining and naming themes, and (6) producing a report. The 6-phase process is not linear but an iteratively developing process that is flexible enough to be tailored for qualitative data [36]:

- Phase 1 (familiarizing with the data): In phase 1, the researcher read and reread the data and made notes on a manual transcript. The researcher also identified connections between participants, data, and existing studies.
- Phase 2 (generating initial codes): The interview transcript was systematically analyzed through coding. Some codes matched the participants’ concepts, while others had to be interpreted by the researcher.
- Phase 3 (searching for themes): All the coded data were identified and clustered into broader topics or themes based on their similarities and overlaps.
- Phase 4 (reviewing potential themes): All the generated themes were reviewed to check whether they worked meaningfully and relevantly with the coded data.
- Phase 5 (defining and naming themes): After reviewing possible themes, the identified themes were subsequently defined and labeled.
- Phase 6 (producing a report): The generated themes answered the research questions about senior adults’ experience, perceptions, and acceptance of adopting exergames for exercise.

Ethical Considerations

This study was approved by the Central Regional Research Ethics Committee China Medical University Taiwan (review number: CRREC-109-090). All participants were informed of the procedures involved in the study. They agreed to participate in the study and signed the consent form. The data in this study is anonymized. This study offered non-monetary incentives to participants as compensation.

Results

Participant Details

In total, 22 participants (n=16, 73% female; n=6, 27% male) were enrolled in the study. Their ages ranged from 60 to 82 years (Table 1).

Table 1. Participants' general information (N=22).

Characteristics	Value
Age (years), mean (SD)	70.4 (6.1)
Male	68.7 (5.4)
Female	71.0 (6.4)
Sex, n (%)	
Male	6 (27)
Female	16 (73)
Education (years), mean (SD)	11.5 (4.5)
BMI (kg/m^2), mean (SD)	
Male	23.3 (4.5)
Female	22.4 (2.8)
Exercise routine, n (%)[a]	
Everyday	6 (27)
3-5 days a week	12 (55)
1-2 days a week	3 (14)
1-3 days a month	1 (5)
Exercise time (minutes/day), n (%)	
>30	16 (73)
20-30	4 (18)
10-20	1 (5)
<10	1 (5)
Exercise intensity, n (%)	
Barely there	11 (50)
Moderate	10 (46)
Harder	1 (5)
Technology products use habits (possession of smartphones/tablets/computers), n (%)	
I do not have any of them.	1 (5)
I have 1 of them.	8 (36)
I feel comfortable using them.	13 (59)
Frequency of using the technology products, n (%)	
Never	1 (5)
1 time a week	3 (14)
2-3 times a week	0 (0)
Many times a week	1 (5)
1 time a day	2 (9)
Many times a day	15 (68)
Experience of playing videogames/digital games (arcade, home console, handheld game console, computer game, or online game), n (%)	12 (55)

[a]The percentages might add up to more than 100 because of rounding.

The background information of the 22 participants showed that only 7 (32%) met the WHO criteria for weekly exercise. The participants were familiar with phone technology and often used it for social media, photography, Google searching, or online shopping. Some of them had previous experience of playing video games. These results showed that the participants were less active and were comfortable using technology products.

Most participants performed regular or mild light-intensity exercise; however, only a few participants met the WHO criteria for senior adult exercise. WHO recommends multicomponent physical activity at a moderate or high intensity [3]; however, the results showed that the exercise intensity of older adults and the time they spend are not enough.

Thematic Analysis

The semistructured interviews raised topics relating to user experience, psychological perceptions, and the adoption of game technology. In phase 2 of the thematic analysis, the manual transcript was systematically analyzed through coding. For example, "The tutorials in the game were not clear and not easy to understand. It is better to have instructors to teach me how to operate the game" was coded as the game tutorials not being supportive, being difficult to understand, and causing the participant to lose the motivation to play. In phase 3, the coded data was grouped into larger themes based on how they were connected. For example, coded data such as "do not know how to play," "not understanding," and "forgot tutorials" were clustered into "incomprehension of game instructions." In phase 4, the identified themes were carefully reviewed to ensure they accurately and meaningfully represented the coded data. In phase 5, the themes were defined and named as (1) incomprehension of game instructions, (2) confusion caused by a complicated interface, (3) frustration caused by the fast game speed, (4) a sense of control and freedom, (5) psychological perception of game technology, (6) social interaction, and (7) preference of game art. These 7 themes could answer the research questions about older adults' experience, perceptions, and acceptance of exergaming. The first 6 themes and the participants' feedback are shown in Table 2.

Table 2. Themes 1-6 generated from the interviews.

Theme and example quotes	Sex (male/female), age (years), game experience (yes/no)
1. Incomprehension of game instructions (n=8, 36%)	
“It is very critical that someone could guide us before playing games [*Just Dance*]. Knowing how to play motivates me to play the game.”	Female, 74, no
“I need someone next to me and tell me how to play.	Female, 67, yes
“I don’t know how to play the game because I don’t understand the game mechanics and rules.”	Male, 68, yes
“I need a detailed game tutorial of the game to support me when playing the game.”	Female, 75, no
“The tutorial in the game is not clear and not easy to understand. It is better to have instructors to teach me how to operate the game [*Ring Fit Adventure*].”	Male, 62, yes
“By the game instructions, I cannot fully understand what it means and how to play. I prefer someone (instructors) to teach me for better understanding.”	Male, 72, yes
“I can't remember things well. So, I forgot tutorials while playing. I might remember how to play after many times gameplay.”	Female, 79, yes
“I need more practice to remember how to play.”	Male, 76, yes
2. Confusion caused by a complicated interface (n=5, 23%)	
“At the beginning, it was difficult to understand the game interface, but after practicing, I got used to controlling it.”	Female, 66, no
“I don’t know the meanings of the icons in the game interface.”	Female, 74, no; male, 76, yes
“There is too much information on the screen; I don’t know what to look at while playing [*Ring Fit Adventure*].”	Female, 75, no
“The game interface was very complicated and difficult to use. I guess it is also difficult to control machines for most of the older adults [*Just Dance*].”	Female, 63, no
3. Frustration caused by the fast game speed (n=8, 36%)	
“I would give up and feel frustrated when I cannot keep up with the speed of the games.”	Female, 82, yes
“I wish I could control the speed of the game from slow to fast so that I could participate in the game, not be excluded since the very beginning.”	Female, 63, no
“I was too slow in the response in the playing. For example, I wanted to jump and get the golden coin, but when I jumped, the golden coin disappeared [*Kinect Adventures!*].”	Female, 78, yes; female, 75, no
“It is very difficult and complicated to do more than 2 actions at the same time. I have to avoid getting bombed and to get golden coins [*Ring Fit Adventure*].”	Female, 74, no
“I feel I am not as agile as I was. The game should provide different levels of difficulty for different ages.”	Female, 79, yes
“After seeing the boxing icon showing up from the bottom of the screen, I need to punch at the right tempo. But I was always too late to react. I need some time to think and then do the action [*Boxing*].”	Female, 66, no
“The speed in the game was too fast for me. I think the game should provide different levels of difficulty or provide different levels according to different age groups [*Boxing*].”	Female, 79, yes
4. A sense of control and freedom (n=4, 18%)	
“I like to use joystick to control games because of the tactile responses, although I feel free when I don’t need to hold anything to control the game. However, I don’t receive tactile sensation feedback to know whether I did the correct movement or not [*Kinect Adventures!*].”	Female, 60, yes
“I like to use my body motion to control the game because I don’t like to hold anything on hands. I don’t know how to use joystick, and I feel it is quite complicated. It is easy, free, and intuitive to control the game by my body.”	Female, 63, no
“I prefer the realistic feeling of pressing and dragging the Ring-Con. It gives me real feedback and is easy to use.”	Male, 71, no
“The vibration of the joystick gives me feedback of correct movement, and I like to receive the feedback [*Boxing*].”	Female, 74, no
5. Psychological perception of game technology (n=9, 41%)	
“Playing the games could make me keep focus, train my brain and coordination.”	Female, 66, no

Theme and example quotes	Sex (male/female), age (years), game experience (yes/no)
"It is so much fun to do exercise by playing games. I didn't expect myself to enjoy doing exercises [*Boxing*, *Kinect Adventures!*]."	Male, 68, yes
"I prefer to play *Boxing* game because it is very challenging."	Male, 62, yes
"I feel it is fashionable to play state-of-the-art games, which could connect me and the younger generations."	Female, 82, yes
"I feel I am keeping up with trends because I can operate game technology device and share with my grandchildren."	Female, 74, no
"I didn't know games could so charming and interesting. It is a fresh experience, and it is quite fashion and trendy to play games for me."	Female, 79, yes
"I would like to try new things, and I wish I could keep up with the times."	Female, 74, no
"I was afraid of new technology in general. I don't know how to use them. But if I have a chance to approach it and someone could teach me, I will be happy to learn."	Female, 60, yes
I feel video games are not for us; it is for the youth.	Female, 82, yes
6. Social interaction (n=5, 23%)	
"It has become a common interest between me and my grandchildren. We can play together, and we have common topics to talk about."	Female, 75, no
"Collaborating with other players is very fun, and it attracts me to continue using exergames."	Female, 69, yes
"I am feeling that I don't want to lose against my friends. I will make every effort to win."	Male, 68, yes
"I felt lonely when I did exercise alone, so I like exergames, which allow me to do exercise with people."	Female, 68, no
"I enjoyed playing with friends who are the same age as me. It's much more fun to play games or do exercise with friends than by myself."	Male, 63, yes

Additionally, the subthemes of game art preference were generated from the "preference of game art" theme. The qualitative data related to game art were collated and coded into game genres, characters, and scenes (Table 3). In the interviews, participants talked about their favorite games from among *Just Dance*, *Zumba*, *Boxing*, *Kinect Adventures!*, *Mini Games*, and *Ring Fit Adventure*. Adventure games were constantly mentioned as participants' favorite games (n=16, 73%, participants): 11 (69%) participants preferred *Kinect Adventures!*, and 5 (31%) participants liked *Ring Fit Adventure*. Only 3 (14%) participants liked the *Zumba* dance game and 3 (14%) liked *Boxing*. Thus, most participants were fond of the adventure game genre compared to the sports genre, such boxing or dance. Most participants (n=13, 59%) preferred true-to-scale 3D human models as game characters instead of cartoons. Some participants (n=8, 36%) preferred to see a real man on the screen to lead them in exercise, that is, the participants preferred the game characters to be real-scale humans or virtual models. Nearly all participants (n=21, 95%) mentioned that they enjoyed and immersed themselves in outdoor or natural scenes in the games, while only 1 (5%) participant preferred a static virtual stage.

Table 3. Subthemes generated from theme 7 (preference of game art; n=8, 36%, participants).

Subtheme and example quotes	Sex (male/female), age (years), game experience (yes/no)
7.1. Favorite game genre	
“My favorite game is adventure games. It made me focus and immersed myself in the gameplay. The adventure game is very exciting, which could also train my brain and improve my coordination. I don’t usually have the experience in my daily life [*Kinect Adventures!*].”	Female, 74, no
“The *Ring Fit Adventure* provides a more intense and challenging exercise for me. I love to take the challenge.”	Male, 76, yes
7.2. Preference of game character	
“I like the game character is made by 3D humanlike model with a realistic ratio, such as the character in *Ring Fit Adventure*. The character looks fit, muscular, and powerful.”	Female, 75, no
“I prefer a real man in the game, such as the coaches in *Zumba*. They are real people, who make it easier to understand and learn their dancing.”	Female, 69, yes
7.3. Preference of game scene	
“I like the game with outdoor natural scenes. I feel like I was doing exercise outdoor rather than in an indoor room.”	Female, 69, yes
“I was very excited and fully immersed in the game with a natural scene, such as *Xbox Adventure!* or *Ring Fit Adventure*.”	Male, 63, yes
“The natural background made me feel comfortable, and I wanted to go forward to see what shows up next. It also keeps me focus on the gameplay [*Kinect Adventures!*].”	Female, 69, yes
“I prefer static stages in the background because then I can focus on my game tasks [*Boxing*].”	Female, 60, yes

Themes Generated From the Interviews

A total of 7 themes were generated from the semistructured interviews. Of these themes, 3 (43%; themes 1, 5, and 7) were initially identified in the study. The remaining 4 themes (57%; themes 2, 3, 4, and 6) aligned with prior research findings and are discussed in the *Comparison With Prior Work* section.

Theme 1: Incomprehension of Game Instructions

Lacking gaming experience in their youth might result in senior adults not understanding the game mechanics, the meaning of visual effects or reward icons, health points (HPs), or experience points (EPs). The instructions in the gameplay were not easy to understand, so most of the participants perceived them negatively. Moreover, although some participants had game experience, it did not seem to support their learning in exergames. They still needed appropriate instructions for gameplay learning.

As the game instructions were not helpful to the participants, they needed instructors to explain and teach them again to learn how to play. However, although the participants were taught how to play before the gameplay, they forgot how to play when playing. It was not easy for some participants to understand and memorize the given information. These findings showed that age-related visual and hearing loss might also contribute to increasing the barriers to understanding the gameplay. Thus, repetitive and real-time tutorials in games might be helpful to support memory, learning, and thinking. Therefore, simple visual instructions and oral guidance using plain language might support senior adults’ game learning.

Theme 2: Confusion Caused by a Complicated Interface

In general, younger adults tend to find game interfaces more intuitive and user friendly compared to senior adults with regard to playing digital games. Senior adults often encounter significant challenges when attempting to engage with game interfaces. For instance, *Ring Fit Adventure* integrates a diverse range of physical tasks within its exercise regimen. Players are required to perform actions such as pressing or dragging the ring with their hands while simultaneously engaging in activities such as jumping, squatting, or running with their feet. The information and tasks presented in the game interface pose a complexity that senior adults may find challenging to comprehend. For instance, 1 (5%) participant (female, age 75 years, nonexperienced player) said:

> *There is too much information on the screen, I do not know what to look at while playing.*

Another participant (female, age 63 years, nonexperienced player) said:

> *The game interface is very complicated and difficult to use.*

Furthermore, senior adults often struggle to grasp the meaning of icons or virtual objects displayed in the game interface. For instance, they may not be aware that collecting golden coins in *Ring Fit Adventure* can lead to earning additional points. Two participants expressed:

> *I do not know the meanings of the icons in the game interface.*

Theme 3: Frustration Caused by the Fast Game Speed

One participant expressed frustration and a sense of exclusion when unable to keep up with the game’s pace. They desired the ability to control the game's speed, allowing for a more inclusive gaming experience. The participant (female, age 66 years, nonexperienced player) playing *Boxing* noted that her ability to perform actions effectively was affected by slower response times. She explained:

> *After seeing the boxing icon showing up from the bottom of the screen, I need to punch at the right tempo. But I was always too late to react. I need some time to think and then do the action.*

Another participant (female, age 82 years, experienced player) said:

> *I would give up and feel frustrated when I cannot keep up with the speed of the games.*

In addition, participants perceived a decrease in agility compared to their previous exercise experiences, which could be attributed to the natural decline associated with aging. Thus, an inappropriate game speed could result in senior adults feeling frustrated and avoiding playing exergames.

The American College of Sports Medicine [37] has classified the characteristics of physical fitness into health-related and skill-related components. The skill-related components of physical fitness include agility, coordination, balance, power, reaction time, and speed, which can be adjusted in-game to meet individual conditions. The results in this study showed that game speed levels should be adjustable according to given skill-related health conditions. Senior adults need more time to understand what is going on and how to react accordingly. Therefore, games should provide a range of difficulty levels specifically designed to accommodate different age groups, thereby ensuring an enjoyable gaming experience for senior adult players.

Theme 4: A Sense of Control and Freedom

Participants shared diverse preferences regarding their choice of gaming controls. Some appreciated the tactile responses provided by a joystick (n=3, 14%) or the Ring-Con (n=7, 32%), while others found freedom in not having to hold anything when controlling the game, such as Kinect motion-sensing interaction (n=12, 55%). Alongside their appreciation for the sense of freedom, 7 (32%) participants favored the Ring-Con because it resembles holding a steering wheel, thus enhancing their perception of control.

However, the absence of tactile feedback in some games left participants uncertain about the accuracy of their movements. For example, 1 (5%) participant said:

> *Although I feel free when I do not need to hold anything to control the game...I do not receive tactile sensation feedback to know whether I did the correct movement or not.*

Theme 5: Psychological Perception of Game Technology

Playing exergames was perceived to relate to enjoyment, socializing, achievement, frustration, defeat, and keeping up with trends. Generally, participants enjoyed playing and socializing with people. In addition, playing exergames and operating technological devices made them feel trendy.

> *I feel I am keeping up with trends because I can operate [a] game technology device and share with my grandchildren.* [Female, 74 years]

Some participants felt frustrated when they could not follow the game speed or obtained lower scores due to their longer reaction time and lack of coordination.

> *I want to have a feeling of achievement, rather than being defeated.*

Therefore, the game design should consider senior adults' psychological perceptions to meet their emotional needs.

Theme 6: Social Interaction

The use of exergames fostered a sense of connection and engagement, both within the family and among peers.

> *I enjoyed playing with friends.* [Male, age 63 years, experienced player]
>
> *I like to play with my family, and it is the biggest motivation to play.*

Some participants enjoyed the collaboration and competition with team players. For instance, 1 (5%) participant (female, age 69 years, experienced player) expressed:

> *Collaborating with other players is fun, and it attracts me to continue using exergames.*

These findings showed that playing exergames serves as a common interest that allows for enjoyable collaborative experiences, providing shared topics for discussion, and enhancing the motivation to continue using exergames.

Theme 7: Preference of Game Art

The game preference for game art and game genres was investigated. The exergames in the workshop included the latest Nintendo Switch games, such as *Just Dance* and *Ring Fit Adventure*, and Xbox Kinect games, such as *Kinect Adventures!* Among these games, 16 (73%) participants were fond of adventure games, including *Ring Fit Adventure* and *Kinect Adventures!* Interestingly, the results showed that younger senior adults (6/16, 37.5%, participants; average age 65.8, SD 5.7 years) preferred *Ring Fit Adventure*, which is a resistance exercise and a high-intensity training game, more than older senior adults (10/16, 62.5%, participants; average age 71.4, SD 5.3 years). Other participants (n=6, 27%) preferred less intensive exercise games, such as *Boxing* and *Zumba*. Of these 6 participants, 5 (83%) had no previous game experience, and the average age was 73.2 (SD 6.2) years. Therefore, younger senior adults might prefer a more intensive resistance exercise, while older senior adults might prefer a gentle exercise with less leg work.

For game characters, most of the participants preferred a human or humanlike character because it was easier to understand their movements. The participants also could reflect themselves as the avatar if the game character looked like a human. Of the 22 participants, 21 (96%) preferred human or humanlike characters, of which 13 (62%) participants preferred a humanlike game character and 8 (38%) preferred real humans as game characters. One participant did not show her preference. These findings showed that among the cute animal characters and human (humanlike) characters, senior adults prefer human or humanlike avatars, which allows them to easily reflect themselves in games and observe exercise movements more clearly.

For game scenes, outdoor nature scenes were the most preferable, such as scenes in *Kinect Adventures!* and *Ring Fit Adventure*. Of the 22 participants, 20 (91%) preferred nature scenes, followed by static scenes. The participants felt the nature scenes made them feel comfortable and helped them focus. Two participants preferred simple and clean colored indoor scenes rather than a sophisticated background, which allowed them to focus on the game mission and not be distracted by the background.

Therefore, adventure games, humanlike characters, and nature scenes could create a sense of reality and players might be more easily immersed in the gameplay.

Discussion

Principal Findings

In this study, most of the participants had no or little experience of playing exergames. However, after playing exergames in the workshop, they found them interesting and appealing. The results of this study reflect factors of the senior technology acceptance model [26]. Most of the participants referenced a positive attitude, perceived usefulness, and social relationships regarding using game technology. However, the participants provided negative feedback for the perceived ease of use and support of technology use. Lacking experience of playing video games might have also resulted in the participants not having the knowledge and skills needed to adopt game technology.

The supports for technology use, such as understandable tutorials or game mechanics, to help senior adult players are not sufficient. Thus, there is a significant need for customized instructions for senior adults. Overall, the reasons the participants did not have the intention to play could be because they thought video games are for younger people, not for them. Therefore, exergames should also meet senior adults' psychological needs to increase their adoption of game technology.

The results of this study present older adults' experience, perceptions, and acceptance of commercial exergames. In total, 7 themes were generated from the semistructured interviews; 4 of these are in line with previous studies, but 3 were identified in this study.

Three distinguished themes generated from this study are incomprehension of game instructions, psychological perception of game technology, and preferences of game art. These themes have rarely been discussed in prior studies. The psychological perception toward exergames that the participants expressed was that it is fashionable and trendy to play the latest digital games, which belong to the youth. They also enjoyed the gameplay by collaborating and competing with their peers. Among the exergames, the participants preferred adventure games much more than other genres. In this study, the adventure games were *Kinect Adventures!* and *Ring-Fit Adventure*, both of which focus on players' action tasks. These types of games typically emphasize storytelling and character development, as players assume the role of a principal character who must overcome various challenges and obstacles to reach their objectives. The exciting experience of outdoor sports motivated them to continue to play, and the virtual environment provided a safe place to perform thrilling activities.

The game characters and scenes are also critical visual elements to immerse senior adult players in games. In this study, senior adults preferred humanlike characters who looked healthy, fit, and muscular. The game character with a healthy outlook inspired senior adult players to reflect themselves as energetic people exercising in the gameplay. The nature scenes in the game also gave the senior adult players feelings of comfort and relaxation during the gameplay. WHO recommends that adjustable intensity and game levels should be carefully designed to support senior adults perform a sufficient amount of physical activity.

However, if senior adults cannot understand how to play a game or react in time, they feel defeated and frustrated. They may give up playing straight away. Therefore, there is a significant need for game tutorials that are easy to understand for senior adults and game levels that are suitable for their mobile ability.

Comparison With Prior Work

Four themes generated from the qualitative data demonstrated similar issues as in previous studies. First, "confusion caused by a complicated interface" (theme 2) and "frustration caused by the fast game speed" (theme 3) are in line with the results of Aarhus et al [38], who found that simultaneously increasing information, speed, and colors would more likely increase cognitive challenges. Although Aarhus et al [38] adopted Nintendo Wii in a physical rehabilitation context, the issue remains in the current Nintendo Switch. Second, the game speed was too fast to follow for most of the participants, which is in line with the results of Brox et al [39]. From the cognitive ability perspective, an age-related decrease in working memory causes a reduction in the amount and speed of information processing [35]. Third, the participants in this study highlighted "a sense of control and freedom" (theme 4) in the gameplay, similar to the results of Thin et al [21], who found that the game experience of motion-sensing games is preferrable due to its greater freedom and holistic movement experience. Fourth, "social interaction" (theme 6) is a prominent motivation for senior adults to adopt exergames, which is in line with previous studies [38]. Therefore, social interactions with peers, friends, the family, and society formed through playing exergames are appreciated.

Previous studies have also discussed the aforementioned themes, although most of the commercial games tested in previous studies were Nintendo Wii, Wii Fit, Wii Balance Board, and Wii Sports. Compared to prior studies, this study adopted the latest Nintendo Switch and Xbox Kinect to investigate the player experience. However, obstacles, such as game speed, interface, and tasks, remained, indicating that the latest games still do not consider the needs of senior adults.

Evaluation of This Study

According to the criteria for scientific rigor (credibility, dependability, confirmability, and transferability), in the qualitative research proposed by Lincoln and Guba [40], consistent outcomes are expected when replicating the research process within the same setting. To enhance credibility in this study, continuous interaction was maintained with each participant throughout the data collection process. Furthermore, the participants were encouraged to provide examples while discussing their gaming behavior and experiences, with the interviewer posing follow-up questions. This approach facilitated the participants' familiarity with both the research setting and the content, thereby ensuring an accurate interpretation of their original perspectives. Regarding transferability, which pertains to applicability, the researcher provided detailed descriptive data, such as participants' demographics, exercise characteristics, technology use habits, inclusion criteria of recruitment, workshop and interview procedures, and the iterative research process. The information helped the researcher explain the participants' behavior and experience within a gaming context, potentially making them meaningful and transferable to an external observer. To ensure dependability, which is related to consistency, a detailed analysis process was used throughout this study [41], thereby establishing the potential for reproducing the outcomes across similar participant cohorts and settings. Given the constraints, concerning confirmability, which relates to maintaining neutrality, although the absence of an external review was recognized, diligent steps were taken to compensate for this limitation by implementing a rigorous process of self-evaluation and critical reflection on the research process and outcomes. This included a careful examination of the data, consistent cross-referencing with established literature, and self-awareness regarding potential biases.

Strengths and Limitations

This study makes a noteworthy contribution by emphasizing the importance of game tutorials, preferences of game art and genres, and the perception of trendiness in the game design for older adults. However, the study does have limitations. First, it is essential to acknowledge that the sample size was restricted to 22 participants, consequently impacting the precision of estimates for main outcomes. Second, evaluating the qualitative study's confirmability would ideally involve external researchers. The advantage of having a single researcher code the content is that it ensures consistency and stability in the use of codes; however, confirmation bias can be manifest during various stages of the research process, such as data coding or interpretation [42]. As previously stated, a single researcher may hold preconceived ideas or preferences about the outcomes they anticipate or desire, and these biases can inadvertently impact their scholarly work. Although this study was conducted by a single author, the inherent limitations associated with solo authorship are recognized. The absence of external researchers in this study could limit the diversity of perspectives during data coding and interpretation. This limitation may result in a narrower scope of interpretation and analysis, potentially overlooking valuable insights. By acknowledging potential biases and preconceptions, steps have been taken to minimize their impact on the study's findings. Although collaborative research may not have been feasible in this study, the research process involved continuous cross-referencing with established literature and a sustained awareness of potential biases. It is hoped that this study will contribute valuable insights to the field of digital exergames for senior adults, despite its single authorship limitations. Additionally, it is anticipated that future research will build upon the study's findings to further enhance our understanding of senior adults' experiences in engaging with exergames.

Conclusion

Exergames could serve as an engaging approach to promote exercise and a healthy life among senior adults. Most of the prior studies have focused on usability and facilities, but senior adults' psychological perception toward the exergame experience is highlighted in this study. The findings of this study have some important practical and research implications for adoption of game technology, as well as for research with senior adults' gameplay experience for future work. First, tailored game tutorials for senior adults could be beneficial for increasing the adoption of exergames to promote physical health. Because of insufficient game experience in their youth and cognitive decline due to aging, there is a significant need for understandable and age-friendly tutorials of exergames to equip senior adults with affordable information and skills to get into the game scenario and mechanics. Second, preferences of game art and genres reveal that adventure games are the most favorable game genre and humanlike avatars in nature scenes are most liked among senior adults. Moreover, evidence from this study and the literature shows that the exercise time and intensity of senior adults in Taiwan are clearly not sufficient according to WHO criteria. Thus, game design should plan appropriate game times and intensities for senior adults, which could support them in gradually performing moderate- or high-intensity exercise to promote their health. In addition to game intensity, the game speed of current off-the-shelf exergames is still not suitable for senior adults. Therefore, intelligent recommended systems of game intensity, speed, or difficulty might be helpful for senior adults with various health conditions. Finally, senior adults also want to be fashionable and keep up with trends, not to be excluded by the market, so exergame developers may consider including senior adults' physical and psychological needs to create age-friendly exergames that are more accessible. Future research could focus on investigating age-friendly game tutorials, developing approachable adventure games, creating adjustable game intensity levels, and designing game artwork for senior adults to enhance their exercise for better health.

Acknowledgments

This research was funded under National Sci-Tech Programs by the Ministry of Science and Technology (ROC; MOST 109-2410-H-025-005).

Data Availability

All data generated and analyzed during this study are included in this published paper.

Authors' Contributions

YHW is the sole author of this manuscript, responsible for conducting the research, analyses, and data verifications. She successfully secured research funding from the Ministry of Science and Technology. Her contributions encompass interpreting the results and approving the latest version of the manuscript.

Conflicts of Interest

None declared.

References

1. Life expectancy in Taiwan. Ministry of Health and Welfare. 2020. URL: https://dep.mohw.gov.tw/DOS/lp-1720-113-xCat-1-1-20.html [accessed 2024-03-11]
2. Healthy life expectancy in Taiwan. Ministry of Health and Welfare. 2020. URL: https://dep.mohw.gov.tw/DOS/cp-4231-45165-113.html [accessed 2024-03-12]
3. Physical activity. World Health Organization. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2024-03-12]
4. Ministry of Health and Welfare. 2015 Taiwan longitudinal study on aging. Health Promotion Administration, Ministry of Health and Welfare. Taiwan; 2015 Jan 28. URL: https://www.hpa.gov.tw/EngPages/Detail.aspx?nodeid=1077&pid=6197 [accessed 2024-03-12]
5. Oh Y, Yang S. 2010 Presented at: Meaningful Play 2010; October 21-23, 2010; East Lansing, MI.
6. Elsayyad MM, Abdel-aal NM, Zamzam SSA. Effect of exergaming and resistance training on obesity : a review article. Med J Cairo Univ 2022 Sep 01;90(9):1591-1596. [doi: 10.21608/mjcu.2022.271511]
7. Cortis C, Giancotti G, Rodio A, Bianco A, Fusco A. Home is the new gym: exergame as a potential tool to maintain adequate fitness levels also during quarantine. Hum Mov 2020;21(4):79-87. [doi: 10.5114/hm.2020.94826]
8. Benzing V, Schmidt M. Exergaming for children and adolescents: strengths, weaknesses, opportunities and threats. J Clin Med 2018 Nov 08;7(11):422 [FREE Full text] [doi: 10.3390/jcm7110422] [Medline: 30413016]
9. Piech J, Czernicki K. Virtual reality rehabilitation and exergames-physical and psychological impact on fall prevention among the elderly—a literature review. Appl Sci 2021 Apr 30;11(9):4098. [doi: 10.3390/app11094098]
10. Dauvergne C, Bégel V, Gény C, Puyjarinet F, Laffont I, Dalla Bella S. Home-based training of rhythmic skills with a serious game in Parkinson's disease: usability and acceptability. Ann Phys Rehabil Med 2018 Nov;61(6):380-385 [FREE Full text] [doi: 10.1016/j.rehab.2018.08.002] [Medline: 30193992]
11. Fernandes CS, Magalhães B, Lima A, Nóbrega P, Silva M, Santos C. Impact of exergames on the mental health of older adults: a systematic review and GRADE evidence synthesis. Games Health J 2022 Jul 11;11(6):355-368. [doi: 10.1089/g4h.2021.0229] [Medline: 35819413]
12. Gallou-Guyot M, Mandigout S, Bherer L, Perrochon A. Effects of exergames and cognitive-motor dual-task training on cognitive, physical and dual-task functions in cognitively healthy older adults: an overview. Ageing Res Rev 2020 Nov;63:101135. [doi: 10.1016/j.arr.2020.101135] [Medline: 32768594]
13. Hou H, Li H. Effects of exergame and video game training on cognitive and physical function in older adults: a randomized controlled trial. Appl Ergon 2022 May;101:103690 [FREE Full text] [doi: 10.1016/j.apergo.2022.103690] [Medline: 35066398]
14. Fang Q, Ghanouni P, Anderson SE, Touchett H, Shirley R, Fang F, et al. Effects of exergaming on balance of healthy older adults: a systematic review and meta-analysis of randomized controlled trials. Games Health J 2020 Feb 01;9(1):11-23. [doi: 10.1089/g4h.2019.0016] [Medline: 31800322]
15. Alhagbani A, Williams A. Home-based exergames for older adults balance and falls risk: a systematic review. Phys Occup Ther Geriatr 2021 Feb 08;39(3):241-257 [FREE Full text] [doi: 10.1080/02703181.2020.1867286]
16. Choi SD, Guo L, Kang D, Xiong S. Exergame technology and interactive interventions for elderly fall prevention: a systematic literature review. Appl Ergon 2017 Nov;65:570-581. [doi: 10.1016/j.apergo.2016.10.013] [Medline: 27825723]
17. Gauthier LV, Kane C, Borstad A, Strahl N, Uswatte G, Taub E, et al. Video game rehabilitation for outpatient stroke (VIGoROUS): protocol for a multi-center comparative effectiveness trial of in-home gamified constraint-induced movement therapy for rehabilitation of chronic upper extremity hemiparesis. BMC Neurol 2017 Jun 08;17(1):109 [FREE Full text] [doi: 10.1186/s12883-017-0888-0] [Medline: 28595611]

18. Schaham N, Zeilig G, Weingarden H, Rand D. Game analysis and clinical use of the Xbox-Kinect for stroke rehabilitation. Int J Rehabil Res 2018;41(4):30. [doi: 10.1097/mrr.0000000000000302]
19. Guimarães V, Oliveira E, Carvalho A, Cardoso N, Emerich J, Dumoulin C, et al. An exergame solution for personalized multicomponent training in older adults. Appl Sci 2021 Aug 29;11(17):7986. [doi: 10.3390/app11177986]
20. Altamimi R, Skinner G. A survey of active video game literature: from theory to technological application. Int J Comp Inf Technol 2012 Sep;1(1):20-35.
21. Thin AG, Brown C, Meenan P. User experiences while playing dance-based exergames and the influence of different body motion sensing technologies. Int J Comp Games Technol 2013;2013:1-7. [doi: 10.1155/2013/603604]
22. Haghighi Osgouei R, Soulsby D, Bello F. Rehabilitation exergames: use of motion sensing and machine learning to quantify exercise performance in healthy volunteers. JMIR Rehabil Assist Technol 2020 Aug 18;7(2):e17289 [FREE Full text] [doi: 10.2196/17289] [Medline: 32808932]
23. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
24. Mathieson K. Predicting user intentions: comparing the TAM with the theory of planned behavior. Inf Syst Res 1991 Sep;2(3):173-191. [doi: 10.1287/isre.2.3.173]
25. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
26. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014 Mar 24;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
27. World population ageing 2019: highlights. United Nations. 2019. URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf [accessed 2024-03-12]
28. Mostaghel R, Oghazi P. Elderly and technology tools: a fuzzyset qualitative comparative analysis. Qual Quant 2017 Jul 23;51(5):1969-1982 [FREE Full text] [doi: 10.1007/s11135-016-0390-6] [Medline: 28867832]
29. Lu S, Lu F, Lee Y. The use and acceptance of smart home by elderly Taipei users : an examination of senior technology acceptance model. J Bus Adm 2020;45(3):95-116. [doi: 10.3966/102596272020090453004]
30. Nelson-Kakulla B. Who Is the 50+ Gamer? Gaming Attitudes and Habits of Adults Ages 50-Plus. AARP. AARP; 2019 Dec. URL: https://www.aarp.org/research/topics/technology/info-2019/2020-gaming-trends-older-americans.html [accessed 2022-01-20]
31. Bossom A, Dunning B. Video games: An Introduction to the Industry. Vol. 16. London, UK: Bloomsbury Publishing; 2016:395-411.
32. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
33. Braun V, Clarke V, Hayfield N, Terry G. Thematic analysis. In: Liamputtong P, editor. Handbook of Research Methods in Health Social Sciences. Singapore: Springer; 2018:1-18.
34. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. London, UK: Sage; 2013.
35. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, DC: American Psychological Association; 2012:57-71.
36. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2020 Aug 12;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
37. American College of Sports Medicine. In: Riebe D, Ehrman J, Liguori G, Magal M, editors. ACSM's Guidelines for Exercise Testing and Prescription. 10th ed. Philadelphia: Wolters Kluwer; 2018.
38. Aarhus R, Grönvall E, Larsen S, Wollsen S. Turning training into play: embodied gaming, seniors, physical training and motivation. Gerontechnology 2011;10(2):110-120. [doi: 10.4017/gt.2011.10.2.005.00]
39. Brox E, Konstantinidis ST, Evertsen G. User-centered design of serious games for older adults following 3 years of experience with exergames for seniors: a study design. JMIR Serious Games 2017 Jan 11;5(1):e2 [FREE Full text] [doi: 10.2196/games.6254] [Medline: 28077348]
40. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic inquiry. Int J Intercult Relat 1985 Jan;9(4):438-439. [doi: 10.1016/0147-1767(85)90062-8]
41. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 4: trustworthiness and publishing. Eur J Gen Pract 2018 Dec;24(1):120-124 [FREE Full text] [doi: 10.1080/13814788.2017.1375092] [Medline: 29202616]
42. McDonald N, Schoenebeck S, Forte A. Reliability and inter-rater reliability in qualitative research: norms and guidelines for CSCW and HCI practice. Proc ACM Hum-Comput Interact 2019 Nov 07;3(CSCW):1-23. [doi: 10.1145/3359174]

Abbreviations

WHO: World Health Organization

Edited by G Eysenbach, T Leung; submitted 23.01.22; peer-reviewed by S Muraki, S Olsen, V Kotrajaras; comments to author 04.07.22; revised version received 25.08.22; accepted 04.09.23; published 05.04.24.

Please cite as:
Wang YH
Understanding Senior Adults' Needs, Preferences, and Experiences of Commercial Exergames for Health: Usability Study
JMIR Serious Games 2024;12:e36154
URL: https://games.jmir.org/2024/1/e36154
doi:10.2196/36154
PMID:38578674

Original Paper

Examining and Comparing the Energy Expenditure of Two Modes of a Virtual Reality Fitness Game (Supernatural): Indirect Calorimetry Study

Tabitha V Craig[1], BSc; Ryan E Rhodes[2], PhD; Wuyou Sui[2], PhD

[1]Department of Exercise Science, Physical & Health Education, University of Victoria, Victoria, BC, Canada

[2]Behavioural Medicine Lab, Department of Exercise Science, Physical & Health Education, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Wuyou Sui, PhD
Behavioural Medicine Lab
Department of Exercise Science, Physical & Health Education
University of Victoria
3800 Finnerty Rd
Victoria, BC, V8P 5C2
Canada
Phone: 1 250 472 5288
Email: yoahsui@gmail.com

Abstract

Background: The effectiveness of virtual reality (VR) fitness games as a form of moderate to vigorous physical activity has yet to be thoroughly quantified through gold standard energy expenditure measures.

Objective: The purpose of this study was to examine the energy expenditure of 2 medium-intensity modes ("Flow and "Boxing") of a VR fitness game, Supernatural, using indirect calorimetry.

Methods: Indirect calorimetry was used to examine relative and objective maximal oxygen consumption (VO_2 max), metabolic equivalents of task (METs), and calories burned during medium-intensity bouts of both Flow and Boxing gameplay modes in young (mean age 25.42, SD 3.25 years), active individuals (n=12 female and n=11 male). METs and calories were also compared using a triaxial waist-worn accelerometer, an Apple smartwatch, and a VR headset. Mood states were assessed pre- and postbout using the shortened Profile of Mood States Questionnaire. Paired 2-tailed *t* tests were used to examine differences in game modes, between sexes, and pre-post exercise sessions.

Results: Objective and relative VO_2 max averaged 1.93 (SD 0.44) L/min and 27.61 (SD 5.60) mL/kg/min, respectively, between modes. Flow (mean 8.2, SD 1.54 METs) and Boxing (mean 7.6, SD 1.66 METs) are both classified as high energy expenditure, vigorous activities. Calorie expenditure data of the accelerometer and VR headset differed significantly from the metabolic cart. Mood changes pre- to post exercise were consistent with expected values for moderate- to vigorous-intensity physical activity, with participants reporting that they felt more "active," "full of pep," "vigorous," and "lively" ($P<.05$) following bouts. Male individuals reported higher objective oxygen consumption (VO_2) for both Flow and Boxing modes; no other sex-specific differences were observed.

Conclusions: Both Flow and Boxing gameplay modes of Supernatural classify as vigorous physical activity and demonstrate the potential to promote mental and physical health benefits. Supernatural may be an effective exercise modality in a VO_2 training program.

(JMIR Serious Games 2024;12:e53999) doi:10.2196/53999

KEYWORDS

energy expenditure; exergaming; indirect calorimetry; virtual reality; VR; VR fitness; VR gaming

Introduction

Regular physical activity (PA), described as a movement-driven increase in energy expenditure [1], is an important health behavior that has well-established mental and physical benefits [2]. Evidence suggests that individuals who partake in regular PA of 150 minutes of moderate-intensity PA or 75-150 minutes of moderate- to vigorous-intensity PA (MVPA) each week [3] have a lower risk of cancer, type 2 diabetes mellitus, ischemic heart disease, and ischemic stroke [4].

While there are diverse modes in which individuals can engage in regular PA, at-home exercise options have increased in popularity recently [5]. Before the COVID-19 pandemic, at-home PA may have been chosen due to the ease of access, high levels of autonomy in modifying both workouts and equipment, and the increased feeling of comfort provided by exercising in a familiar environment [6]. Due to COVID-19–related shutdowns of traditionally frequented centers for PA (eg, gyms, recreation centers, and organized sports leagues), individuals had to develop new routines to stay active. Many turned to home fitness workouts and exergames, including virtual reality (VR) fitness options, to maintain their PA levels [7,8]. Even beyond the pandemic, home workout options, such as VR gaming and internet-based fitness workouts, continue to appeal to individuals, in part due to their convenience and accessibility.

VR can be described as a head-mounted amalgamation of the human perception of sensation into an interactive virtual environment [9]. Contemporary iterations of VR, specifically immersive VR (such as the Oculus Quest 2 or HTC Vive), involve a head-mounted device inlaid with cameras that track the wearer's environment, in order to project a virtual environment within the headset. VR gaming (also known as VR exercising, VR game exercising, or active VR games) refers to the use of a VR headset (and often handheld controllers) to engage in a single immersive, digital fitness experience consisting of both video game and exercise elements [10]. VR gaming differs from traditional exergaming (eg, Wii Fit and Just Dance) in that the gaming environment is visually immersive and interactive, rather than having visuals confined to a screen and movements or interactivity restricted by the fidelity of an external camera. These limitations of exergaming may contribute to their variability in energy expenditure [11]. Popular VR games include Beat Saber (Beat Games), Fruit Ninja VR (Halfbrick Studios), HOLOFIT (Holodia Holofit), Dance Central (Harmonix Music Systems), and BOXVR (FitXR). VR fitness demonstrated both mental and physical health during the COVID-19 lockdown [12]. Thus, VR fitness continues to remain popular as we emerge from the era of COVID-19 lockdowns, in part due to its enjoyable, flexible, and motivating nature [13].

While VR is gaining in popularity as a form of PA, there are, unfortunately, several factors that currently limit our understanding of whether it can serve as an effective means of achieving MVPA intensity. There is a tremendous amount of variance in the equipment involved within a VR gaming setup. Many studies use specialized equipment (eg, cycle ergometer [14], bespoke machinery [15], or sport-specific VR setups [16]), which would be unrealistic or inaccessible for most users to acquire or comfortably use. Among studies that examined a "regular" VR gaming setup, other confounding factors may also be at play. For example, some studies examine the use of additional accessories to existing VR setups, such as hand weights [17], which further obfuscate the energy expenditure of VR games.

Even in previous research that has used more consumer-friendly VR (eg, Oculus Quest and HTC Vive), the VR games themselves are quite variable in what they demand from the user. For example, Beat Saber, a popular VR rhythm game, requires users to hit targets in a virtual space with either one or both of their handheld controllers. However, the intensity with which these targets must be hit to register as a valid hit can be quite low. Additionally, the movements facilitated by VR games can vary, leading to a change in energy expenditure. Stewart [18] compared the energy expenditure from 3 different VR games, including Beat Saber, and found that each game demonstrated a significantly different energy expenditure for a similar playtime. Given the nature of VR setups and the range of movement they require [19], few studies have used the gold-standard measures of energy expenditure (eg, indirect calorimetry) within VR gaming studies. Furthermore, the presence of sex and gender differences within the field of VR gaming remains a controversial topic, with mixed evidence supporting a higher rate of physical discomfort and symptoms for female individuals [20,21]. Combined with the tendency for male individuals to demonstrate higher maximal oxygen consumption (VO_2 max) values when controlling for age [22], investigating the presence of any sex-specific differences within VR gaming is of importance.

Supernatural is a VR fitness service that is available on the Oculus Quest (Meta) headsets, a popular consumer VR headset. Supernatural was chosen due to its specificity as a fitness service [23]. Other contemporary titles, like Beat Saber and Fruit Ninja VR, play closer to rhythm or arcade games in that the goal of the game is to achieve a high score, which can be done by hitting more targets. Supernatural is similar in that target accuracy is a metric; however, the key difference is that Supernatural records a metric of movement power (ie, power for the "Flow" mode and speed for the "Boxing" mode). In this way, Supernatural provides a means to compare effort, not just accuracy. Hence, limiting our study to just Supernatural was deemed appropriate. Thus, the primary purpose of this study was to examine and compare the energy expenditure of a bout of the VR fitness game Supernatural through indirect calorimetry; specifically, we examined the energy expenditure of both the Flow and Boxing modes, selecting workout durations and intensities (ie, medium intensity) based on the game's average workout data. Comparisons between sexes (ie, male and female) were also examined as part of this primary objective.

The emerging popularity and novelty of VR gaming presents a further point of interest insofar as how accurate traditional PA monitoring devices are when applied to VR contexts. Traditional devices, such as waist-worn accelerometers and wrist-worn accelerometers (eg, smartwatches), demonstrate reasonable validity for capturing walking and running behaviors [24]; however, whether this extends to the more space-restricted, calisthenic-focused movements of VR gaming is unclear [25,26]. The inclusion of an onboard accelerometer in many consumer VR headsets is also a noteworthy activity monitor that is worth comparing to gold-standard energy expenditure measures [23]. Thus, the secondary purpose of this study was to examine the energy expenditure of the VR game as measured by accelerometry, consumer activity monitor (ie, Apple Watch 2), and the built-in accelerometer in the VR headset to provide a comparison to the gold-standard measure of indirect calorimetry.

The tertiary purpose of this study was to examine any changes to mood as a result of engagement in Flow and Boxing, as compared to the contemporary literature on VR exercise and mood [27,28].

Methods

Participants

Inclusion criteria for participants were (1) being 19-40 years of age; (2) self-reporting a minimum of 150 minutes of MPVA per week; (3) self-identifying as not being at increased risk for contracting COVID-19 or being a part of an immunocompromised population; and (4) being considered to have a minimal risk of an exercise-induced adverse outcome. To ensure participants with a broad range of VR experiences were recruited, recruitment documentation described the study as a "digital fitness experience" to avoid confounding any measurements collected. Participants became aware of the use of VR technology upon their receipt of the informed consent documentation before their first visit. Participants were recruited between September 2022 and December 2022 using a combination of social media postings, physical postings on the host university's campus, and word of mouth.

VR Headset and Game

Participants engaged with the VR fitness game Supernatural [29], which was played using the Oculus Quest 2 VR headset [30]. The Oculus Quest 2 uses a head-mounted display and 2 handheld controllers to provide an immersive VR experience for users. For this study, a 5-foot (1.5 m) by 5-foot space was marked onto the floor with tape to calibrate the in-headset play area boundaries (Multimedia Appendix 1). For each participant, the floor level was calibrated, and headset straps were adjusted to fit comfortably.

The Supernatural game has 2 exercise modes: Flow and Boxing. Both modalities cue participants to arm movements with color-coordinated orbs that have directional arrows to indicate movements to participants. There are also horizontal bars and triangles to encourage squatting, lunging, and the dodging of obstacles. During their first session, participants engaged in the workout mode Flow. Flow is an aerobic workout that involves both the upper and lower body and footwork in 360 degrees (Figure 1). Participants wield a virtual bat in each hand, striking targets in a variety of patterns and intensities. Lower-body movements are incorporated into each sequence, requiring participants to squat or lunge to hit some targets. During the second session, participants challenged a Boxing workout. Boxing requires participants to punch, uppercut, or swing through color-coordinated orbs and has horizontal and diagonal bars that participants have to maneuver their head and torso under and around (Figure 2). Note that still captures of Supernatural provide only an approximate impression of the actual, first-person in-app perspective of a Supernatural user on modern Meta Quest hardware.

Figure 1. Supernatural Flow game mode.

Figure 2. Supernatural Boxing game mode.

Experimental Overview

Overview

Interested participants contacted the researcher by email to arrange the initial session. Participants attended 2 sessions at the host institution which occurred a minimum of 24 hours apart. The first visit centered around the Supernatural Flow session, and the second visit was around the Supernatural Boxing session. This order was maintained for all participants. Participants were asked to abstain from engaging in MVPA for 12 hours before each session. During the first study visit, participants signed the informed consent document, completed a Get Active Questionnaire (GAQ [31]), had their anthropometric data collected, and completed a baseline survey assessing previous knowledge and exposure to VR. Before and after each session, participants' mood states using the shortened version of the Profile of Mood States (POMS-SF [32]) were assessed. After initial surveys, participants were fitted with a heart rate (HR) monitor and the Oculus Quest 2 VR headset and then followed a standardized progression up to the measurement intensity workout until meeting the predetermined threshold for competency (see Zones of Competency section). Upon meeting the zone of competency, participants were fitted with the indirect calorimetry mask and headpiece, Apple Watch, and waist-worn accelerometer and completed the measurement session.

VR Game Stage Progression

Following initial surveys at each visit, participants watched a series of tutorial videos for the respective mode of the session (ie, Flow for session 1 and Boxing for session 2) which are available on Supernatural's YouTube page. Participants then engaged in an in-game tutorial (approximately 5 minutes). Participants then proceeded to a low-intensity "Quick Hits" workout and a medium-intensity "Quick Hits" ramp-up workout, both approximately 5 minutes long, to further familiarize them with the mode. Participants who were able to meet the minimum zones of competency for each mode (as detailed in Zones of Competency section) proceeded to the measurement workout. Participants could repeat the "Quick Hits" workouts until the zones of competency were obtained. The measurement session was a medium-intensity workout between 14 and 17 minutes in length. All workouts were selected from a predetermined list of relatively similar-intensity workouts. In correspondence with Supernatural, the makeup of the different workouts within the same intensity is generally similar to one another in the range of motion, target origination, pace, and number of targets delivered. Participants chose the genre of music (eg, pop, rock, hip-hop, and electronic) they preferred, and a researcher would select a workout matching the description.

Zones of Competency

To advance to the measurement session, participants had to meet the predetermined zones of competency for each mode. Flow measures participant competency in accuracy (ie, percentage of targets hit) and power (ie, how fast targets were hit). Participants had to obtain a minimum of 92% accuracy to participate in the measurement session. Power was not used in the determination of participant readiness. Boxing measures participant competency in accuracy (ie, percentage of targets hit) and speed (ie, how fast targets were hit). Participants had to achieve a minimum of 94% accuracy in Boxing to progress to the measurement round. During preliminary testing, it was also determined that a minimum speed goal should be achieved to ensure participants were all working out at a similar intensity. Hence, a minimum score of 70% speed was determined to be sufficient, as this best reflected the lower end of the speed an average user would hit the targets at, according to correspondence with Supernatural.

Instrumentation and Measurement

Primary Outcome: Energy Expenditure (Breath-by-Breath Oxygen Consumption)

A Parvo Medics TrueOne 2400 metabolic cart was used to assess oxygen consumption (VO_2; ie, volume of oxygen) during both measurement sessions [33]. A 2-meter hose and a 3-meter hose were joined together using a polyvinyl chloride elbow joint and hose clamps to allow the hose to reach from the metabolic cart to loop suspended above the participant, then down to the participant to allow VO_2 to be assessed with minimal interference during the testing (Multimedia Appendices 1 and 2). The metabolic cart was calibrated to room air and known gas concentrations before testing, followed by individual setup according to participants' height, weight, age, and sex. Both objective VO_2 max and relative VO_2 max were collected. Relative VO_2 max was also compared to age-predicted VO_2 max. VO_2 averages were recorded in 30-second intervals.

Secondary Outcomes

Energy Expenditure (Metabolic Equivalents and Caloric Expenditure)

Metabolic equivalents of task (METs) and caloric expenditure were assessed using the metabolic cart, an ActiGraph GTX3 triaxial waist-worn accelerometer, a Series 7 Apple Watch, and the built-in accelerometer on board the Oculus Quest 2 headset (as displayed through the Oculus Move in-headset app). Time spent in low-, moderate-, and high-intensity exercise was recorded by the waist-worn accelerometer. The waist-worn accelerometer was calibrated according to each participant's age, height, weight, and sex. The Apple Watch and Oculus Move accounts were calibrated to a reference individual (female, height 181 cm, weight 78 kg) to avoid the need for an individual Apple and Oculus Move account for each participant. The Oculus Quest 2 boundaries were standardized to a 1.5 meter by 1.5 meter square.

Heart Rate

A Polar H10 HR monitor was used to measure average and maximum beats per minute (bpm) and the percentage of time spent in HR zones. Polar classifies HR into 5 zones by both levels of intensity and percentage of age-predicted maximum HR (HRmax). Specifically, these zones are categorized as zone 1: very light, 50%-60% HRmax; zone 2: light, 60%-70% HRmax; zone 3: moderate, 70%-80% HRmax; zone 4: hard, 80%-90% HRmax; and zone 5: maximum, 90%-100% HRmax [34]. The Polar H10 HR monitor was calibrated to a reference individual (female, height 181 cm, weight 78 kg) to avoid the need for an individual Polar account for each participant.

Tertiary Outcomes

Demographics

Participants' age, sex, height, weight, waist circumference, and PA levels were recorded during the initial visit. Age, sex, and PA level were self-reported using a single item. Height, weight, and waist circumference were measured by a researcher.

Mood States

Participants completed surveys at the beginning and end of each session to assess potential changes in mood states, which were assessed with the POMS-SF [32]. Our specific subscales of interest were vigor (ie, lively, active, energetic, full of pep, and vigorous) and fatigue (ie, worn out, fatigued, exhausted, weary, and bushed), as these were thought to be the most receptive to participants' exertion within our acute aerobic intervention and nonclinical sample [35,36]; however, the entire POMS-SF was completed by each participant.

Data Handling and Analysis

Energy Expenditure Outcomes

Energy expenditure outcomes were analyzed descriptively (mean and SD) and assessed for normality. Winsorization [37] of any outliers was planned. An average of the objective VO_2 max and relative VO_2 max for Flow and Boxing sessions was calculated. The relative VO_2 max was also calculated as a percentage of the age-predicted VO_2 max. Accelerometer data were analyzed using the Freedson Adult MV3 cut points [38]. The accelerometer, Apple Watch, and Oculus Move data were compared to the metabolic cart data to determine the percentage difference between the measurement tools. Comparisons between modes (Flow and Boxing) and measurement modalities (ie, metabolic cart, waist-worn accelerometer, Apple Watch, and Oculus Move) were compared for the 21 complete data sets (n=11, 52% male individuals) using paired sample 2-tailed *t* tests using a Bonferroni correction (ie, α=.0125). Missing data were excluded case-wise.

Psychological Outcomes

Changes in mood state (ie, vigor and fatigue subscales) from pre- to postsession for both Flow and Boxing modalities were assessed using paired 2-tailed *t* tests.

Sample Size Determination

To be sufficiently powered to do a sex-specific subanalysis, we aimed to collect at least 20 full data sets (ie, 10 male and 10 female individuals). This number was determined based on an investigation of previous energy expenditure and VR fitness publication recruitment numbers and was deemed appropriate to capture sex-specific differences in energy expenditure outcomes [39-41]. Sex-specific subanalyses were only performed for the energy expenditure outcomes, as the sample was deemed to be too small to capture small-to-medium–sized effects on psychological outcomes.

Ethical Considerations

This prospective, single-group experimental study was approved by the host University of Victoria's research ethics board (22-0213), and all participants provided written, informed consent before their involvement in the study and for inclusion in the publication of any research findings, as indicated by the Declaration of Helsinki.

Results

A CONSORT (Consolidated Standards of Reporting Trials)-eHealth checklist for this study can be found in Multimedia Appendix 3.

Participant Characteristics

A total of 12 male and 12 female individuals who met inclusion criteria and provided informed consent were recruited for this study. One male participant dropped out before beginning the Flow data collection due to issues with the VR environment. Overall, 2 female participants completed the Flow session but not the Boxing session due to factors unrelated to the study. A total of 23 (n=11, 48% male) participants completed the Flow session, and a total of 21 participants (n=11, 52% male) completed the Boxing session.

Participant demographics are presented in Table 1. The average age of participants was 25.42 (SD 3.25) years. Participants reported participating in MVPA 4.71 (SD 1.62) days a week for an average of 72.38 (SD 39.04) minutes per session, classifying our participants as active [3]. Our participant pool was very naive to VR and VR fitness, however, with 96% (n=23) reporting no previous familiarity with VR fitness products before study participation. There were no outliers in the data set.

Table 1. Participant demographics.

Characteristic	Male participants (n=12), mean (SD)	Female participants (n=12), mean (SD)
Age (years)	25.5 (3.1)	25.2 (3.0)
Height (cm)	178.1 (9.9)	166.6 (7.9)
Weight (kg)	76.7 (14.1)	64.36 (7.9)
Waist circumference (cm)	85.5 (6.2)	76.0 (5.1)

Workout Characteristics

The average Flow workout was 15.48 (SD 1.31) minutes in duration and was completed with 94.57% (SD 2.35%) accuracy and 85.52% (SD 5.27%) power. The average Boxing workout was slightly longer than Flow at 16.91 (SD 1.51) minutes in length, with 96.67% (SD 1.77%) accuracy and 78.86% (SD 7.89%) speed.

Energy Expenditure

Oxygen Consumption

Objective VO_2 max, as measured by the metabolic cart, was 1.98 (SD 0.44) L/min for Flow, 1.88 (SD 0.45) L/min for Boxing, and 1.93 (SD 0.44) L/min overall. Relative VO_2 max was 28.52 (SD 5.39) mL/kg/min for Flow, 26.70 (SD 5.79) mL/kg/min for Boxing, and 27.61 (SD 5.60) mL/kg/min overall. There was a significant difference between Flow and Boxing for both objective (mean difference [M_{diff}]=0.14, 95% CI 0.05-0.24; *P*=.006) and relative VO_2 max (M_{diff}=2.05, 95% CI 0.53-3.56; *P*=.01). The percentage of age-predicted VO_2 max was 59.39% (SD 11.75%) for Flow, 55.42% (SD 12.45%) for Boxing, and 57.41% (SD 12.12%) overall. There was no difference between Flow and Boxing with respect to the percentage of age-predicted VO_2 max (M_{diff}=4.20%, 95% CI 1.34%-7.05%; *P*=.006).

With respect to sex-specific differences, a significant difference was observed for objective VO_2, with male individuals demonstrating a significantly higher objective VO_2 for Flow than female individuals (M_{diff}=0.46, 95% CI 0.13-0.78; *P*=.009). Sex differences for objective VO_2 for Boxing were not statistically significant (M_{diff}=0.36, 95% CI –0.02 to 0.74; *P*=.06). No significant differences between sexes were revealed for any other VO_2 outcome (*P*>.05).

Outcomes for METs

METs were collected by both the metabolic cart and the waist-worn accelerometer. The metabolic cart recorded average METs to be 8.15 (SD 1.54) for Flow, 7.63 (SD 1.66) for Boxing, and 7.89 (SD 1.60) overall, while the waist-worn accelerometer recorded averages as 4.31 (SD 0.56) for Flow, 4.78 (SD 0.57) for Boxing, and 4.55 (SD 0.65) overall. For data recorded by the metabolic cart, there was a significant difference between Flow and Boxing modes (M_{diff}=0.58, 95% CI 0.15-1.02; *P*=.01). There was a significant difference between the values recorded by the metabolic cart and the waist-worn accelerometer for both Flow (M_{diff}=3.69, 95% CI 3.12-4.25; *P*<.001) and Boxing (M_{diff}=2.84, 95% CI 2.21-3.47; *P*<.001), with the accelerometer reporting a percent of metabolic cart reading of 45.26% (SD 14.55%) for Flow, 56.03% (SD 13.80%) for Boxing, and 50.64% (SD 15.05%) overall.

With respect to sex-specific differences, no significant differences were revealed for METs as assessed by either the metabolic cart or waist-worn accelerometer (*P*>.05). Oxygen consumption and metabolic equivalent data are given in Table 2.

Table 2. Oxygen consumption (VO2) and metabolic equivalent of task (MET) data.

Energy expenditure outcome	Flow, mean (SD)	Boxing, mean (SD)	Overall, mean (SD)
Objective VO_2 (L/min)	1.98 (0.44)[a]	1.88 (0.45)[a]	1.93 (0.44)
Relative VO_2 (mL/kg/min)	28.52 (5.39)	26.70 (5.79)	27.61 (5.60)
Age-predicted VO_2 max[b] (%)	59.39 (11.75)	55.42 (12.45)	57.41 (12.12)
METs (metabolic cart)	8.15 (1.54)[a]	7.63 (1.66)[a]	7.89 (1.60)
METs (accelerometer)	4.31 (0.56)	4.78 (0.57)	4.55 (0.65)

[a]Values represent a significant difference between modes (ie, Flow and Boxing).

[b]VO_2 max: maximal oxygen consumption.

Calories

The estimated caloric expenditure of the metabolic cart, waist-worn accelerometer, Apple Watch, and Oculus are presented in Table 3. For Flow, both the accelerometer (M_{diff}=79.25, 95% CI 68.30-90.20; *P*<.001) and Oculus Move (M_{diff}=30.79, 95% CI 12.84-48.74; *P*=.002) were found to be significantly different from the metabolic cart, while the Apple Watch (M_{diff}=–4.83, 95% CI –18.30 to 8.66; *P*=.47) was not.

Similarly, for Boxing, both the accelerometer (M_{diff}=70.56, 95% CI 56.59-84.53; *P*<.001) and Oculus Move (M_{diff}=62.89, 95% CI 45.52-80.26; *P*<.001) were found to be significantly different from the metabolic cart, while the Apple Watch (M_{diff}=–14.24, 95% CI –32.22 to 3.73; *P*=.11) was not.

Table 3. Caloric expenditure data.

Device	Flow (kcal)		Boxing (kcal)		Overall (kcal)	
	Mean (SD)	% Metabolic cart	Mean (SD)	% Metabolic cart	Mean (SD)	% Metabolic cart
Metabolic cart	151.22 (35.52)	N/A[a]	159.76 (38.829)	N/A	155.49 (36.95)	N/A
Accelerometer	70.60 (22.34)[b]	45.26[b]	89.20 (28.892)[b]	56.03[b]	79.90 (27.12)[b]	50.46[b]
Apple Watch	155.41 (24.87)	107.56	174.00 (39.607)	112.34	164.71 (33.59)	109.95
Oculus Move	121.27 (21.98)[b]	82.88[b]	102.00 (17.914)[b]	63.75[b]	111.63 (22.06)[b]	73.32[b]

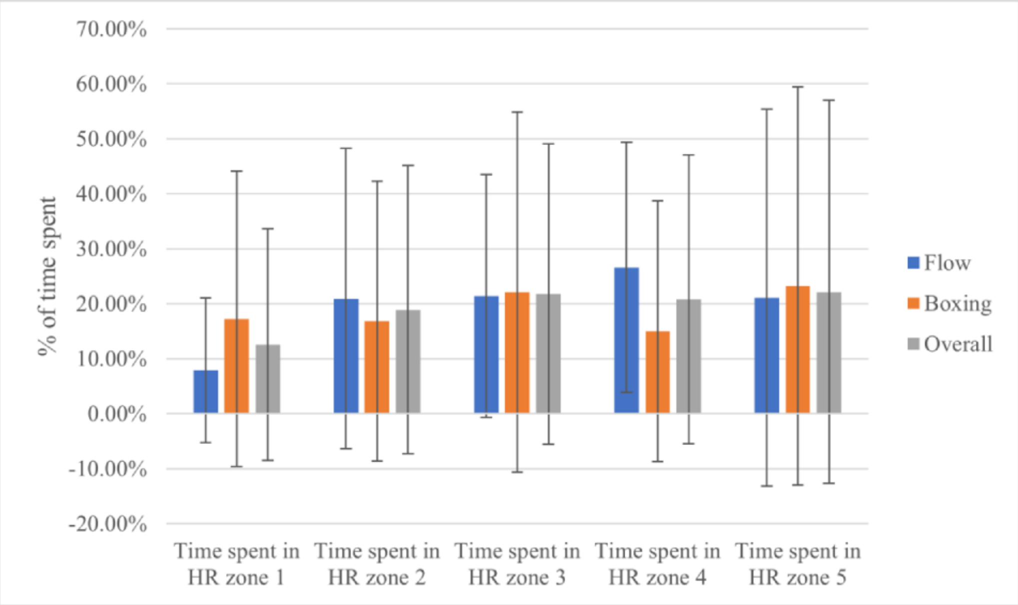
[a]N/A: not applicable.

[b]Values represent a significant difference from the metabolic cart (*P*≤.0125).

Heart Rate

The average HR for Flow was 151.13 (SD 23.12) bpm, and HRmax was 169.39 (SD 20.04) bpm. The average HR for Boxing was 143.29 (SD 26.36) bpm, with an HRmax of 161.43 (SD 26.36) bpm. Flow demonstrated a significantly higher average HR (M_{diff}=8.43, 95% CI 3.45-13.40; *P*=.002) and HRmax (M_{diff}=8.38, 95% CI 3.12-13.64; *P*=.003) when compared to Boxing. The overall average HR across both modalities was 147.21 (SD 23.12) bpm, and the average overall HRmax was 165.41 (SD 20.04). Sex-specific differences in HR were observed for Boxing average HR (M_{diff}=25.61, 95% CI 1.99-49.23; *P*=.04) and HRmax (M_{diff}=25.53, 95% CI 4.04-47.01; *P*=.02), with male individuals demonstrating a higher HR during Boxing. Time spent in HR zones is presented in Figure 3.

Figure 3. Time spent in heart rate zones. Error bars represent the SD. HR: heart rate.



Changes in Mood States

Between baseline and postsession 1, on average, participants reported that they felt more "lively" (M_{diff}=0.46, 95% CI 0.50-0.86; *P*=.03), more "active" (M_{diff}=0.67, 95% CI 0.32-1.07; *P*<.001), more "full of pep" (M_{diff}=0.57, 95% CI 0.14-0.99; *P*=.01), and more "vigorous" (M_{diff}=0.48, 95% CI 0.90-0.87; *P*=.02). Between presession 2 and postsession 2, on average, participants reported feeling more "active" (M_{diff}=0.76, 95% CI 0.36-1.17; *P*<.001), more "full of pep" (M_{diff}=0.38, 95% CI 0.02-0.75; *P*=.04), and more "vigorous" (M_{diff}=0.52, 95% CI 0.18-0.87; *P*=.004). No item on the fatigue subscale of the POMS-SF changed significantly for either session 1 or 2.

Discussion

VR fitness games have seen an increase in popularity as a mode of PA in recent years [7]. However, variability among game demands and objectives, along with difficulties in measuring energy expenditure with gold standard assessments (ie, indirect calorimetry), have limited our understanding of the intensity of these games [19], in turn limiting recommendations involving this mode of PA with respect to MVPA guidelines. Hence, the primary aim of this study was to examine the energy expenditure of one of the most popular VR fitness games, Supernatural. Specifically, we assessed the energy expenditure of a session of medium intensity for both the Flow and Boxing modes of Supernatural using indirect calorimetry. We also examined how other popular measures of energy expenditure compared relative to indirect calorimetry.

For Flow, average relative VO_2 max was 28.52 (SD 5.39) mL/kg/min, which translated to approximately 8.2 (SD 1.54) METs, classifying it as vigorous intensity (ie, >6 METs) [42]. Compared to other forms of PA, Flow was akin to climbing stairs (8.0 METs) or general circuit training (8.0 METs) [43]. Moreover, the percent of age-predicted VO_2 max (mean 59.39%, SD 11.75%) points to the potential use of this mode of Supernatural within a VO_2 training program [44]. Notably, only objective VO_2 was found to be significantly different between male and female participants, which suggests that this mode of PA demonstrates a significant sex difference in objective energy expenditure. This finding is likely due to the body weight nature of the game (ie, movements were relative to participants' own body weight), as this significant difference was not evident when examining the relative VO_2 values or METs (ie, accounting for participants' weight). This is encouraging, suggesting that young, active individuals engaging in a bout of Supernatural Flow should receive a relatively similar aerobic workout, independent of sex. This indicates that during a flow session, participants were averaging a high enough percentage of VO_2 to potentially improve maximum oxygen uptake.

For Boxing, average relative VO_2 max was 26.70 (SD 5.79) mL/kg/min, which translated to approximately 7.6 (SD 1.66) METs, also classifying it as vigorous intensity (ie, >6 METs) [42]. Compared to other modes of PA, Boxing was higher than high-impact aerobics (7.3 METs) and close to sparring while boxing (7.8 METs) [43]. Interestingly, both the Flow and Boxing modes demonstrated a higher energy expenditure than "activity-promoting video or arcade game (eg, Exergaming and Dance Dance Revolution), vigorous effort" (7.2 METs), which further speaks to the heterogeneity in demands among available VR fitness games [45]. For example, the lower average energy expenditure of the Boxing mode may be due to the difference in physical demands between modes, with Flow incorporating more frequent multimuscle group movements (eg, squats and arm swings) than Boxing, which consists primarily of slips and punches. Similar to Flow, there was a trend in objective VO_2 between sexes (P=.06) favoring male individuals but no significant differences in any other VO_2 outcome or METs, for the Boxing session. Furthermore, while the objective VO_2 max and METs were significantly lower for Boxing than Flow, the data suggests that individuals were still exercising at a high enough percentage of VO_2 to improve maximal oxygen uptake [46]. Hence, Boxing also has implications as a candidate for use in aerobic training programs, independent of sex.

Given the lack of research comparing indirect calorimetry to device-based measures when examining VR fitness, we also aimed to compare the findings from the metabolic cart to that of a triaxial waist-worn accelerometer, an Apple Watch, and the Oculus Move. On average, the accelerometers underestimated the energy expenditure of the metabolic cart by approximately 65% (SD 14.55%) for Flow and approximately 45% (SD 13.8%) for Boxing. This discrepancy and variability were somewhat surprising, given the relative reliability and validity of this particular accelerometer in assessing aerobic PA (eg, running and walking) [47]. However, this underestimation is likely due to the placement of the accelerometer (ie, waist-worn), which is unlikely to capture the full range of dynamic upper and lower body movements characteristic of a VR fitness game. Notably, several other studies have used waist-worn accelerometers within VR research [25,26], though these studies have examined time spent in MVPA rather than an energy expenditure outcome (eg, METs). Despite this, an underrepresentation of MVPA as captured by accelerometry still appears to be evident; Sousa et al [26] reported an average of 4.10 (SD 4.93) minutes of MVPA for a 20-minute VR fitness session (approximately 21%), while Giakoni-Ramirez and colleagues [25] reported an average of 3.57 minutes of MVPA for a 9-minute "intermediate" VR fitness session (approximately 40%). Hence, our findings suggest that the use of accelerometers during this form of PA be used cautiously.

Compared to the accelerometer, the Apple Watch was relatively more accurate, estimating approximately 108% (SD 24.88%) and approximately 112% (SD 25.41%) of the metabolic cart for Flow and Boxing, respectively. This may be due to the wrist-worn placement of the Apple Watch, which makes it more sensitive to the movements of the VR fitness games. Notably, there was a considerable degree of variability within the Apple Watch measures, ranging from a 28% underestimation to a 56% overestimation of caloric expenditure. Our findings are consistent with previous validation work using the Apple Watch to measure MVPA and energy expenditure [48]. Work by Bai and colleagues [49] does support the validity of the Apple Watch as a measure of MVPA; importantly, however, their work uses a waist-worn accelerometer as the criterion measure, which limits the interpretability of their findings in a VR fitness context. Hence, while our results support the relative accuracy of the Apple Watch, its usefulness as a precise measure of energy expenditure during VR fitness is limited.

On average, the Oculus Move accelerometer underestimated the energy expenditure of the VR fitness games measured by the metabolic cart by 27% (SD 23.1%) and 44% (SD 12.48%) for Flow and Boxing, respectively. Like the waist-worn accelerometer, the placement of the accelerometer within the Oculus headset may have impacted the accuracy of its measurements. Similar to the other measures of energy expenditure, the use of this outcome as a measure of energy expenditure is limited.

Lastly, preliminary comparisons of pre-post VR fitness measurement session mood states revealed reported changes in mood that are consistent with the beneficial changes we would expect from a bout of moderate-intensity PA [36] and with previous VR exergaming research [28,35]. These changes provide preliminary evidence for the positive mental health benefits of even single sessions of VR fitness, which is encouraging given the relationship between positive affect and adherence to PA [50].

Though our study contained many strengths, such as a gold-standard measure of energy expenditure (ie, indirect calorimetry) and a relatively homogeneous sample, there are limitations to our work. One limitation of this study is that all participants were naive to VR fitness. This may have resulted in a variable amount of energy expended compared to someone who regularly engages in VR fitness. Although we implemented a strict threshold for inclusion in the measurement session, discrepancies in the final accuracy and power scores, along with direct observation of participants, suggest that some participants spent more or less energy adjusting to the difficulty of the measurement session workout. In other words, the final score across participants suggests that some struggled more than others in acclimating to the difficulty of the measurement session. Further, all participants were active individuals (ie, meeting the weekly PA guidelines). As a result, the measured metabolic and cardiorespiratory responses to the measurement session may not be reflective of the average new user.

Both the Flow and Boxing medium-intensity modes of Supernatural demonstrated relatively high energy expenditures (8.2, SD 1.54 METs, and 7.6, SD 1.66 METs, respectively), classifying as vigorous-intensity PA (ie, >6.0 METs) [42]. Device-based measures of energy expenditure varied considerably both between participants and when compared to the metabolic cart results. Hence, caution should be used when using and interpreting device-based measures of energy expenditure within VR fitness games, including Supernatural. This study also provides preliminary evidence to support the physical and mental health benefits of engaging in VR fitness games like Supernatural. These findings are encouraging, given the increasing popularity and accessibility of VR fitness games as a means of achieving MVPA. For individuals who are interested in being physically active at home or are unable to access traditional forms of exercise, VR fitness presents a potential supplement or alternative to achieving the recommended levels of MVPA.

Acknowledgments

This research was funded by Supernatural.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to privacy restrictions in accordance with the study sponsor but are available from the corresponding author on reasonable request.

Conflicts of Interest

This research was commissioned and funded by Supernatural. Supernatural was involved in the design and development of the protocol but was not involved in the conduction, analysis, or publication of the study or study data.

Multimedia Appendix 1

Experimental setup in action detailing the Oculus Quest 2 VR headset, metabolic cart apparatus, Apple Watch 2, and ActiGraph waist-worn accelerometer (located beneath model’s shirt).

[PNG File , 3761 KB - games_v12i1e53999_app1.png]

Multimedia Appendix 2

Experimental setup including Oculus Quest 2 VR headset and metabolic cart headgear and mouthpiece.

[PNG File , 4287 KB - games_v12i1e53999_app2.png]

Multimedia Appendix 3

CONSORT (Consolidated Standards of Reporting Trials)-eHealth Checklist.

[PDF File (Adobe PDF File), 8798 KB - games_v12i1e53999_app3.pdf]

References

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]
2. Rhodes RE, Janssen I, Bredin SSD, Warburton DER, Bauman A. Physical activity: health impact, prevalence, correlates and interventions. Psychol Health 2017;32(8):942-975. [doi: 10.1080/08870446.2017.1325486] [Medline: 28554222]
3. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
4. Kyu HH, Bachman VF, Alexander LT, Mumford JE, Afshin A, Estep K, et al. Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. BMJ 2016;354:i3857 [FREE Full text] [doi: 10.1136/bmj.i3857] [Medline: 27510511]

5. Nyenhuis SM, Greiwe J, Zeiger JS, Nanda A, Cooke A. Exercise and fitness in the age of social distancing during the COVID-19 pandemic. J Allergy Clin Immunol Pract 2020;8(7):2152-2155 [FREE Full text] [doi: 10.1016/j.jaip.2020.04.039] [Medline: 32360185]
6. Kaushal N, Rhodes RE. The home physical environment and its relationship with physical activity and sedentary behavior: a systematic review. Prev Med 2014;67:221-237. [doi: 10.1016/j.ypmed.2014.07.026] [Medline: 25084562]
7. Mutz M, Gerke M. Sport and exercise in times of self-quarantine: how Germans changed their behaviour at the beginning of the COVID-19 pandemic. Int Rev Sociol Sport 2020;56(3):305-316 [FREE Full text] [doi: 10.1177/1012690220934335]
8. Sui W, Rush J, Rhodes RE. Engagement with web-based fitness videos on YouTube and Instagram during the COVID-19 pandemic: longitudinal study. JMIR Form Res 2022;6(3):e25055 [FREE Full text] [doi: 10.2196/25055] [Medline: 35258459]
9. Isenstein EL, Waz T, LoPrete A, Hernandez Y, Knight EJ, Busza A, et al. Rapid assessment of hand reaching using virtual reality and application in cerebellar stroke. PLoS One 2022;17(9):e0275220 [FREE Full text] [doi: 10.1371/journal.pone.0275220] [Medline: 36174027]
10. Mologne MS, Hu J, Carrillo E, Gomez D, Yamamoto T, Lu S, et al. The efficacy of an immersive virtual reality exergame incorporating an adaptive cable resistance system on fitness and cardiometabolic measures: a 12-week randomized controlled trial. Int J Environ Res Public Health 2022;20(1):210 [FREE Full text] [doi: 10.3390/ijerph20010210] [Medline: 36612530]
11. Bailey BW, McInnis K. Energy cost of exergaming: a comparison of the energy cost of 6 forms of exergaming. Arch Pediatr Adolesc Med 2011;165(7):597-602 [FREE Full text] [doi: 10.1001/archpediatrics.2011.15] [Medline: 21383255]
12. Menhas R, Qin L, Saqib ZA, Younas M. The association between COVID-19 preventive strategies, virtual reality exercise, use of fitness apps, physical, and psychological health: testing a structural equation moderation model. Front Public Health 2023;11:1170645 [FREE Full text] [doi: 10.3389/fpubh.2023.1170645] [Medline: 37483921]
13. Seo EY, Kim YS, Lee YJ, Hur MH. Virtual reality exercise program effects on body mass index, depression, exercise fun and exercise immersion in overweight middle-aged women: a randomized controlled trial. Int J Environ Res Public Health 2023;20(2):900 [FREE Full text] [doi: 10.3390/ijerph20020900] [Medline: 36673656]
14. Zeng N, Pope Z, Gao Z. Acute effect of virtual reality exercise bike games on college students' physiological and psychological outcomes. Cyberpsychol Behav Soc Netw 2017;20(7):453-457. [doi: 10.1089/cyber.2017.0042] [Medline: 28715263]
15. Feodoroff B, Konstantinidis I, Froböse I. Effects of full body exergaming in virtual reality on cardiovascular and muscular parameters: cross-sectional experiment. JMIR Serious Games 2019;7(3):e12324 [FREE Full text] [doi: 10.2196/12324] [Medline: 31464194]
16. Akbaş A, Marszałek W, Kamieniarz A, Polechoński J, Słomka KJ, Juras G. Application of virtual reality in competitive athletes - a review. J Hum Kinet 2019 Oct;69:5-16. [doi: 10.2478/hukin-2019-0023] [Medline: 31666884]
17. Polechoński J, Zwierzchowska A, Makioła Ł, Groffik D, Kostorz K. Handheld weights as an effective and comfortable way to increase exercise intensity of physical activity in virtual reality: empirical study. JMIR Serious Games 2022;10(4):e39932 [FREE Full text] [doi: 10.2196/39932] [Medline: 36416873]
18. Stewart TH. Effect of exercise setting on energy expenditure and enjoyment during active virtual reality gaming [thesis]. San Francisco State University. San Francisco; 2019. URL: https://scholarworks.calstate.edu/concern/theses/2801pj090 [accessed 2024-03-12]
19. Stewart TH, Villaneuva K, Hahn A, Ortiz-Delatorre J, Wolf C, Nguyen R, et al. Actual vs. perceived exertion during active virtual reality game exercise. Front Rehabil Sci 2022;3:887740 [FREE Full text] [doi: 10.3389/fresc.2022.887740] [Medline: 36189005]
20. Grassini S, Laumann K. Are modern head-mounted displays sexist? A systematic review on gender differences in HMD-mediated virtual reality. Front Psychol 2020;11:1604 [FREE Full text] [doi: 10.3389/fpsyg.2020.01604] [Medline: 32903791]
21. Bannigan GM, de Sousa AA, Scheller M, Finnegan DJ, Proulx MJ. Potential factors contributing to observed sex differences in virtual-reality-induced sickness. Exp Brain Res 2024;242(2):463-475 [FREE Full text] [doi: 10.1007/s00221-023-06760-0] [Medline: 38170233]
22. Malek MH, Berger DE, Housh TJ, Coburn JW, Beck TW. Validity of VO2max equations for aerobically trained males and females. Med Sci Sports Exerc 2004;36(8):1427-1432 [FREE Full text] [doi: 10.1249/01.mss.0000135795.60449.ce] [Medline: 15292753]
23. FAQ's. Supernatural. 2024. URL: https://www.getsupernatural.com/faq [accessed 2024-01-24]
24. Wu WJ, Yu HB, Tai WH, Zhang R, Hao WY. Validity of actigraph for measuring energy expenditure in healthy adults: a systematic review and meta-analysis. Sensors (Basel) 2023;23(20):8545 [FREE Full text] [doi: 10.3390/s23208545] [Medline: 37896640]
25. Giakoni-Ramírez F, Godoy-Cumillaf A, Fuentes-Merino P, Farías-Valenzuela C, Duclos-Bastías D, Bruneau-Chávez J, et al. Intensity of a physical exercise programme executed through immersive virtual reality. Healthcare (Basel) 2023;11(17):2399 [FREE Full text] [doi: 10.3390/healthcare11172399] [Medline: 37685433]
26. Sousa CV, Hwang J, Cabrera-Perez R, Fernandez A, Misawa A, Newhook K, et al. Active video games in fully immersive virtual reality elicit moderate-to-vigorous physical activity and improve cognitive performance in sedentary college students. J Sport Health Sci 2022;11(2):164-171 [FREE Full text] [doi: 10.1016/j.jshs.2021.05.002] [Medline: 34004390]

27. Hensley Z. The effect of virtual reality immersion level on mood, enjoyment level, and intentions of future engagement in exergames [thesis]. California State University, Stanislaus State. Turlock; 2019. URL: https://scholarworks.calstate.edu/concern/theses/3484zh74g [accessed 2024-03-12]
28. Plante TG, Aldridge A, Bogden R, Hanelin C. Might virtual reality promote the mood benefits of exercise? Comput Human Behav 2003;19(4):495-509. [doi: 10.1016/S0747-5632(02)00074-2]
29. Supernatural. URL: https://www.getsupernatural.com [accessed 2023-10-18]
30. Meta Quest 2. Meta. 2024. URL: https://www.meta.com/ca/quest/products/quest-2/ [accessed 2024-01-24]
31. Get active questionnaire. Canadian Society for Exercise Physiology. URL: https://csep.ca/2021/01/20/pre-screening-for-physical-activity/ [accessed 2024-03-12]
32. Shacham S. A shortened version of the Profile of Mood States. J Pers Assess 1983;47(3):305-306. [doi: 10.1207/s15327752jpa4703_14] [Medline: 6886962]
33. Kaviani S, Schoeller DA, Ravussin E, Melanson EL, Henes ST, Dugas LR, et al. Determining the accuracy and reliability of indirect calorimeters utilizing the methanol combustion technique. Nutr Clin Pract 2018;33(2):206-216 [FREE Full text] [doi: 10.1002/ncp.10070] [Medline: 29658183]
34. Heart rate zones | The basics. Polar. 2023. URL: https://www.polar.com/blog/running-heart-rate-zones-basics/ [accessed 2023-10-19]
35. Ochi G, Kuwamizu R, Fujimoto T, Ikarashi K, Yamashiro K, Sato D. The effects of acute virtual reality exergaming on mood and executive function: exploratory crossover trial. JMIR Serious Games 2022;10(3):e38200 [FREE Full text] [doi: 10.2196/38200] [Medline: 36169992]
36. Berger BG, Motl RW. Exercise and mood: a selective review and synthesis of research employing the profile of mood states. J Appl Sport Psychol 2000;12(1):69-92. [doi: 10.1080/10413200008404214]
37. Dixon WJ, Yuen KK. Trimming and winsorization: a review. Statistische Hefte 1974;15(2-3):157-170. [doi: 10.1007/bf02922904]
38. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport 2011;14(5):411-416. [doi: 10.1016/j.jsams.2011.04.003] [Medline: 21616714]
39. Bird JM, Karageorghis CI, Baker SJ, Brookes DA. Effects of music, video, and 360-degree video on cycle ergometer exercise at the ventilatory threshold. Scand J Med Sci Sports 2019;29(8):1161-1173 [FREE Full text] [doi: 10.1111/sms.13453] [Medline: 31050032]
40. Farrow M, Lutteroth C, Rouse PC, Bilzon JLJ. Virtual-reality exergaming improves performance during high-intensity interval training. Eur J Sport Sci 2019;19(6):719-727 [FREE Full text] [doi: 10.1080/17461391.2018.1542459] [Medline: 30403927]
41. Hall C, Figueroa A, Fernhall B, Kanaley JA. Energy expenditure of walking and running: comparison with prediction equations. Med Sci Sports Exerc 2004;36(12):2128-2134 [FREE Full text] [doi: 10.1249/01.mss.0000147584.87788.0e] [Medline: 15570150]
42. Haapala EA, Gao Y, Vanhala A, Rantalainen T, Finni T. Validity of traditional physical activity intensity calibration methods and the feasibility of self-paced walking and running on individualised calibration of physical activity intensity in children. Sci Rep 2020;10(1):11031 [FREE Full text] [doi: 10.1038/s41598-020-67983-7] [Medline: 32620888]
43. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 compendium of physical activities: a second update of codes and MET values. Med Sci Sports Exerc 2011;43(8):1575-1581 [FREE Full text] [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
44. Scribbans TD, Vecsey S, Hankinson PB, Foster WS, Gurd BJ. The effect of training intensity on VO2max in young healthy adults: a meta-regression and meta-analysis. Int J Exerc Sci 2016;9(2):230-247 [FREE Full text] [Medline: 27182424]
45. Mouatt B, Smith AE, Mellow ML, Parfitt G, Smith RT, Stanton TR. The use of virtual reality to influence motivation, affect, enjoyment, and engagement during exercise: a scoping review. Front Virtual Real 2020;1:564664 [FREE Full text] [doi: 10.3389/frvir.2020.564664]
46. Maud PJ, Shultz BB. Gender comparisons in anaerobic power and anaerobic capacity tests. Br J Sports Med 1986;20(2):51-54 [FREE Full text] [doi: 10.1136/bjsm.20.2.51] [Medline: 3730753]
47. Migueles JH, Cadenas-Sanchez C, Ekelund U, Nyström CD, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017;47(9):1821-1845 [FREE Full text] [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]
48. Kwon S, Kim Y, Bai Y, Burns RD, Brusseau TA, Byun W. Validation of the Apple watch for estimating moderate-to-vigorous physical activity and activity energy expenditure in school-aged children. Sensors (Basel) 2021;21(19):6413 [FREE Full text] [doi: 10.3390/s21196413] [Medline: 34640733]
49. Bai Y, Tompkins C, Gell N, Dione D, Zhang T, Byun W. Comprehensive comparison of Apple Watch and Fitbit monitors in a free-living setting. PLoS One 2021;16(5):e0251975 [FREE Full text] [doi: 10.1371/journal.pone.0251975] [Medline: 34038458]
50. Stevens CJ, Baldwin AS, Bryan AD, Conner M, Rhodes RE, Williams DM. Affective determinants of physical activity: a conceptual framework and narrative review. Front Psychol 2020;11:568331 [FREE Full text] [doi: 10.3389/fpsyg.2020.568331] [Medline: 33335497]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
GAQ: Get Active Questionnaire
HR: heart rate
HRmax: maximum heart rate
Mdiff: mean difference
MET: metabolic equivalent of task
MVPA: moderate to vigorous intensity physical activity
PA: physical activity
POMS-SF: shortened version of Profile of Mood States
VO2 max: maximal oxygen consumption.
VO2: oxygen consumption.
VR: virtual reality

Edited by R Baranyi; submitted 26.10.23; peer-reviewed by J Simmich, C Wimmer; comments to author 18.12.23; revised version received 30.01.24; accepted 09.02.24; published 04.06.24.

Please cite as:
Craig TV, Rhodes RE, Sui W
Examining and Comparing the Energy Expenditure of Two Modes of a Virtual Reality Fitness Game (Supernatural): Indirect Calorimetry Study
JMIR Serious Games 2024;12:e53999
URL: https://games.jmir.org/2024/1/e53999
doi:10.2196/53999
PMID:38833285

Original Paper

Exploring the Psychological Effects and Physical Exertion of Using Different Movement Interactions in Casual Exergames That Promote Active Microbreaks: Quasi-Experimental Study

Aseel Berglund[1], PhD; Helena Orädd[1], MSc

Department of Computer and Information Science, Linköping University, Linköping, Sweden

Corresponding Author:
Aseel Berglund, PhD
Department of Computer and Information Science
Linköping University
Campus Valla, B-house
Linköping, 58329
Sweden
Phone: 46 0732703862
Email: aseel.berglund@liu.se

Abstract

Background: Prolonged sedentary behavior, such as sitting or reclining, has consistently been identified as a stand-alone risk factor for heightened cardiometabolic risk and overall mortality. Conversely, interrupting sedentary periods by incorporating short, active microbreaks has been shown to mitigate the negative effects of sedentary behavior. Casual exergames, which mix elements of casual gaming with physical activity, are one prospective intervention to reduce sedentary behavior because they require physical exertion. Casual exergames have shown promise in fostering emotional and physical advantages when played in specific circumstances. However, little research exists on how different types of movement interactions impact the psychological effects as well as the physical exertion of playing casual exergames.

Objective: The primary aim of this work was to explore the psychological effects and physical exertion of playing casual exergames lasting 2 minutes. More precisely, the investigation focused on comparing upper body and full body movement interactions. In addition, the work examined variations in body positions, considering both standing and seated positions during upper body movement interactions.

Methods: Two casual exergames were developed and investigated through 2 quasi-experimental studies. In study 1, we investigated how players' perceptions of control, exertion, and immersion were affected by using upper body as opposed to full body exergame controllers when playing casual exergames. In study 2, we investigated differences in positive affect, performance, enjoyment, and exertion when playing casual exergames with upper body movement interactions in seated and standing positions.

Results: Study 1 showed that perceived control was significantly higher for upper body movement interactions than for full body movement interactions (*P*=.04), but there were no significant differences regarding perceived exertion (*P*=.15) or immersion (*P*=.66). Study 2 showed that positive affect increased significantly for both standing (*P*=.003) and seated (*P*=.001) gameplay. The participants in the standing gameplay group showed slightly higher actual exertion; however, there were no differences between the groups in terms of positive affect, perceived exertion, enjoyment, or performance.

Conclusions: Casual exergames controlled by upper body movement interactions in seated gameplay can produce similar psychological effects and physical exertion as upper body movement interactions in standing gameplay and full body movement interactions. Therefore, upper body and seated casual exergames should not be overlooked as a suitable microbreak activity.

(JMIR Serious Games 2024;12:e55905) doi:10.2196/55905

KEYWORDS

physical activity; exergames; casual exergames; enjoyment; exertion; motion-based games

Introduction

Background

High levels of sedentary behavior (such as sitting or lying down) [1] and low levels of moderate to vigorous physical exercise are characteristics of an inactive lifestyle [1,2]. The people most vulnerable to the negative effects of an inactive lifestyle are those who have highly sedentary lives with minimal exercise [3,4]. The advances in domestic and workplace technologies, along with the changes in personal and public transportation, have decreased the necessity for physical activity, leading to increased sedentary lifestyles among populations [5]. The frequency of light-intensity activities (such as walking and performing household chores) in outdoor and nonoffice occupations has decreased in the last few decades [5]. Sedentary behavior is a widespread issue affecting various groups of people (eg, children [6-8], youth [9], and adults [10,11]) across different demographics and settings. While watching television, reading, using a computer, and playing video games are discretionary activities [12], sitting during school [6-8], university [9], or work [10,11] hours is generally nondiscretionary [12]. Therefore, researchers have investigated the impact of sedentary behavior [5]. Numerous studies have demonstrated the connection between excessive sedentary behavior and unfavorable health indicators or results [13-15], regardless of how much time is spent exercising to counteract the sedentary behavior [16]; for instance, when physical activity is taken into account, adults who sit 10 hours a day, as opposed to 1 hour a day, still show a 34% higher risk of increased mortality [17].

Incorporating active microbreaks has demonstrated the potential to mitigate the adverse consequences linked to sedentary behavior [14,18] and improve people's moods [19]; for example, 2- to 3-minute microbreaks with a light activity level every 30 minutes can positively impact mental and physical health [20-22]. Another study showed that over the course of 8 weeks, individuals who were encouraged to stand up and move every 30 minutes for 1 to 2 minutes during the workday significantly decreased their sedentary behavior by approximately 36 minutes per workday [23]. Thus, encouraging frequent, short breaks from sitting could improve health outcomes related to the risk of chronic diseases [23].

Exergames—video games that use movement and physical exertion during gameplay—have been shown to be effective in providing a good psychological experience [24] while also offering advantages for the player's health [25]. Exergames have been created with a variety of health-related goals in mind, including rehabilitation [26] and better mental health [27]. Exergames can increase physical activity [28,29], decrease sedentary behavior [30-33], and promote more active breaks [34-37]. Exergaming has been shown to achieve both moderate [31,38] and vigorous [38] levels of exertion, suggesting that exergames have the potential to serve as an alternative to conventional exercise [39].

Casual exergames are easily learned and quickly accessible exergames with simple rules, designed to motivate players to engage in moderate-intensity exercise during short play sessions [33]. Incorporating active microbreaks using casual exergames has demonstrated the potential to interrupt sedentary behavior [32,37,40]. Playing casual exergames can also induce positive affective states [31,32,41], which can reduce stress levels [42] and enhance overall well-being [43]. Previous studies have demonstrated that casual exergames can be enjoyable and generate appropriate levels of exertion [31,32,40]. However, little research has been conducted on the psychological experience and physical exertion when playing casual exergames [44-73] or on using different types of movement interactions in casual exergames [45].

When developing exergames, it is important to consider both the psychological and physiological aspects of the player's experience [46-48]. According to the dual flow model, the attractiveness of an exergame is determined by balancing the player's skill level with the exergame's challenge level, resulting in a state of psychological flow; and the effectiveness of the exergame is determined by balancing the player's fitness level with the exergame's physical intensity level [35,49]. As the optimal psychological experience and the optimal physiological experience are not necessarily aligned, focusing exclusively on boosting in-game experience may risk reducing health-related advantages [48]. On the basis of the dual flow model, movements with great psychological attractiveness and physiological effectiveness should be used [48]. The psychological and physiological aspects are also connected because the level of enjoyment experienced during playing exergames is correlated with increased exertion levels [44].

One aspect of psychological attractiveness in the exergame experience is immersion, defined in the context of video games as the degree to which the player participates in a game [50] and loses awareness of their surroundings while playing [51]. Immersion is closely related to the concept of flow, which is defined as the psychological state in which one is fully involved in an activity, losing self-consciousness [52]. Immersion impacts the psychological attractiveness of exergames [53] and can be achieved when players have control over the game's actions [50]. While both flow and immersion revolve around player engagement in a game, flow pertains to the general motivation and enjoyment of an activity, while immersion is more directly tied to the user experience and can be seen as a quantifiable aspect of the game experience that relates to both flow and motivation [44]. Movement-based controllers can enhance immersion levels by allowing for natural interaction, leading to increased well-being associated with physical exercise [53]. As control is one of the first stages in creating an immersive experience, it should be considered as a factor that might affect the player's immersive experience and, in turn, the casual exergame's psychological attractiveness [50]. Higher control can also contribute to improved performance, which is positively related to game enjoyment [54-56].

The characteristics of players' movements when exergaming are influenced by the characteristics of the exergame, the difficulty of remaining motivated without sacrificing the quality of the movements, the exergaming experience even when it is of short duration, and the scoring of points [57]. For the physical exertion of playing exergames, the type of movement interaction used in the exergame is an important design decision [47,58]. One decision revolves around using upper body or full body movements. Specifically, for casual exergames designed to target upper body movements, it is worth investigating whether the use of upper body movements can encourage sufficient levels of exertion. To induce positive health-related outcomes, it has previously been stated that casual exergame play should produce a moderate level of exertion [33]. Previous research has shown that upper body movements in exergames have lower potential for increasing exertion compared to full body movements [46,59]. A meta-analysis discovered that lower body and full body exergames produced more energy expenditure than upper body exergames; thus, the authors concluded that upper body exergaming movements are insufficient for achieving adequate energy expenditure [23]. Another study measured movements during gameplay of 3 different exergames using upper body and full body movements and discovered that adequate physical activity required full body movements [58]. However, another meta-analysis showed that exergames with continuous upper body movements have the potential to meet the recommendations for a moderate-intensity activity level [39]; for example, playing Wii boxing exergames with upper body movement interactions has been shown to elicit a moderate-intensity activity level [39,60-63]. On the basis of the mixed results of previous studies, it is worth further investigating whether upper body movements can elicit similar levels of exertion as full body movements for casual exergames as well as how the psychological experience is affected.

Another factor that can affect the psychological and physiological aspects of casual exergames is whether the exergame is designed to be played in a seated or standing position. One study comparing seated and standing gameplay in an exergame for individuals with mobility impairments found that the participants playing seated had higher perceived exertion [64]. However, because the participants chose gameplay mode depending on their abilities instead of being randomly selected for each condition, other factors apart from the seated or standing position could have affected the results. In addition, the psychological effects experienced could have influenced exertion levels because the seated players rated the exergame as more usable than the standing players. By contrast, another study comparing seated with standing positions when playing an adapted game mat exergame showed that participants had lower energy expenditure, perceived exertion, and heart rate when they played seated than when they played standing [65]. This aligns with research outside of the exergame domain that shows that performing an otherwise sedentary activity while standing requires more energy expenditure than performing the same activity while seated [66,67]. However, when we compared upper body movement interactions in seated and standing positions for a casual exergame, we found that there were no significant differences in perceived exertion between seated and standing gameplay [41]. Furthermore, an additional study compared playing 8 to 10 minutes of a boxing exergame in standing and seated positions and found that the seated position resulted in lower energy expenditure than the standing position, while there was no significant difference in perceived exertion [68]. As “Every move counts towards better health” according to the World Health Organization [69], even if exergame activities do not exceed recommended intensity levels, playing exergames is superior to being inactive [39] and might lead to higher exercise adherence [70]. Furthermore, microbreaks with light-intensity exercise every 30 minutes during the day offer both physical and mental health benefits [20]. Therefore, it could be argued that playing upper body casual exergames while seated is better than not playing at all.

In terms of psychological effects, research shows that both standing and seated gameplay generally have the same level of enjoyment [41,64,65,71,72]. However, in 3 of these previous studies [64,65,72], no statistical analysis comparing sitting and standing positions was performed, and in 2 of the studies, different movements were used in the sitting and standing conditions [65,71]. Thus, the generalizability of the results regarding the effect of standing versus sitting positions on enjoyment is limited. On the basis of the limitations of previous studies, we statistically compared 2-minute seated and standing casual exergame play with the same movement interactions in a previous study and found no significant difference between playing seated or standing in terms of positive affect and enjoyment [41].

Objectives

This work is an extension of a conference paper on full body and upper body movement interactions [45]. The aim of this work was to study the psychological effects and physical exertion of playing short-duration casual exergames lasting 2 minutes based on upper body movement interactions designed to promote active microbreaks to interrupt sedentary behavior. Considering the limited research and the mixed results of the psychological effects and physical exertion of different movement interactions in casual exergames, we specifically wanted to explore the psychological effects and physical exertion of upper body compared to full body movement interactions (study 1) and standing compared to seated positions for upper body movement interactions (study 2). The following research questions (RQs) were addressed:

- RQ1: How do casual exergames based on upper body compared to full body movements differ in terms of their psychological effects and physical exertion?
- RQ2: How do upper body casual exergames played in seated positions compared to standing positions differ in terms of their psychological attractiveness and physical exertion levels?

Methods

Overview

To answer the RQs, we conducted 2 studies using distinct in-house–developed casual exergames (Crossing and Beaver) for each study. The casual exergames were developed using the open-source game engine Godot. Each round of play lasts 2 minutes. Both casual exergames can be played on a computer screen equipped with a webcam that captures players’ movements. To enable players to observe their movements in relation to on-screen events, the webcam feed is displayed in the top left corner of the casual exergame interface. The actual time and collected points are displayed in the upper right corner of the interface.

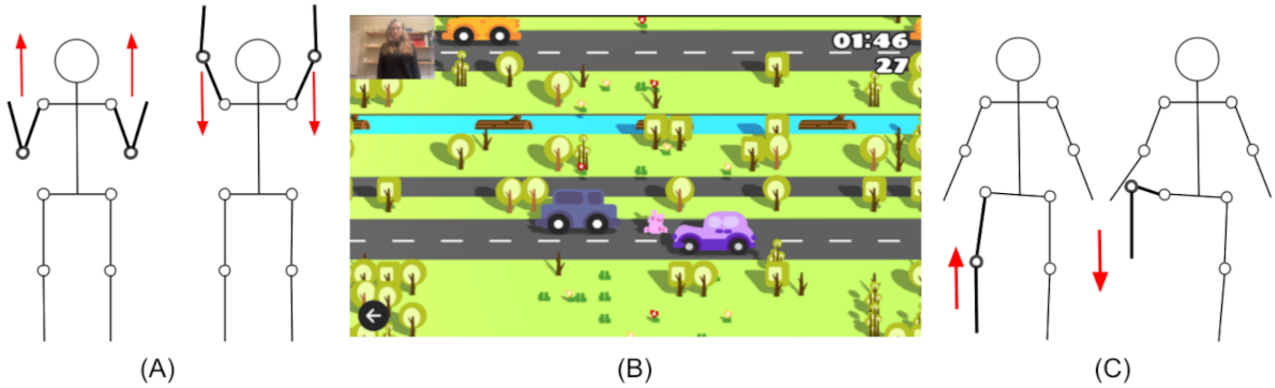
Both studies were performed at a game and cosplay festival. People walking by were asked whether they wanted to participate in a study about exergames. Those interested were provided with detailed information, and if willing to participate, they read through the information materials and signed a consent form. Participants aged <18 years needed their guardian to sign the form on their behalf. By recruiting participants who were naturally interested in games, we aimed to evaluate the impact of exergames on a population that is likely to engage with such interventions in real-world settings. This setting provided a unique opportunity to observe the effects of exergames on individuals already motivated by gaming, which we believe is crucial for understanding the design and immediate effects of these interventions.

Study 1: Upper Body Versus Full Body Movement Interactions

Casual Exergame (Crossing): Design and Interaction

The casual exergame Crossing is based on the classic arcade game Frogger (Figure 1). Using a well-known game genre and basic game mechanics is known to make it easier for players to identify and comprehend the game mechanics [73,74]. In Crossing, players control a rabbit navigating roads, rivers, and rail lines, earning points for forward jumps. Players can also perform sideways jumps, which aid in avoiding obstacles but do not contribute to the accumulation of points. If players collide with vehicles or fall into rivers, the rabbit dies, prompting a 5-second restart. The points acquired before each death are retained, but the waiting period during the restart impacts overall performance.

Figure 1. (A) Demonstration of forward movement for upper body movement interactions for (B) the casual exergame Crossing. (C) Demonstration of forward movement for full body movement interactions. Arrows show the direction of movement.



To support our research, 2 different movement controllers were developed for Crossing: an upper body condition and a full body condition. In the upper body condition, players raise their hands from near their bodies up to their shoulders and above their heads, starting with their arms stretched downward. The rabbit on the screen leaps forward as soon as the hands cross over the head. The player must repeat the movement and extend their arms to the starting position to make the rabbit jump again. In the full body condition, players raise their left or right knee over the hip to make the rabbit jump forward. The rabbit jumps forward when the knee is lifted above the hip. The player must stretch their leg back down to the starting position and repeat the maneuver with either leg to make the rabbit jump once more. To make the rabbit jump to the side in both upper body and full body conditions, players extend a single arm corresponding to the intended direction—raising the right arm for a jump to the right and the left for a jump to the left, unlike extending both arms upward.

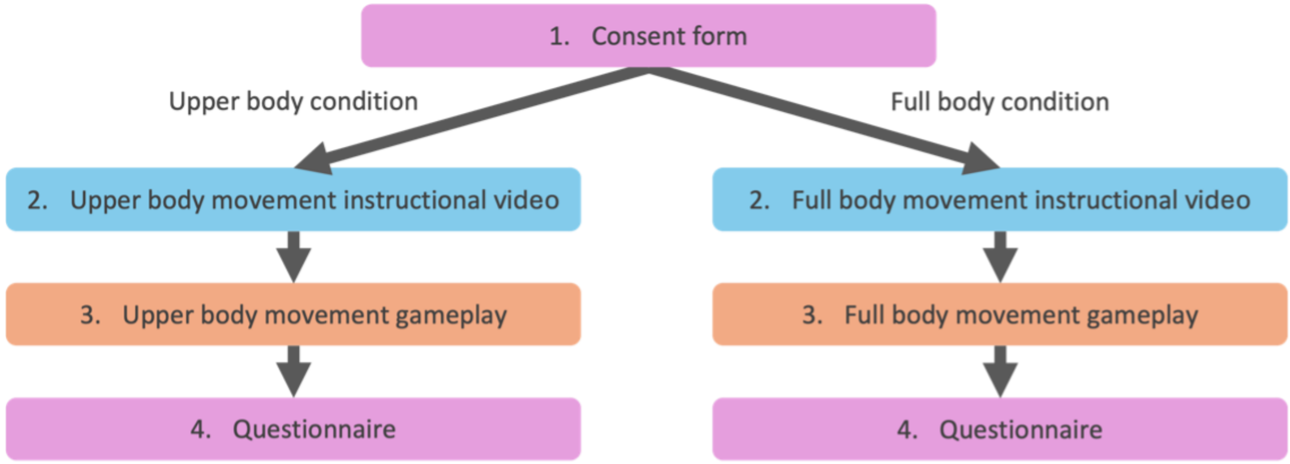
Previous research has indicated the importance of shifting the player's focus away from the physical exertion they are experiencing while playing the exergame [58,73]. When it comes to body focus, Crossing's fast-paced, time-limited gameplay forces players to pay attention to their movements and prevents them becoming self-conscious. The game's obstacles might *kill* players if they remain motionless for an extended period of time, further contributing to the game's fast pace [73].

Participants and Procedure

In all, 80 people participated in the study. The participants' ages ranged from 11 to 54 (mean 25, SD 7) years. Of the 80 participants, 56 (70%) identified as man, 17 (21%) as woman, and 5 (6%) as nonbinary, while 2 (3%) reported that they were unsure about their gender identity. Of the 80 participants, 51 (64%) reported having played exergames multiple times, 25 (31%) reported having played exergames once, and 4 (5%) reported never having played exergames.

After signing the consent form, all participants watched an instructional video about how to play and interact with the exergame. Participants were then asked to stand 1.5 meters away from a computer screen equipped with a webcam. The webcam was adjusted so that the participant was clearly visible, and they were asked to demonstrate the movements they should use. Next, participants played one 2-minute round and then filled out a questionnaire (Figure 2) that included (1) demographic questions, (2) the Borg rating of perceived exertion (RPE) scale [40], and (3) the perceived immersion and control subscales of the Exergame Enjoyment Questionnaire (EEQ) [44]. We alternated between controllers for each new participant, resulting in half the participants (40/80, 50%) playing Crossing in the upper body condition and the other half (40/80, 50%) playing in the full body condition.

Figure 2. Study 1 procedure.



Measures

Physical exertion and psychological effects were the main measures in the study. To measure the physical exertion of playing Crossing, the Borg RPE scale was used to measure perceived exertion [75,76] because it is a valid instrument for measuring exercise intensity [76] and has been used previously to measure perceived exertion when playing casual exergames [33]. The Borg RPE scale ranges from 6 (indicating no exertion at all) to 20 (indicating maximal exertion) [77], with ratings of 11 to 12 corresponding to light intensity, 13 to 14 corresponding to moderate intensity, and 15 to 16 corresponding to hard intensity [78] (Table 1).

Table 1. Correspondence between the Borg rating of perceived exertion (RPE) score and percentage of maximum heart rate [78-80].

Maximum heart rate (%)	Intensity	Borg RPE score
20-39	No exertion at all	6-7
40-59	Very light	8-10
60-69	Light	11-12
70-79	Moderate (somewhat hard)	13-14
80-89	Hard (heavy)	15-16
90-99	Very hard	17-18
100	Maximal	19-20

To measure psychological effects, the perceived immersion and control subscales of the EEQ were used (Table 2) [81]. The EEQ perceived immersion subscale consists of 5 items and measures the degree to which the player is fully engaged and involved in the activity [81]. The EEQ perceived control subscale consists of 4 items and measures the degree to which players can directly affect the outcome of the exergame [81]. All items were measured on a 5-point Likert scale: strongly disagree (score=1), disagree (score=2), neutral (score=3), agree (score=4), and strongly agree (score=5).

Table 2. The Exergame Enjoyment Questionnaire (EEQ) perceived immersion and control items [81].

Item	Scale	Item wording
Immersion 1[a]	EEQ immersion	I did not feel like I wanted to keep playing.
Immersion 2	EEQ immersion	I felt like I lost track of time while playing.
Immersion 3	EEQ immersion	I felt a strong sense of being in the world of the game to the point that I was unaware of my surroundings.
Immersion 4	EEQ immersion	I felt emotionally attached to the game.
Immersion 5	EEQ immersion	I was focused on the game.
Control 1	EEQ control	I felt that it was easy to familiarize myself with the game controls.
Control 2[a]	EEQ control	I felt that it was difficult to understand how the game works.
Control 3	EEQ control	I felt in control over the game.
Control 4	EEQ control	I felt that the game reacted quickly to my movements.

[a]Reversed item.

To provide descriptive data, the number of movements during the play session was measured by counting the number of movements each participant executed while playing the exergame, encompassing both the forward and sideways jumping movements for the rabbit. In addition, each participant's performance was measured by recording their final score.

Statistical Analysis

A between-subjects analysis was conducted for the players in the upper body condition compared to those in the full body condition using SPSS software (version 29.0; IBM Corp). Before the analysis, the items immersion 1 and control 2 were reversed, and internal consistency was measured for the perceived immersion and control subscales of the EEQ using Cronbach α. The Cronbach α value for the 5-item perceived immersion subscale was 0.66, and the Cronbach α value for the 4-item perceived control subscale was 0.68, both close to the satisfactory threshold value of 0.70 [82]. As several of the variables did not meet the assumption of normality and had no significant outliers [83], we chose to use the nonparametric Mann-Whitney *U* test to answer RQ1. The condition was used as the independent variable and perceived control, perceived immersion, and perceived exertion as the dependent variables. Approximate values of the effect size *r* were calculated by dividing the z score for each test by the square root of the number of cases (n=80) [84], following the guidelines of 0.2 corresponding to a small effect size, 0.5 to a medium effect size, and 0.8 to a large effect size [85]. To complement the analysis, a Spearman correlation matrix was produced looking at the correlation between perceived control, perceived immersion, and perceived exertion for both conditions, following the guidelines of ≤0.35 corresponding to a weak correlation, 0.36 to 0.67 corresponding to a moderate correlation, and 0.68 to 1 corresponding to a strong correlation [86].

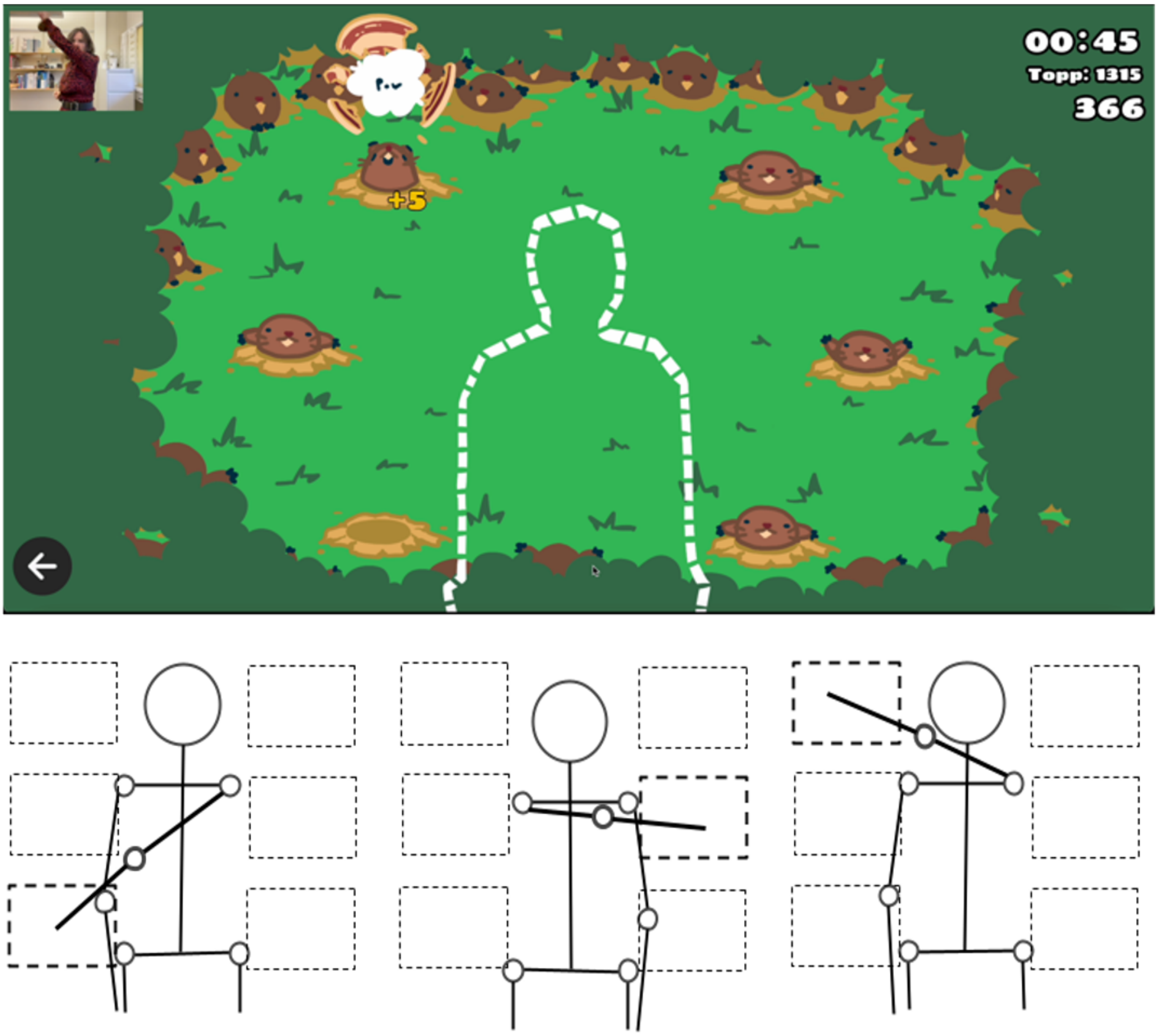
Study 2: Standing Versus Seated Positions for Upper Body Movement Interactions

The second study replicated our previous study on seated and standing exergames [41], using a different casual exergame and with the addition of using the participant's actual heart rate as an objective measure of exertion.

Casual Exergame (Beaver): Design and Interaction

The in-house–developed casual exergame Beaver was used in this study. The goal of Beaver is to hit targets on the screen as fast as possible within the time limit of 2 minutes. In the exergame, beavers holding target boards appear on the screen in 6 different positions according to a specific pattern. The player receives points based on how fast the target is hit (slow hit=1 point, moderately fast hit=3 points, and very fast hit=5 points). The order in which the targets appear follows different patterns but is repeated over 2 rounds, allowing players to learn the current pattern and hit the targets faster in the second round. To hit the targets, the player moves the arm on the opposite side across the body toward the target's position, as shown in Figure 3. Targets alternate between the left and right side of the player's head, shoulders, and hips so that the player keeps hitting the targets across the torso with both arms. This movement is similar to boxing because prior research on boxing exergames has demonstrated that boxing-like movements can result in moderate levels of exertion [39,60-63]. The decision to avoid a boxing game graphics theme was to maintain consistency with the game-like, rather than sport-like, theme and appearance used in the Crossing exergame in study 1.

Figure 3. The casual exergame Beaver with boxing-like upper body movement interactions.

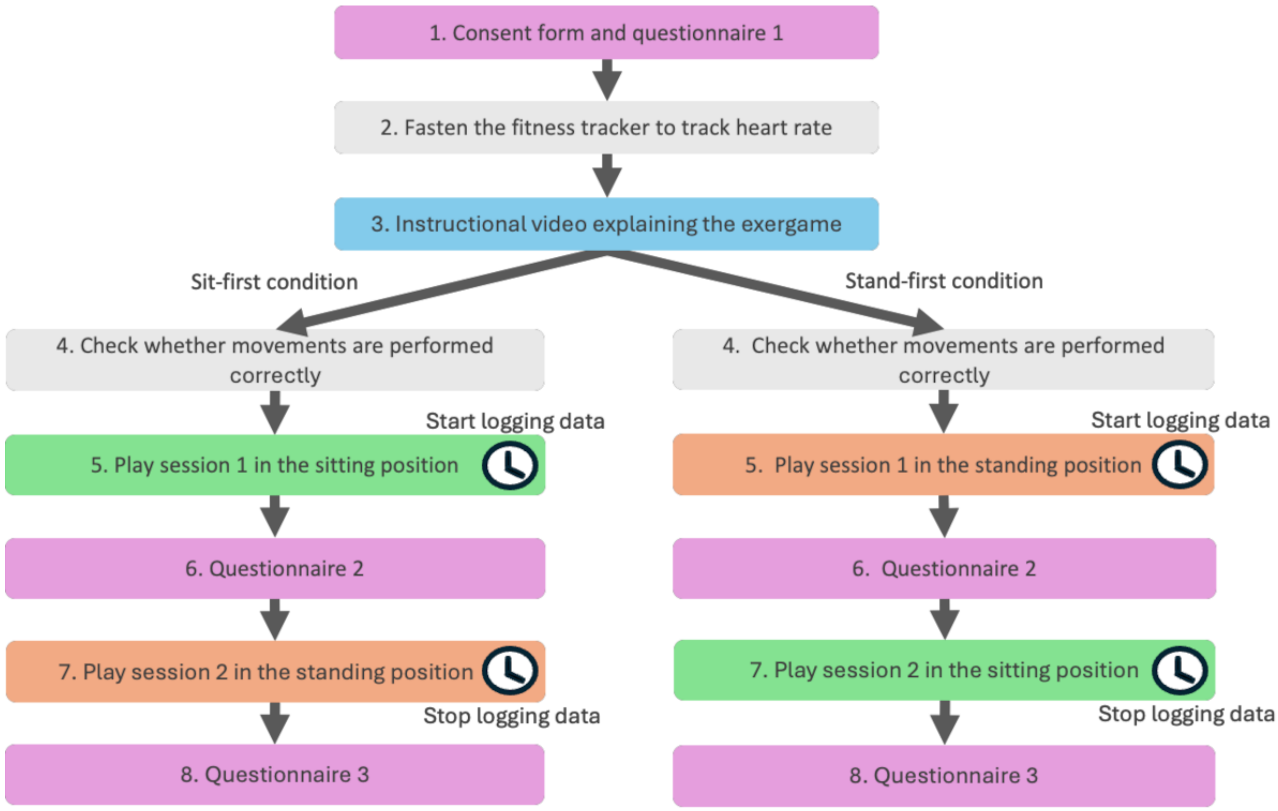


Participants and Procedure

In all, 40 people participated in the study, and their ages ranged from 18 to 47 (mean 25.27, SD 6.357) years. Of the 40 participants, 23 (58%) identified as man, 12 (30%) as woman, and 4 (10%) as nonbinary, while 1 person (2%) reported that they were unsure about their gender identity. Of the 40 participants, 19 (48%) reported as work was their main occupation, 17 (42%) were students, 2 (5%) worked as well as studied, and 2 (5%) had some other main occupation. Of the 40 participants, 4 (10%) had never played an exergame, 17 (42%) had played an exergame at some point, 14 (35%) sometimes played exergames, 4 (10%) played exergames somewhat often, and 1 (2%) often played exergames.

After signing the consent form, participants answered a questionnaire that included (1) demographic questions and (2) the international, shortened version of the Positive and Negative Affect Schedule (PANAS) [87,88]. A Polar Unite fitness tracker was fastened around each participant's wrist to track their heart rate during the play sessions. The participants were then shown an instructional video that explained how to play the exergame. To avoid possible bias, the video was recorded in such a way that it was impossible to tell whether the instructor was sitting or standing. Conditions were alternated for each participant, resulting in half of the participants (20/40, 50%) playing Beaver in the sit-first condition (seated in a chair) and the other half (20/40, 50%) playing in the stand-first condition. Participants were placed 1.5 meters away from a computer screen equipped with a webcam in both conditions. The webcam was then adjusted so that the participant was clearly visible, and the participant was asked to demonstrate the movements they should use. Once the participant had confirmed that they had understood the exergame and movements by demonstrating them correctly, the fitness tracker was set to start logging the heart rate, and the participant started the first game session. The time was also noted to correlate the heart rate data to the questionnaire results. After the session ended, the final score was noted, while the participant filled in the second questionnaire, which included (1) the PANAS [87], (2) the Borg RPE scale [40], and (3) the Physical Activity Enjoyment Scale (PACES) [70,89]. Next, the participant was again guided to sit or stand in front of the computer in the position they had not played before. The webcam was adjusted to the new position, and the participant played a second round of the exergame. The starting time of the second gameplay session was again noted for purposes of later identification in the data. When the gameplay session ended, the heart rate data logging was stopped, and the final score of the second gameplay session was noted while the participant filled out the third and final questionnaire, which included the same scales as the second questionnaire. The entire procedure is shown in Figure 4.

Figure 4. Study 2 procedure.



Measures

Physical exertion of playing exergames can be measured using subjective perception (eg, using the Borg RPE scale [75-77] as in study 1) and objective measures (eg, capturing the actual heart rate) [79]. As there is a strong correlation between heart rate and energy expenditure, heart rate monitors can be used to measure actual exertion in exergames [90]. Borg RPE scores and heart rate measures are positively correlated during exergame play [91]. Thus, integrating both subjective and objective data offers a comprehensive assessment of physical exertion, enhancing accuracy in evaluating exergame physical exertion.

The participant's actual heart rate was measured using a Polar Unite fitness tracker fastened around their wrist during the gameplay sessions, while the perceived exertion was measured using the Borg RPE scale [75-77], which corresponds to the maximum heart rate [75,78] (Table 1). The fitness tracker logs the wearer's heart rate during a recorded training session and displays the heart rate over time in a graph. The highest and average heart rates during the first gameplay session and the second gameplay session were compared to the participant's estimated maximum heart rate (calculated using the following formula: maximum heart rate = 220 – age in years) to obtain a percentage used to determine exercise intensity.

After study 1, some potential issues with measuring immersion in very short exergames were discovered, as described in the Discussion section. Therefore, in study 2, psychological effects were instead measured in terms of affect, enjoyment, and performance. Positive and negative affect before and after playing Beaver were measured using the international, shortened version of the PANAS [88]. The PANAS was selected because of its established reliability and validity in assessing affect [92], its proven efficacy in measuring changes in affect before and after shorter interventions [93], and its frequent use in research on exergames [94]. The shortened PANAS contains 10 validated items—5 for positive affect and 5 for negative affect—and has a total score for both positive and negative affect that ranges from 5 to 25 based on the summation of its items [87]. In this study, only the positive affect scale of the PANAS was analyzed because the negative affect scale exhibited low variation and, as a result, issues with internal consistency, with values close to 0. As the positive and negative affect scales are uncorrelated and independent [88], eliminating the negative affect scale does not impact the findings related to the positive affect scale.

Physical activity enjoyment was measured using the shortened version of the PACES [70,89]. Five of the original 18 statements are included in the shortened version, which has demonstrated strong reliability for exergames across age groups [70]. In the PACES, the player rates how much they agree with each statement using a 7-point Likert scale.

Finally, performance was measured using the participants' final scores. These scores indicated the number of times the participant had successfully jumped forward on the path.

Statistical Analyses

Two main statistical analyses were conducted. First, a within-subjects comparison of the difference in positive affect before and after playing the first round was carried out for both the participants playing the first gameplay session seated and the participants playing the first gameplay session standing. The second analysis was a within-subjects comparison of the differences in heart rate, perceived exertion, affect, enjoyment, and performance between the seated and standing positions for each participant.

For the first analysis, Cronbach α was calculated for positive and negative affect before and after the first gameplay session for participants playing the first gameplay session standing and seated. Internal consistency was acceptable for positive affect before playing for both seated (0.67) and standing (0.75) gameplay, but while negative affect had acceptable consistency for seated (0.70) gameplay, it was low for standing gameplay (0.22). After playing 1 session of the exergame, the internal consistency for positive affect was high for both seated (0.83) and standing (0.86) gameplay, but, once again, the consistency for negative affect was low for standing (0.15) gameplay while remaining high for seated (0.89) gameplay. Because of the low internal consistency for negative affect, only positive affect was analyzed. To evaluate the difference in positive affect after playing 1 session of the exergame, the difference between the positive affect before and after playing the exergame was calculated. The assumption behind the dependent 2-tailed *t* test was checked in terms of outliers and normality. No outliers were found for either sitting or standing gameplay, and the Shapiro-Wilk test statistic showed that the standing group had no significant deviation from normality at a significance of P=.94. However, the seated group deviated from normality with a significance of P=.04, which is less than the threshold of P=.50. Therefore, the nonparametric Wilcoxon signed rank test was run on the positive affect before and after the first seated and standing gameplay sessions. Effect size was calculated using the formula $r = z/\text{sqrt}(N)$ [84].

For the second analysis, the differences between each participant's seated and standing gameplay were calculated. Cronbach α was calculated for the PACES (seated=0.83 and standing=0.83) as well as positive (seated=0.89 and standing=0.88) and negative (seated=0.78 and standing=0.73) affect. As all scales showed high internal consistency, their items were summed. Next, the differences between standing and seated values were calculated for the PACES, positive and negative affect, Borg RPE, percentage of maximum heart rate reached (highest and average), and game score for each participant and analyzed to check assumptions of normality. Negative affect had multiple extreme outliers and was heavily skewed with a Shapiro-Wilk significance of <.001 and was thus excluded from further analysis. No other score showed deviation from normality according to the Shapiro-Wilk test statistic (positive affect=.13, PACES=.50, Borg RPE=.07, highest heart rate percentage=.69, average heart rate percentage=.39, and game score=.50). While the game score had 1 outlier, it was not extreme, and we decided to keep it for analysis due to its being a possible increase in score between gameplay sessions. The dependent *t* test could thus be used to analyze the differences between standing and seated gameplay.

Ethical Considerations

According to the Swedish Ethical Review Act [95], this study did not require ethical review, as it posed no apparent risk of physical or psychological harm to the research subjects, did not involve a physical intervention, and did not involve sensitive personal data. Nonetheless, all procedures adhered to ethical standards outlined in Swedish law (SFS 2003:460). Participants were verbally invited to join the study and were provided with an overview of the project and study procedures. Upon agreeing to participate, individuals received a written information letter and were asked to sign the written consent form. Participants were informed of their right to withdraw from the study at any time without providing a reason. To ensure anonymity, each participant was assigned a unique code that could not be traced back to them. Only the project team had access to these coded data.

Results

Study 1: Upper Body Versus Full Body Movement Interactions

Overall, the participants perceived the exergame as controllable and immersive, with perceived exertion corresponding to a low-intensity activity for both the upper body and full body movement interactions (Table 3). In the upper body condition, participants on average agreed that the casual exergame was controllable (median 4.00, IQR 0.88) and were between being neutral and somewhat agreeing that the casual exergame was immersive (median 3.40, IQR 0.90), with a perceived exertion corresponding to a low-intensity activity (median 11.00, IQR 2.00). In the full body condition, participants perceived the exergame as less controllable (median 3.50, IQR 1.00) than, and as immersive (median 3.40, IQR 1.00) as, the upper body condition, with a perceived exertion corresponding to a moderate-intensity activity (median 13.00, IQR 3.50). The statistical analysis (Table 4) showed that the exergame was perceived as significantly more controllable in the upper body condition than in the full body condition (U=589.00; P=.04), corresponding to a small effect size (0.23). However, there was no statistically significant difference between the conditions regarding perceived immersion (U=754.00; P=.66) and perceived exertion (U=652.00; P=.15).

Table 3. Descriptive statistics for upper body and full body movement interactions.

Measure	Upper body		Full body	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Sideways movements, n	12.88 (6.57)	11.50 (7.50)	9.58 (4.86)	9.00 (8.00)
Forward movements, n	62.30 (13.54)	62.50 (16.00)	56.98 (11.38)	57.50 (16.50)
Performance	58.53 (12.93)	58.50 (16.00)	47.60 (13.98)	46.50 (22.50)
Perceived control	3.78 (0.61)	4.00 (0.88)	3.48 (0.69)	3.50 (1.00)
Perceived immersion	3.24 (0.62)	3.40 (0.90)	3.15 (0.65)	3.40 (1.00)
Perceived exertion	11.47 (2.10)	11.00 (2.00)	12.00 (2.60)	13.00 (3.50)

Table 4. Differences between upper body and full body movement interactions.

Variable	Δmedian (UB[a]–FB[b])	Mann-Whitney *U* test	z score	*P* value	Effect size (*r*)
Perceived control[c]	0.5	589.00	–2.05	.04[c]	0.23
Perceived immersion	0.0	754.00	–0.45	.66	0.05
Perceived exertion	−2.0	652.00	–1.45	.15	0.16

[a]UB: upper body game controller.

[b]FB: full body game controller.

[c]Significant effect at *P*<.05.

The Spearman correlation matrix (Table 5) showed a significant moderate correlation between perceived immersion and perceived control (0.44; *P*<.001). No significant correlation was found between perceived exertion and perceived control (0.07; *P*=.51) and between perceived exertion and perceived immersion (0.11; *P*=.33).

Table 5. Spearman correlation matrix of perceived control, immersion, and exertion.

Variable	Perceived control	Perceived immersion	Perceived exertion
Perceived control			
r	—[a]	0.44	0.07
P value	—	<.001[b]	.51
Perceived immersion			
r	0.44	—	0.11
P value	<.001[b]	—	.33
Perceived exertion			
r	0.07	0.11	—
P value	.51	.33	—

[a]Not applicable.

[a]*P* value met the threshold for significance.

Study 2: Standing Versus Seated Positions for Upper Body Movement Interactions

There was an increase in positive affect scores from before (mean 13.90, SD 3.43; median 14.00, IQR 4.50) to after (mean 16.00, SD 4.61; median 15.00, IQR 7.00) playing one 2-minute session of the exergame in the seated position. The Wilcoxon signed rank test showed that this difference was significant (*z*=–3.288; *P*=.001), and the effect size (*r*=0.74) corresponded to a large effect [85]. There was also an increase in positive affect scores from before (mean 13.45, SD 3.69; median 13.50, IQR 5.00) to after (mean 16.00, SD 4.19; median 16.00, IQR 7.00) playing one 2-minute session of the exergame in the standing position. The Wilcoxon signed rank test showed that this was also significant (*z*=–2.947; *P*=.003), and the effect size (*r*=0.66) corresponded to a large effect [85].

The highest percentage of their maximum heart rate that participants reached during standing gameplay (mean 63.29%, SD 7.40%) was slightly higher than that reached during seated gameplay (mean 60.46%, SD 7.38%), which was significant (t_{39}=2.805; *P*=.008), with an effect size (Cohen *d*=0.44) corresponding to a small effect. Similarly, the average heart rate (as a percentage of the estimated maximum heart rate) was also slightly higher in standing gameplay (mean 55.71%, SD

7.78%) than in seated gameplay (mean 52.63%, SD 6.92%). This was again statistically significant (t_{39}=2.542; *P*=.02), with an effect size (Cohen *d*=0.40) corresponding to a small effect.

Standing gameplay also showed slightly higher scores in positive affect, enjoyment, and perceived exertion than seated gameplay, while seated gameplay had slightly higher scores in performance than standing gameplay. However, none of these differences were significant (Table 6).

Table 6. Statistical analysis of seated gameplay compared to standing gameplay.

Measure	Seated gameplay, mean (SD)	Standing gameplay, mean (SD)	Seated gameplay, median (IQR)	Standing gameplay, median (IQR)	*t* test (*df*)	*P* value	Effect size (Cohen *d*)
Positive affect	15.80 (4.85)	16.13 (4.82)	15.00 (8.50)	15.50 (7.00)	1.114 (39)	.27	0.18
Enjoyment	25.38 (5.67)	25.90 (5.66)	25.00 (9.50)	25.50 (8.50)	0.894 (39)	.38	0.14
Perceived exertion	10.60 (2.04)	10.95 (2.40)	11.00 (3.00)	11.00 (3.00)	1.300 (39)	.20	0.21
Average heart rate (%)	52.63 (6.92)	55.71 (7.78)	51.93 (1.74)	55.21 (0.97)	2.542 (39)	.02[a]	0.40
Highest heart rate (%)	60.46 (7.38)	63.29 (7.40)	60.21 (8.44)	62.24 (8.93)	2.805 (39)	.008[a]	0.44
Performance	702.30 (109.52)	695.50 (94.13)	716.50 (154.50)	706.00 (116)	–0.409 (39)	.69	0.07

[a]*P* value met the threshold for significance.

Discussion

This work aimed to study the psychological effects and physical exertion of playing short-duration casual exergames lasting 2 minutes based on upper body movement interactions designed to promote active microbreaks for people who are sedentary. In the first study, upper body movement interactions were compared to full body movement interactions, and in the second study, seated gameplay was compared to standing gameplay when using upper body movement interactions.

Principal Findings

This work demonstrated that (1) upper body movement interactions in casual exergames can be as effective and appealing as full body movement interactions; and (2) playing upper body casual exergames in standing positions can result in slightly higher effectiveness than, and a similar degree of attractiveness as, playing upper body casual exergames in seated positions.

In terms of the psychological effects of casual exergames, study 1 showed that the upper body casual exergame controller was perceived as more controllable than the full body controller and that the level of immersion was similar for both conditions. As perceived control is seen as a prerequisite to an immersive experience [50] and was found to positively correlate to immersion in this study, longer-duration gameplay could have resulted in the upper body condition being perceived as more immersive over time; for example, 1 study found that short-duration gameplay lasting 3 minutes resulted in less immersion than long-duration gameplay lasting 7 minutes [96]. The short-duration gameplay session lasting 2 minutes in this study could thus have resulted in lower immersion than if the exergame had been played for longer durations. Further research would be needed to investigate whether this holds true. In study 2, the psychological effects were similar for the seated and standing upper body movement conditions, with no significant difference in positive affect, enjoyment, or performance. Furthermore, the results showed that both seated and standing upper body casual exergame movement interactions significantly increased positive affect after participants played a 2-minute session. The result signifies that seated and standing casual exergame play could both be valid movement interaction options for upper body casual exergames.

In terms of physiological effectiveness, study 1 showed that perceived exertion did not differ for the upper body and full body casual exergame conditions. Similarly, in study 2, there was no significant difference between perceived exertion for the standing and seated upper body casual exergame conditions. However, there was a difference in terms of the objective measure of exertion. The standing group had a significantly higher heart rate than the sitting group for both the highest and average heart rates, although the effect size was small for both. As no objective measure of exertion was taken in study 1, a similar pattern could have emerged for the upper body conditions compared to the full body conditions. The upper body movement interactions provided a light level of perceived exertion in both study 1 (mean 11.47, SD 2.10), with the upper body movement corresponding to players lifting their arms, and study 2 (mean 10.60, SD 2.04 for seated gameplay and mean 10.95, SD 2.40 for standing gameplay), with the upper body movement corresponding to boxing-like movements. However, while the highest heart rates reached also showed a light level of exertion (mean 60.46%, SD 7.38% for seated gameplay and mean 63.29%, SD 7.40% for standing gameplay), the average heart rates in study 2 (mean 52.63%, SD 6.92% for seated gameplay and mean 55.71%, SD 7.78% for standing gameplay) only indicate a very light–intensity activity, which might also be the case for study 1. On the basis of previous research claiming that casual exergames should produce at least a moderate level of exertion [33], neither exergame thus reached sufficient levels of exertion. Despite this, both exergames may be considered appropriate tools for microbreaks during the day, considering that every move contributes to better health [69], and microbreaks (2-3 min of light-intensity exercise) during the day (every 30 min) still offer both physical and mental health benefits [20-22]. The higher heart rate averages displayed in study 2 imply that people could gain some benefit from playing upper body casual exergames in standing positions compared to seated positions. However, due to the small difference in effect size and with both measures corresponding to very light activity, playing seated should be seen as a reasonable option for those who prefer to do so. Further research is necessary to ascertain the long-term effects on physical health of playing in seated or standing positions.

Comparison With Prior Work

The importance of considering both the psychological effects and the physical exertion of playing exergames has been highlighted previously [46-48, 73]. This study focused on how the movement interactions used in casual exergames could affect these dimensions.

Previous research has shown that playing exergames in both standing and seated positions produces equivalent levels of enjoyment (psychological attractiveness) [41,64,65,71,72]. The results from this work further support these findings, showing no difference between the seated and standing positions in terms of enjoyment; in addition, they support the notion that the psychological effects are also similar in terms of positive affect and performance [41]. Previous research has also shown that playing casual exergames can induce positive affective states [31,32,41]. The results of this study support this notion, with both the seated and standing positions generating an increase in positive affect. As positive affect can reduce stress levels [42], enhance overall well-being [43], and improve work performance [97], casual exergames could be promising for implementation in a workplace context to promote active microbreaks.

Regarding physiological effectiveness, previous studies have found mixed results when comparing standing and seated exergaming. While some studies have found higher energy expenditure in seated gameplay [64,65], others have found higher energy expenditure in standing gameplay [68,72]. Our results in study 2 support that standing gameplay involves higher objective exertion than seated gameplay, with no difference in subjective measures. The higher exertion in the standing position is likely due to using the leg muscles to stand (although they are not used to play the exergame) because even sedentary activities require more energy when standing instead of sitting [66,67]. When rating perceived exertion, this difference might be small enough that people do not register it because the exertion in their arms is more noticeable after playing, thus leading to similarly rated exertion despite differing heart rates. The results show that the difference in heart rate percentage is only approximately 3% for both the highest and average heart rates, which may feel very similar. As for why the participants rated their exertion higher than their heart rates show, it is possible that the average over the 2 minutes of gameplay does not reflect the exertion level at the end, which is when the participants rated their exertion. When beginning an activity, the heart rate rises gradually. With the short playtimes, this initial increase might affect the average heart rate more than it would in a longer activity. The highest heart rate reached was typically toward the end of the sessions and was also more closely matched to the perceived exertion.

Previous studies have also shown mixed results when comparing upper body to lower body movement interactions. It has been suggested by some studies that upper body exergaming might not be sufficient for achieving adequate energy expenditure [46,59]. Meanwhile, other studies have suggested that certain upper body movements could produce sufficient exertion levels [39,60-63]. This study found no difference in perceived exertion between upper body and full body movement interactions, somewhat supporting the notion that upper body movements could produce sufficient exertion. However, because no objective measures of exertion were recorded in study 1, more research would be needed to validate that this holds true not only for subjective measures of exertion but also for objective measures.

A meta-review shows that upper body movement interactions characterized by continuous movements can result in greater energy expenditures and intensity levels [39]. Previous studies on boxing exergames have shown that boxing-like movements can achieve a moderate intensity of exertion [39,60-63]. However, both the perceived exertion and the heart rate data showed that the exergame in study 2, which involved using boxing-like movements, only reached a light or very light intensity of exertion. A potential explanation for this difference could be the time participants spent playing because a 2-minute exergame session will feel less exerting than the same exergame played for twice as long or more. Another explanation could be that the design of the exergame influenced the amount of physical effort the participants were willing to exert; for example, the exergame in study 2 features short pauses between sets of targets, which could have contributed to participants not reaching moderate levels of exertion. The interplay between gameplay duration, movement design, and exergame design is an interesting avenue for further research.

Limitations

Integrating active microbreaks into the day via casual exergames has shown promising advantages [32,37]. However, both studies took place at a game and cosplay festival at a specific point in time. The 2-minute casual exergames may influence positive affect in distinct ways when experienced in other settings; for example, having an audience of peers sharing an interest in gaming might be experienced as more positive than playing alone. Furthermore, because the number of people observing the players varied throughout the day depending on other activities at the festival, participants may also have had different experiences depending on their reactions to feeling observed. Conducting the studies in a more controlled setting might produce different results. In addition, continuous gameplay might play a role in influencing changes in positive affect over time. Further research should also be undertaken to quantify the findings of these studies by using other casual exergames.

Participants in both studies played the exergames for the first time. Thus, the lack of familiarity may have influenced their impression of the movement interactions in both study 1 (upper body movement interactions compared to full body movement interactions) and study 2 (upper body movement interactions in seated compared to standing positions). Further research should incorporate a phase allowing participants to familiarize themselves with the exergame controllers before the start of the study. This would help determine whether the findings remain applicable when participants have fully mastered the exergame controls. As this study only covered players' initial experience with casual exergames, there is a need for longitudinal studies on the psychological attractiveness and physiological effectiveness of different movement interactions in casual exergames and how they should be constructed to increase the psychological attractiveness and physical effectiveness over time.

The results regarding exertion in study 1 are also limited due to a lack of objective measures of exertion. As seen in study 2, there could be differences in the full body and upper body movement interactions in terms of exertion that are not captured through only perceived exertion. Further research using objective measures (eg, heart rate) should be undertaken to better understand the effectiveness of full body and upper body movement interactions in short casual exergames. Furthermore, both studies could greatly benefit from using more extensive measures of physical activity, such as metabolic equivalents of tasks (METs) and maximal oxygen consumption, to examine the extent of any health benefits of playing short-duration casual exergames in different player modes.

Conclusions

Short-duration casual exergames lasting 2 minutes with upper body movements may hold potential in promoting active microbreaks during sedentary periods. As this study showed, upper body casual exergame play can produce light exertion levels and an increase in positive affect after compared to before playing. Upper body movement interactions may also be more suitable for casual exergames used for microbreaks because they are perceived as easier to control than full body movement interactions, while reaching similar (light) exertion levels and immersion. Playing with upper body movement interactions in standing gameplay could involve slightly higher objective exertion (as measured by the heart rate) than when playing the same exergame in a seated position; however, due to the small difference, seated positions should not be neglected as a viable alternative for players who prefer them. Furthermore, upper body movement interactions in both seated and standing positions elicit comparable psychological effects, resulting in similar levels of enjoyment, positive affect, and performance.

Acknowledgments

The authors would like to thank the users for participating in the study and East Sweden Game for supporting and enabling the data collection. The authors would also like to thank Erik Berglund who developed the exergames used in this research and Izabella Jedel for assisting in the data collection and analysis.

Authors' Contributions

AB and HO contributed to the conception and design of the study. AB contributed to the design of the exergames used in the study. AB and a research assistant, Izabella Jedel, collected the data for study 1. AB and HO collected the data for study 2. HO and a research assistant, Izabella Jedel, performed the statistical analysis. AB wrote the first draft of the manuscript. Both AB and HO then carefully reviewed and edited the manuscript.

Conflicts of Interest

AB is a cofounder and the chief executive officer of Liopep, a spin-off company based on her research projects about exergames that aim to increase physical activity. Liopep is commercializing an app that includes the 2 exergames evaluated in this research. While the authors have taken steps to ensure objectivity, readers should be aware of this affiliation. No revenue was paid (or promised to be paid) to any of the authors or the university for the submission of this work. Transparency and research integrity have been maintained through collaborative efforts and rigorous peer review. AB was excluded from the data analysis phase of both studies to ensure that the conflict of interest did not bias the research process and outcomes. In addition, AB did not collect the data alone; data collection was conducted collaboratively by the entire author team. The study design was also a joint effort by the author team, further ensuring the integrity and impartiality of the research findings. HO declares no conflicts of interest.

References

1. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75. [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
2. Sedentary Behaviour Research Network. Letter to the editor: standardized use of the terms "sedentary" and "sedentary behaviours". Appl Physiol Nutr Metab 2012 Jun;37(3):540-542 [FREE Full text] [doi: 10.1139/h2012-024] [Medline: 22540258]
3. Ekelund U, Tarp J, Steene-Johannessen J, Hansen BH, Jefferis B, Fagerland MW, et al. Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. BMJ 2019 Aug 21;366:l4570 [FREE Full text] [doi: 10.1136/bmj.l4570] [Medline: 31434697]
4. Stamatakis E, Gale J, Bauman A, Ekelund U, Hamer M, Ding D. Sitting time, physical activity, and risk of mortality in adults. J Am Coll Cardiol 2019 Apr 30;73(16):2062-2072 [FREE Full text] [doi: 10.1016/j.jacc.2019.02.031] [Medline: 31023430]

5. Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. Exerc Sport Sci Rev 2010 Jul;38(3):105-113 [FREE Full text] [doi: 10.1097/JES.0b013e3181e373a2] [Medline: 20577058]
6. Bauman AE, Petersen CB, Blond K, Rangul V, Hardy LL. The descriptive epidemiology of sedentary behaviour. In: Leitzmann MF, Jochem C, Schmid D, editors. Sedentary Behaviour Epidemiology. Cham, Switzerland: Springer; 2018:73-106.
7. Egan CA, Webster CA, Beets MW, Weaver RG, Russ L, Michael D, et al. Sedentary time and behavior during school: a systematic review and meta-analysis. Am J Health Educ 2019 Sep 03;50(5):283-290. [doi: 10.1080/19325037.2019.1642814]
8. Kuzik N, da Costa BG, Hwang Y, Verswijveren SJ, Rollo S, Tremblay MS, et al. School-related sedentary behaviours and indicators of health and well-being among children and youth: a systematic review. Int J Behav Nutr Phys Act 2022 Apr 05;19(1):40 [FREE Full text] [doi: 10.1186/s12966-022-01258-4] [Medline: 35382825]
9. Castro O, Bennie J, Vergeer I, Bosselut G, Biddle SJ. How sedentary are university students? A systematic review and meta-analysis. Prev Sci 2020 Apr 23;21(3):332-343. [doi: 10.1007/s11121-020-01093-8] [Medline: 31975312]
10. Prince SA, Elliott CG, Scott K, Visintini S, Reed JL. Device-measured physical activity, sedentary behaviour and cardiometabolic health and fitness across occupational groups: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2019 Apr 02;16(1):30 [FREE Full text] [doi: 10.1186/s12966-019-0790-9] [Medline: 30940176]
11. Parry S, Straker L. The contribution of office work to sedentary behaviour associated risk. BMC Public Health 2013 Apr 04;13(1):296 [FREE Full text] [doi: 10.1186/1471-2458-13-296] [Medline: 23557495]
12. Pettee Gabriel KK, Morrow JR, Woolsey AL. Framework for physical activity as a complex and multidimensional behavior. J Phys Act Health 2012 Jan;9 Suppl 1(s1):S11-S18 [FREE Full text] [doi: 10.1123/jpah.9.s1.s11] [Medline: 22287443]
13. Healy GN, Matthews CE, Dunstan DW, Winkler EA, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06. Eur Heart J 2011 Mar;32(5):590-597 [FREE Full text] [doi: 10.1093/eurheartj/ehq451] [Medline: 21224291]
14. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
15. Park JH, Moon JH, Kim HJ, Kong MH, Oh YH. Sedentary lifestyle: overview of updated evidence of potential health risks. Korean J Fam Med 2020 Nov;41(6):365-373 [FREE Full text] [doi: 10.4082/kjfm.20.0165] [Medline: 33242381]
16. Koster A, Caserotti P, Patel KV, Matthews CE, Berrigan D, Van Domelen DR, et al. Association of sedentary time with mortality independent of moderate to vigorous physical activity. PLoS One 2012 Jun 13;7(6):e37696 [FREE Full text] [doi: 10.1371/journal.pone.0037696] [Medline: 22719846]
17. Chau JY, Grunseit AC, Chey T, Stamatakis E, Brown WJ, Matthews CE, et al. Daily sitting time and all-cause mortality: a meta-analysis. PLoS One 2013 Nov 13;8(11):e80000 [FREE Full text] [doi: 10.1371/journal.pone.0080000] [Medline: 24236168]
18. Saunders TJ, McIsaac T, Douillette K, Gaulton N, Hunter S, Rhodes RE, et al. Sedentary behaviour and health in adults: an overview of systematic reviews. Appl Physiol Nutr Metab 2020 Oct;45(10 (Suppl. 2)):S197-S217 [FREE Full text] [doi: 10.1139/apnm-2020-0272] [Medline: 33054341]
19. Giurgiu M, Koch ED, Plotnikoff RC, Ebner-Priemer UW, Reichert M. Breaking up sedentary behavior optimally to enhance mood. Med Sci Sports Exerc 2020 Feb;52(2):457-465. [doi: 10.1249/MSS.0000000000002132] [Medline: 31453884]
20. Radwan A, Barnes L, DeResh R, Englund C, Gribanoff S. Effects of active microbreaks on the physical and mental well-being of office workers: a systematic review. Cogent Eng 2022 Jan 19;9(1):1-12. [doi: 10.1080/23311916.2022.2026206]
21. Dunstan DW, Kingwell BA, Larsen R, Healy GN, Cerin E, Hamilton MT, et al. Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care 2012 May;35(5):976-983 [FREE Full text] [doi: 10.2337/dc11-1931] [Medline: 22374636]
22. Peddie MC, Bone JL, Rehrer NJ, Skeaff CM, Gray AR, Perry TL. Breaking prolonged sitting reduces postprandial glycemia in healthy, normal-weight adults: a randomized crossover trial. Am J Clin Nutr 2013 Aug;98(2):358-366 [FREE Full text] [doi: 10.3945/ajcn.112.051763] [Medline: 23803893]
23. Mailey EL, Rosenkranz SK, Casey K, Swank A. Comparing the effects of two different break strategies on occupational sedentary behavior in a real world setting: a randomized trial. Prev Med Rep 2016 Dec;4:423-428 [FREE Full text] [doi: 10.1016/j.pmedr.2016.08.010] [Medline: 27583200]
24. Bianchi-Berthouze N, Kim WW, Patel D. Does body movement engage you more in digital game play? And why? In: Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction. 2007 Presented at: ACII '07; September 12-14, 2007; Lisbon, Portugal p. 12-14 URL: https://link.springer.com/chapter/10.1007/978-3-540-74889-2_10 [doi: 10.1007/978-3-540-74889-2_10]
25. Sween J, Wallington SF, Sheppard V, Taylor T, Llanos AA, Adams-Campbell LL. The role of exergaming in improving physical activity: a review. J Phys Act Health 2014 May;11(4):864-870 [FREE Full text] [doi: 10.1123/jpah.2011-0425] [Medline: 25078529]
26. Wüest S, van de Langenberg R, de Bruin ED. Design considerations for a theory-driven exergame-based rehabilitation program to improve walking of persons with stroke. Eur Rev Aging Phys Act 2014 Dec 7;11(2):119-129 [FREE Full text] [doi: 10.1007/s11556-013-0136-6] [Medline: 25309631]

27. Huang K, Zhao Y, He R, Zhong T, Yang H, Chen Y, et al. Exergame-based exercise training for depressive symptoms in adults: a systematic review and meta-analysis. Psychol Sport Exerc 2022 Nov;63:102266. [doi: 10.1016/j.psychsport.2022.102266]
28. Oh Y, Yang S. Defining exergames and exergaming. In: Proceedings of the 2010 Conference on Meaningful Play. 2010 Presented at: MP '10; October 21-23, 2010; East Lansing, MI p. 21-23 URL: https://meaningfulplay.msu.edu/proceedings2010/
29. Klompstra L, Jaarsma T, Piepoli MF, Ben Gal T, Evangelista L, Strömberg A, et al. Objectively measured physical activity in patients with heart failure: a sub-analysis from the HF-Wii study. Eur J Cardiovasc Nurs 2022 Jun 30;21(5):499-508 [FREE Full text] [doi: 10.1093/eurjcn/zvab133] [Medline: 34993536]
30. Oyebode O, Orji R. A persuasive mobile game for reducing sedentary behaviour and promoting physical activity. In: Proceedings of the 15th International Conference on Persuasive Technology. 2020 Presented at: ICPT '20; April 20-23, 2020; Aalborg, Denmark p. 1-4 URL: https://ceur-ws.org/Vol-2629/11_poster_oyebode.pdf
31. Gao Y, Gerling KM, Mandryk R, Stanley KG. Decreasing sedentary behaviours in pre-adolescents using casual exergames at school. In: Proceedings of the 1st ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play. 2014 Presented at: CHI PLAY '14; October 19-21, 2014; Toronto, ON p. 97-106 URL: https://dl.acm.org/doi/10.1145/2658537.2658693 [doi: 10.1145/2658537.2658693]
32. Gao Y, Mandryk R. The acute cognitive benefits of casual exergame play. In: Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 1863-1872 URL: https://dl.acm.org/doi/10.1145/2207676.2208323 [doi: 10.1145/2207676.2208323]
33. Gao Y, Mandryk RL. GrabApple: the design of a casual exergame. In: Proceedings of the 10th International Conference on Entertainment Computing. 2011 Presented at: ICEC '11; October 5-8, 2011; Vancouver, BC p. 35-46 URL: https://link.springer.com/chapter/10.1007/978-3-642-24500-8_5 [doi: 10.1007/978-3-642-24500-8_5]
34. Ren X, Hollander L, Van Marel R, Molenaar L, Lu Y. Step-by-step: exploring a social exergame to encourage physical activity and social dynamics among office workers. In: Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI EA '19; May 4-9, 2019; Glasgow, UK p. 1-6 URL: https://dl.acm.org/doi/abs/10.1145/3290607.3312788 [doi: 10.1145/3290607.3312788]
35. Yoo S, Gough P, Kay J. Embedding a VR game studio in a sedentary workplace: use, experience and exercise benefits. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI p. 1-14 URL: https://dl.acm.org/doi/10.1145/3313831.3376371 [doi: 10.1145/3313831.3376371]
36. Cambo SA, Avrahami D, Lee ML. BreakSense: combining physiological and location sensing to promote mobility during work-breaks. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO p. 3595-3607 URL: https://dl.acm.org/doi/10.1145/3025453.3026021 [doi: 10.1145/3025453.3026021]
37. Berglund A, Jedel I, Orädd H, Fallström J, Berglund E. Liopep: a gamified casual exergame application to help office workers not be active couch potatoes. In: Proceedings of the 2023 IEEE 11th International Conference on Serious Games and Applications for Health. 2023 Presented at: SeGAH '23; April 21-26, 2023; Athens, Greece p. 1-8 URL: https://ieeexplore.ieee.org/document/10253779 [doi: 10.1109/segah57547.2023.10253779]
38. Wu PT, Wu WL, Chu IH. Energy expenditure and intensity in healthy young adults during exergaming. Am J Hlth Behav 2015 Jul 01;39(4):556-561. [doi: 10.5993/ajhb.39.4.12]
39. Davis JC, Killen LG, Green JM, Waldman HS, Renfroe LG. Exergaming for physical activity: a systematic review. J Am Coll Health 2022 Jul 26:1-9. [doi: 10.1080/07448481.2022.2103377] [Medline: 35882066]
40. Berglund E, Jedel I, Berglund A. Using mediapipe machine learning to design casual exertion games to interrupt prolonged sedentary lifestyle. In: Proceedings of the 5th International Conference, HCI-Games 2023, Held as Part of the 25th HCI International Conference on HCI in Games. 2023 Presented at: HCII '23; July 23-28, 2023; Copenhagen, Denmark p. 237-251 URL: https://link.springer.com/chapter/10.1007/978-3-031-35930-9_16 [doi: 10.1007/978-3-031-35930-9_16]
41. Berglund E, Orädd H, Jedel I, Berglund A. The potential of seated and standing short duration casual exergames to increase positive affect. In: Proceedings of the 26th International Academic Mindtrek Conference. 2023 Presented at: Mindtrek '23; October 3-6, 2023; Tampere, Finland p. 25-34 URL: https://dl.acm.org/doi/10.1145/3616961.3616964 [doi: 10.1145/3616961.3616964]
42. Gloria CT, Faulk KE, Steinhardt MA. Positive affectivity predicts successful and unsuccessful adaptation to stress. Motiv Emot 2012 Apr 6;37(1):185-193. [doi: 10.1007/s11031-012-9291-8]
43. Meyers MC, van Woerkom M. Effects of a strengths intervention on general and work-related well-being: the mediating role of positive affect. J Happiness Stud 2016 Apr 30;18(3):671-689. [doi: 10.1007/s10902-016-9745-x]
44. Lee S, Kim W, Park T, Peng W. The psychological effects of playing exergames: a systematic review. Cyberpsychol Behav Soc Netw 2017 Sep;20(9):513-532. [doi: 10.1089/cyber.2017.0183] [Medline: 28837353]
45. Berglund A, Jedel I, Orädd H, Berglund E. The attractiveness and effectiveness of upper body and full body casual exergame controllers. In: Proceedings of the 2023 IEEE 11th International Conference on Serious Games and Applications for Health. 2023 Presented at: SeGAH '23; August 28-30, 2023; Athens, Greece p. 28-30 URL: https://ieeexplore.ieee.org/document/10253764 [doi: 10.1109/segah57547.2023.10253764]

46. Peng W, Lin JH, Crouse J. Is playing exergames really exercising? A meta-analysis of energy expenditure in active video games. Cyberpsychol Behav Soc Netw 2011 Nov;14(11):681-688. [doi: 10.1089/cyber.2010.0578] [Medline: 21668370]
47. Subramanian S, Skjæret-Maroni N, Dahl Y. Systematic review of design guidelines for full-body interactive games. Interact Compute 2020;32(4):367-406. [doi: 10.1093/iwc/iwaa026]
48. Sinclair J, Hingston P, Masek M. Considerations for the design of exergames. In: Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia. 2007 Presented at: GRAPHITE '07; December 1-4, 2007; Perth, Australia p. 289-295 URL: https://dl.acm.org/doi/10.1145/1321261.1321313 [doi: 10.1145/1321261.1321313]
49. Sinclair J, Hingston P, Masek M. Exergame development using the dual flow model. In: Proceedings of the 6th Australasian Conference on Interactive Entertainment. 2009 Presented at: IE '09; December 17-19, 2009; Sydney, Australia p. 1-7 URL: https://dl.acm.org/doi/10.1145/1746050.1746061 [doi: 10.1145/1746050.1746061]
50. Brown E, Cairns P. A grounded investigation of game immersion. In: Proceedings of the 2004 Extended Abstracts on Human Factors in Computing Systems. 2004 Presented at: CHI EA '04; April 24-29, 2004; Vienna, Austria p. 1297-1300 URL: https://dl.acm.org/doi/10.1145/985921.986048 [doi: 10.1145/985921.986048]
51. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, et al. Measuring and defining the experience of immersion in games. Int J Hum Comput Interact 2008 Sep;66(9):641-661. [doi: 10.1016/j.ijhcs.2008.04.004]
52. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. New York, NY: Harper Perennial Modern Classics; 1990.
53. Pasch M, Bianchi-Berthouze N, van Dijk B, Nijholt A. Movement-based sports video games: investigating motivation and gaming experience. Entertain Comput 2009 Apr;1(2):49-61. [doi: 10.1016/j.entcom.2009.09.004]
54. Trepte S, Reinecke L. The pleasures of success: game-related efficacy experiences as a mediator between player performance and game enjoyment. Cyberpsychol Behav Soc Netw 2011 Sep;14(9):555-557. [doi: 10.1089/cyber.2010.0358] [Medline: 21342012]
55. Klimmt C, Blake C, Hefner D, Vorderer P, Roth C. Player performance, satisfaction, and video game enjoyment. In: Proceedings of the 8th International Conference on Entertainment Computing. 2009 Presented at: ICEC '09; September 3-5, 2009; Paris, France p. 3-5 URL: https://link.springer.com/book/10.1007/978-3-642-04052-8 [doi: 10.1007/978-3-642-04052-8_1]
56. Lemmens JS, von Münchhausen CF. Let the beat flow: How game difficulty in virtual reality affects flow. Acta Psychol (Amst) 2023 Feb;232:103812 [FREE Full text] [doi: 10.1016/j.actpsy.2022.103812] [Medline: 36566612]
57. Skjæret-Maroni N, Vonstad EK, Ihlen EA, Tan X, Helbostad JL, Vereijken B. Exergaming in older adults: movement characteristics while playing stepping games. Front Psychol 2016 Jun 24;7:964 [FREE Full text] [doi: 10.3389/fpsyg.2016.00964] [Medline: 27445926]
58. Isbister K, Mueller FF. Guidelines for the design of movement-based games and their relevance to HCI. Hum Comput Interact 2014 Dec 22;30(3-4):366-399. [doi: 10.1080/07370024.2014.996647]
59. Graves LE, Ridgers ND, Stratton G. The contribution of upper limb and total body movement to adolescents' energy expenditure whilst playing Nintendo Wii. Eur J Appl Physiol 2008 Nov 8;104(4):617-623. [doi: 10.1007/s00421-008-0813-8] [Medline: 18607619]
60. Barkley JE, Penko AL. Physiologic responses, perceived exertion, and hedonics of playing a physical interactive video game relative to a sedentary alternative and treadmill walking in adults. J Exerc Physiol Online 2009;12(3):12-23 [FREE Full text]
61. Miyachi M, Yamamoto K, Ohkawara K, Tanaka S. METs in adults while playing active video games: a metabolic chamber study. Med Sci Sports Exerc 2010 Jun;42(6):1149-1153. [doi: 10.1249/MSS.0b013e3181c51c78] [Medline: 19997034]
62. O'Donovan C, Hussey J. Active video games as a form of exercise and the effect of gaming experience: a preliminary study in healthy young adults. Physiotherapy 2012 Sep;98(3):205-210. [doi: 10.1016/j.physio.2012.05.001] [Medline: 22898576]
63. Perusek K, Sparks K, Little K, Motley M, Patterson S, Wieand J. A comparison of energy expenditure during "wii boxing" versus heavy bag boxing in young adults. Games Health J 2014 Feb;3(1):21-24. [doi: 10.1089/g4h.2013.0059] [Medline: 26197251]
64. Malone LA, Mendonca CJ, Mohanraj S, Misko SR, Moore J, Brascome JM, et al. Usability of the GAIMplank video game controller for people with mobility impairments: observational study. JMIR Serious Games 2023 Jan 10;11:e38484 [FREE Full text] [doi: 10.2196/38484] [Medline: 36626195]
65. Malone LA, Davlyatov GK, Padalabalanarayanan S, Thirumalai M. Active video gaming using an adapted gaming mat in youth and adults with physical disabilities: observational study. JMIR Serious Games 2021 Aug 26;9(3):e30672 [FREE Full text] [doi: 10.2196/30672] [Medline: 34435962]
66. Roemmich JN. Height-adjustable desks: energy expenditure, liking, and preference of sitting and standing. J Phys Act Health 2016 Oct;13(10):1094-1099. [doi: 10.1123/jpah.2015-0397] [Medline: 27256708]
67. Gibbs BB, Kowalsky RJ, Perdomo SJ, Grier M, Jakicic JM. Energy expenditure of deskwork when sitting, standing or alternating positions. Occup Med (Lond) 2017 Mar 01;67(2):121-127. [doi: 10.1093/occmed/kqw115] [Medline: 27515973]

68. Kafri M, Myslinski MJ, Gade VK, Deutsch JE. Energy expenditure and exercise intensity of interactive video gaming in individuals poststroke. Neurorehabil Neural Repair 2014 Jan 29;28(1):56-65. [doi: 10.1177/1545968313497100] [Medline: 23897906]
69. Every move counts towards better health – says WHO. World Health Organization. 2020. URL: https://www.who.int/news/item/25-11-2020-every-move-counts-towards-better-health-says-who#:~:text=%E2%80%9CPhysical%20activity%20of%20any%20type,more%20physical%20activity%20to%20counter [accessed 2024-04-29]
70. Graves LE, Ridgers ND, Williams K, Stratton G, Atkinson G, Cable NT. The physiological cost and enjoyment of Wii Fit in adolescents, young adults, and older adults. J Phys Act Health 2010 May;7(3):393-401. [doi: 10.1123/jpah.7.3.393] [Medline: 20551497]
71. Xu W, Liang H, He Q, Li X, Yu K, Chen Y. Results and guidelines from a repeated-measures design experiment comparing standing and seated full-body gesture-based immersive virtual reality exergames: within-subjects evaluation. JMIR Serious Games 2020 Jul 27;8(3):e17972 [FREE Full text] [doi: 10.2196/17972] [Medline: 32716004]
72. Malone LA, Thirumalai M, Padalabalanarayanan S, Neal WN, Bowman S, Mehta T. Energy expenditure and enjoyment during active video gaming using an adapted Wii Fit balance board in adults with physical disabilities: observational study. JMIR Serious Games 2019 Feb 01;7(1):e11326 [FREE Full text] [doi: 10.2196/11326] [Medline: 30707098]
73. Berglund E, Jedel I, Orädd H, Berglund A. Considerations for player enjoyment and exertion in casual exergames. In: Proceedings of the 57th Hawaii International Conference on System Sciences. 2024 Presented at: HICSS '24; January 3-6, 2024; Hilton Hawaiian Village, HI p. 1328-1337 URL: https://hdl.handle.net/10125/106541
74. 2008 - 2009 casual games white paper. International Game Developers Association. 2008. URL: https://www.org.id.tue.nl/IFIP-TC14/documents/IGDACasualGames-WhitePaper-2008.pdf [accessed 2024-04-29]
75. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [doi: 10.1249/00005768-198205000-00012]
76. Chen MJ, Fan X, Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. J Sports Sci 2002 Nov 09;20(11):873-899. [doi: 10.1080/026404102320761787] [Medline: 12430990]
77. Williams N. The Borg rating of perceived exertion (RPE) scale. Occup Med 2017;67(5):404-405. [doi: 10.1093/occmed/kqx063]
78. Borg G. Borg's Perceived Exertion and Pain Scales. Champaign, IL: Human Kinetics Publishers; 1998.
79. Yoo S, Ackad C, Heywood T, Kay J. Evaluating the actual and perceived exertion provided by virtual reality games. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing. 2017 Presented at: CHI EA '17; May 6-11, 2017; Denver, CO p. 3050-3057 URL: https://dl.acm.org/doi/10.1145/3027063.3053203 [doi: 10.1145/3027063.3053203]
80. Stewart TH, Villaneuva K, Hahn A, Ortiz-Delatorre J, Wolf C, Nguyen R, et al. Actual vs. perceived exertion during active virtual reality game exercise. Front Rehabil Sci 2022 Aug 8;3:887740 [FREE Full text] [doi: 10.3389/fresc.2022.887740] [Medline: 36189005]
81. Fitzgerald A, Huang S, Sposato K, Wang D, Claypool M, Agu E. The exergame enjoyment questionnaire (eeq): an instrument for measuring exergame enjoyment. In: Proceedings of the 53rd Hawaii International Conference on System Sciences. 2020 Presented at: HICSS '20; January 7-10, 2020; Maui, HI p. 3397-3406 URL: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/fc4f8fcf-b80b-4a7c-abff-92d61ad8b4a1/content [doi: 10.24251/hicss.2020.416]
82. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572 [FREE Full text] [doi: 10.1136/bmj.314.7080.572] [Medline: 9055718]
83. Field A. Discovering Statistics using IBM SPSS Statistics. 4th edition. Thousand Oaks, CA: Sage Publications; 2013.
84. McKnight PE, Najab J. Mann-whitney U test. In: Weiner IB, Craighead WE, editors. The Corsini Encyclopedia of Psychology. Hoboken, NJ: John Wiley & Sons; 2010.
85. Sawilowsky SS. New effect size rules of thumb. J Mod App Stat Meth 2009 Nov 01;8(2):597-599. [doi: 10.22237/jmasm/1257035100]
86. Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr 2016 Jul 01;6(1):35-39. [doi: 10.1177/875647939000600106]
87. Thompson ER. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). J Cross Cult Psychol 2016 Jul 26;38(2):227-242. [doi: 10.1177/0022022106297301]
88. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
89. Kendzierski D, DeCarlo KJ. Physical activity enjoyment scale: two validation studies. J Sport Exerc Psychol 1991;13(1):50-64. [doi: 10.1123/jsep.13.1.50]
90. Whitehead A, Johnston H, Nixon N, Welch J. Exergame effectiveness: what the numbers can tell us. In: Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games. 2010 Presented at: Sandbox '10; July 28-29, 2010; Los Angeles, CA p. 55-62 URL: https://dl.acm.org/doi/10.1145/1836135.1836144 [doi: 10.1145/1836135.1836144]
91. Pollock BS, Barkley JE, Potenzini N, Desalvo RM, Buser SL, Otterstetter R, et al. Validity of borg ratings of perceived exertion during active video game play. Int J Exerc Sci 2013;6(2):164-170 [FREE Full text] [Medline: 27293499]

92. Crawford JR, Henry JD. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol 2004 Sep 24;43(Pt 3):245-265. [doi: 10.1348/0144665031752934] [Medline: 15333231]
93. van Geest J, Samaritter R, van Hooren S. Move and be moved: the effect of moving specific movement elements on the experience of happiness. Front Psychol 2020 Jan 15;11:579518 [FREE Full text] [doi: 10.3389/fpsyg.2020.579518] [Medline: 33584414]
94. Kirk A, MacMillan F, Rice M, Carmichael A. An exploratory study examining the appropriateness and potential benefit of the nintendo wii as a physical activity tool in adults aged ≥ 55 years. Interact Comput 2013;25(1):102-114. [doi: 10.1093/iwc/iws004]
95. Ethical review of research involving humans. Swedish Ethical Review Authority. URL: https://etikprovningsmyndigheten.se/for-forskare/utbildningsmaterial/
96. Zhang C, Hoel AS, Perkis A, Zadtootaghaj S. How long is long enough to induce immersion? In: Proceedings of the 10th International Conference on Quality of Multimedia Experience. 2018 Presented at: QoMEX '18; May 29-31, 2018; Cagliari, Italy p. 1-6 URL: https://ieeexplore.ieee.org/document/8463397 [doi: 10.1109/qomex.2018.8463397]
97. Kim S, Park Y, Headrick L. Daily micro-breaks and job performance: general work engagement as a cross-level moderator. J Appl Psychol 2018 Jul;103(7):772-786. [doi: 10.1037/apl0000308] [Medline: 29595289]

Abbreviations

EEQ: Exergame Enjoyment Questionnaire
MET: metabolic equivalent of task
PACES: Physical Activity Enjoyment Scale
PANAS: Positive and Negative Affect Schedule
RPE: rating of perceived exertion
RQ: research question

Edited by G Eysenbach, N Cahill; submitted 28.12.23; peer-reviewed by C Aygün, F Xu; comments to author 25.04.24; revised version received 19.06.24; accepted 15.07.24; published 26.08.24.

Please cite as:
Berglund A, Orädd H
Exploring the Psychological Effects and Physical Exertion of Using Different Movement Interactions in Casual Exergames That Promote Active Microbreaks: Quasi-Experimental Study
JMIR Serious Games 2024;12:e55905
URL: https://games.jmir.org/2024/1/e55905
doi:10.2196/55905
PMID:

Original Paper

Perceived Barriers and Facilitators Regarding the Implementation of Gamification to Promote Physical Activity in the Neighborhood: Interview Study Among Intermediaries

Ayla Schwarz[1*], PhD; Kirsten Verkooijen[1*], PhD; Emely de Vet[1,2*], PhD; Monique Simons[1*], PhD

[1]Wageningen University & Research, Wageningen, Netherlands

[2]Tilburg School of Humanities and Digital Sciences, University College Tilburg, Tilburg, Netherlands

[*]all authors contributed equally

Corresponding Author:
Ayla Schwarz, PhD
Wageningen University & Research
Hollandseweg 1
Wageningen, 6706KN
Netherlands
Phone: 31 0639187108
Email: ayla.schwarz@wur.nl

Abstract

Background: In the Netherlands, neighborhood sport coaches (NSCs) play an important role as intermediaries in promoting physical activity (PA) in the neighborhood. Gamification is the use of game elements in nongame contexts; it can be implemented with or without technology and holds promise for promoting PA. NSCs infrequently make use of this option.

Objective: This study aims to understand barriers to, and facilitators of, using gamification to promote PA, as perceived by NSCs.

Methods: A total of 25 semistructured interviews were conducted with NSCs in the Netherlands. The interviews were audiotaped, transcribed, and analyzed by means of thematic analysis using ATLAS.ti (version 22; ATLAS.ti Scientific Software Development GmbH) software. The deductive coding was informed by the capability, opportunity, motivation, behavior model and the theoretical domains framework, complemented by inductive coding.

Results: Barriers and facilitators identified as factors influencing the implementation of gamification were related to 7 themes. NSCs required technical, creative, and promotion skills; knowledge about existing gamification tools; and social support from their employer and professional network. Financial costs were identified as a barrier to the successful implementation of gamification. Lack of clarity regarding stakeholders' responsibility to implement gamification could further hamper implementation. In general, NSCs were positive about investing time in implementing gamification and expected positive effects from implementing it.

Conclusions: To overcome identified barriers, a clear overview of tools, best practices, and available subsidies must be created, a gamification network must be established, the responsibility of NSCs must be clarified, and guidance must be offered on the promotion of gamification.

(JMIR Serious Games 2024;12:e52991) doi:10.2196/52991

KEYWORDS

gamification; gamified apps; physical activity; implementation; intermediaries; interview

Introduction

Background

Physical activity (PA) supports both physical and mental health. It can prevent noncommunicable diseases, such as cardiovascular diseases, type 2 diabetes, hypertension, and certain types of cancer [1-3], and it reduces the risk of it reduces the risk of becoming overweight or developing obesity [4]. However, worldwide PA prevalence is low, especially in high-income countries [5]. PA guidelines for children recommend daily moderate to vigorous PA for 60 minutes and vigorous activities 3 days per week. PA guidelines for adults recommend at least 150 minutes of moderate-intensity aerobic

PA each week, a minimum of 75 minutes of vigorous-intensity aerobic PA or an equivalent combination of both. Muscle-strengthening activities involving major muscle groups should be performed at least 2 days a week [6]. In the Netherlands, 43% of children (aged 4-11 years), 62% of youths (aged 12-17 years), and 54% (aged 18-64 years) of adults did not meet the PA recommendations in 2022 [7]. Therefore, PA interventions are needed to increase PA levels in all populations.

Intermediaries or health brokers play an important role in offering accessible (ie, geographically close and inexpensive) PA interventions that are tailored to local communities' needs and wishes [8,9]. In the Netherlands, the Dutch Ministry of Health, Welfare, and Sport introduced neighborhood sport coaches (NSCs; in Dutch *buurtsportcoach*) with the aim of connecting the sport sector with other sectors (eg, daycare and welfare) and of promoting PA and sports participation in the neighborhood. NSCs are responsible for supporting local sports clubs, promoting a specific sport (eg, hockey or soccer), or promoting informal PA (eg, physical exertion such as running, jumping, and dancing). These coaches are 40% funded by the state and 60% funded by the municipality or other local organizations. Currently, NSCs are strategically part of the local policy of almost all municipalities in the Netherlands [10]. In each municipality, the NSC's role is unique, thereby allowing municipalities to implement activities in line with local needs and contexts. Examples of activities organized by NSCs include sports clinics, PAs during or after school, and sports tournaments. The latest trend, gamification, is increasingly gaining popularity and interest among NSCs [11,12] and is increasingly applied in PA interventions [13].

Gamification is the use of game elements in nongame contexts [14]. It draws from various theories, with the self-determination theory being the most frequently applied [15], particularly in PA interventions [13]. Self-determination theory supports the proposition that games and gamification can intrinsically motivate individuals to initiate and sustain health behaviors [16] by leveraging people's natural desires for autonomy, competence, and relatedness [17]. Gamification is a set of motivational techniques that can enhance long-term engagement with apps [18], and it offers a promising route for behavioral change [17,19]. It can be implemented either technologically or nontechnologically. Examples of nontechnological gamification include points, badges, challenges, or leaderboards. Technological implementations include, for example, urban dance grounds, digital game walls or playgrounds, or gamified apps. These gamified apps apply gamified elements to enhance user engagement and enjoyment, as seen in popular apps such as Strava (Strava, Inc), Geocaching (Groundspeak, Inc), or Pokémon Go (Niantic, Inc). Gamification is most often applied in PA interventions, and systematic reviews and meta-analyses indicate that it has small to medium effects on various PA outcomes [20-22]. Increased use of gamification elements was associated with greater effectiveness in promoting moderate to vigorous PA [18]. These gamified PA interventions have been shown to affect PA behavior in a broad range of populations, regardless of age or health status [13,18,20], and to be more effective compared with equivalent nongamified interventions [13,18]. Johnson et al [20] identified 7 promises of gamification: enhancing intrinsic motivation, broad accessibility, broad appeal, broad applicability, cost-benefit efficiency, everyday-life fit, and supporting well-being. These benefits of gamification [13,20] align well with NSCs' function of aligning their activities with the local population's needs, ensuring the everyday-life fit of their activities, engaging a wider population, and supporting well-being. Therefore, implementing gamification through NSCs may hold promise.

Despite these benefits and the availability of gamification tools such as interactive (outdoor) play equipment (eg, interactive game walls) or gamified apps (eg, Seppo [Seppo.io]), NSCs' actual use of gamification appears to remain low (Schwarz, AF, unpublished data, December 2022). Insight into factors underlying of implementation of gamification is needed to determine whether gamification aligns with NSCs' intentions and, if so, to better understand how to encourage NSCs to implement gamification. Offering gamification via intermediaries (eg, health professionals, NSCs, and health brokers) has been barely researched [23]. Instead, gamification research has so far focused largely on the development of gamified interventions from end users' perspectives and on the interventions' effectiveness on behavior change in end users [24]. Consequently, knowledge on the strategic implementation of gamified interventions at a systemic level is still lacking, despite this being invaluable for a successful and sustainable uptake of innovations [24].

This Study

Implementation science, which comprises the study of methods to promote the translation and uptake of evidence-based research into practice, has grown as a research field in the last decades. The aim of implementation science is 2-fold. On one hand, it aims to identify barriers to, and facilitators of, uptake. On the other hand, it aims to develop and apply implementation strategies to overcome identified barriers and enhance facilitators in the adoption of innovations [25]. Various implementation frameworks, such as the capability, opportunity, motivation, behavior (COM-B) model and the theoretical domains framework (TDF), have been developed to identify and understand barriers and facilitators [26,27]. Both frameworks can be classified as determinant frameworks that aim to understand influences, including barriers to, and facilitators of, implementation outcomes. This study applied both COM-B and TDF to better understand the perceived barriers to, and facilitators of, intermediaries' gamification implementation, a topic that is barely researched.

Methods

Overview

This research is part of a larger project called mVital@2040 [28], which investigates gamification to promote PA in the neighborhood among youth (aged 10-14 years) living in low-income neighborhoods. In the first step, the project aimed to investigate how NSCs can implement gamification. For this study, we conducted semistructured interviews with NSCs in 5 municipalities in the Netherlands. These 5 municipalities, ranging from small (approximately 55,000 inhabitants) to large

(>250,000 inhabitants), had 56,000; 238,000; 244,000; 368,000; and 918,000 inhabitants, respectively.

Ethical Considerations

All NSCs provided informed consent with the option to opt out, and the study was approved by the Social Sciences Ethics Committee of Wageningen University & Research on March 23, 2022 (2022-45-Schwarz). NSCs did not receive any compensation for their participation, and all personally identifiable information was pseudonymized. Informed consent was obtained by accepting the Microsoft Teams invitation. At the start of each interview, the researcher once again asked for verbal informed consent.

Selection and Study Population

For the purpose of this study, 6 intermediary organizations (eg, youth club and sports club) that employed NSCs were contacted across 5 municipalities in the Netherlands. At each organization, 1 contact person made a preselection of available NSCs, including NSCs both with and without prior experience with gamification or gamified apps to research experienced and expected barriers to, and facilitators of, gamification implementation. All age ranges of NSCs' target group, ranging from young children to older persons, were included. Preselected candidates were invited by email to take part in an interview. A total of 31 NSCs were contacted, and 25 (81%) interviews were held.

Procedure

The interviews were held between April and September 2022. The interviews took place via Microsoft Teams and lasted approximately 1 hour. All interviews were conducted by 1 researcher (AS).

The interview guide can be found in Multimedia Appendix 1. Topics discussed were related to NSCs' function profile, the target group that they aim to reach, their definition of gamification, the role of gamification in their work, inhibiting and promoting factors regarding gamification implementation, and the envisioned future of gamification. The interview guide was informed by the COM-B model [29]. The model proposes that behavior arises from a result of the interaction among 3 components: capability (ie, people must have the knowledge, skills, and strength to perform the behavior), opportunity (ie, the physical and social environment must enable the behavior to occur, by, eg, being accessible, affordable, or acceptable), and motivation (ie, people must be motivated to perform the relevant time). Although best known for its application to understand behavior change in the receivers of health interventions, COM-B has been deemed useful to focus on professionals that incorporate health interventions targeted at a variety of audiences [30-32]. Designing the interview guide based on the COM-B model allowed us to create open questions to identify perceived barriers and facilitators.

Data Analysis

The interviews were audiotaped and transcribed verbatim by an external company. The researcher who conducted the interviews checked the transcripts for accuracy. The transcripts were coded and analyzed using ATLAS.ti (version 22; ATLAS.ti Scientific Software Development GmbH) software. Data were inductively and deductively coded [33] by 1 researcher (AS) and double-coded by a second researcher (KV). Researchers compared the deductive part of the coding until a consensus was reached. In total, 3 (12%) interviews were double-coded.

Deductive coding was based on the COM-B [29] model and TDF [34] (Table 1). Both models allow the identification of barriers and facilitators influencing the uptake and use of gamification. COM-B maps all barriers and facilitators in a simple model of behavior, and TDF complements COM-B with a detailed outline of psychological capability and reflective motivation [34]. TDF contains 14 domains, derived from 33 theories and 128 psychological constructs. These can be mapped onto the components of the COM-B model [29]. The data analysis was based on thematic analysis [35], shifting from codes (TDF) to themes, which are patterns deriving from the large sample of inductive and deductive codes. In the Results section, nonspecific terms of semiquantification (eg, few, several, some, or many) are used to convey general patterns within the data [36].

Table 1. COM-B[a] and TDF[b] as outlined in the study by Michie et al [29].

COM-B component and domains (TDF)	Definition
Physical capability	
Physical skills	An ability or proficiency acquired through practice
Psychological capability	
Knowledge	An awareness of the existence of something
Cognitive and interpersonal skills	An ability or proficiency acquired through practice
Memory, attention, and decision processes	The ability to retain information, focus selectively on aspects of the environment, and choose between ≥2 alternatives
Behavioral regulation	Anything aimed at managing or changing objectively observed or measured actions
Physical opportunity	
Environmental context and resources	Any circumstance in a person's situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behavior
Social opportunity	
Social influences	Those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviors
Automatic motivation	
Emotions	A complex reaction pattern, involving experiential, behavioral, and physiological elements, by which the individual attempts to deal with a personally significant matter or event
Reinforcement	Increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus
Reflective motivation	
Social and professional role and identity	A coherent set of behaviors and displayed personal qualities of an individual in a social or work setting
Beliefs about capabilities	Acceptance of the truth, reality, or validity of an ability, talent, or facility that a person can put to constructive use
Optimism	The confidence that things will happen for the best or that desired goals will be attained
Intentions	A conscious decision to perform a behavior or a resolve to act in a certain way
Goals	Mental representations of outcomes or end states that an individual wants to achieve
Beliefs about consequences	Acceptance of the truth, reality, or validity of outcomes of a behavior in a given situation

[a]COM-B: capability, opportunity, motivation, behavior.

[b]TDF: theoretical domains framework.

Results

NSCs' Work Profile

We interviewed 25 NSCs from 5 municipalities in the Netherlands. The spread of participants (4, 8, 8, 4, 1) differed depending on the size of the municipality. Of these 25 NSCs, 18 (72%) indicated that they had worked <5 years at their organization, representing the level of work experience. NSC teams typically consisted of 1 coordinating NSC that supervised multiple executive NSCs. The NSCs had predominately an executive role (17/25, 68%), including tasks such as organizing sports activities and events in the neighborhood (eg, at schools or public parks) and recruiting members of the target group for activities and events. The other 8 (32%) NSCs had a coordination role, consisting of project planning, coordinating, and connecting different organizations and parties in the neighborhood, and supporting executive NSCs. Responsible roles and tasks differed between municipalities, and not all municipalities differentiated between executive and coordinating NSCs.

NSCs' Target Groups

Many NSCs focused on a predefined age group (eg, children aged 6-12 years, youth, children and adults, and older persons), which was determined by the responsible municipality or organization. Some, however, indicated that they were responsible for all age groups. NSCs affirmed that they had perceived a shift over the last few years from a focus on younger children (eg, aged 6-12 years) to adolescents (aged 12-16 years). Besides specific age groups, NSCs targeted people who were physically inactive or people in a low socioeconomic position. NSCs focused on hard-to-reach groups, such as adolescents (especially female adolescents), children who were physically inactive, people with a low socioeconomic position or multiple problems or constraints, people with a language barrier, young adults with depression, older people, and young adult refugees.

Level of Experience With Gamification

Quotes relating to the questions posed to NSCs are provided in Textbox 1. Overall, 20 (80%) of the 25 NSCs had some experience with gamification in general, and gamification was integrated into their program more naturally (quote 1, Textbox 1). In general, NSCs wanted to learn more about the possibility of applying gamified apps and were positive about committing themselves to adopting gamification but were still looking to see how they could apply it in practice (quote 2, Textbox 1). Examples of current applications of gamification were (1) translating web-based games (eg, Fortnite [Epic Games, Inc]) to real-life activities; (2) adding points systems, competition, and prize-winning to their current activities (eg, photo challenges and social media leaderboards); (3) implementing smart devices, such as light poles (eg, Interactive Play System cones changing colors) or smart walls; (4) reinventing a common sport (eg, James Bond game); and (5) gamified bicycle routes. Many NSCs reported a lack of experience with gamified apps. Those who had some experience implemented gamified apps sporadically. Examples of gamified apps were (1) popular social and dance-related media apps (eg, TikTok [ByteDance] and Just Dance [Ubisoft Inc]), (2) apps particularly designed to increase PA (eg, Missie Master [8D Games BV], Fitcoins [It's My Life–Corporate & Individual well-being BV], Strava [Strava Inc], City Legends [CityLegends BV], Jachtseizoen [StukTV], and VR game), or (3) apps designed by the employer to link a reward system to their current PA offer.

Textbox 1. Numbered quotes, listed per theme level of experience with gamification, defining gamification, skills, and knowledge.

Level of experience with gamification

- Quote 1: "I think [gamification] comes back, in a light form in a lot of activities in our work as [NSCs], making an activity playful." [Executive neighborhood sport coach; NSC]
- Quote 2: "I would like to broaden my er, my world. Ehm, with regard to app use.... I think there is much more to be achieved than we as an organization are currently doing." [Executive NSC]

Defining gamification

- Quote 3: "You have certain game elements...that you have in games, such as competition, leaderboards...and translate that into our own sports-promotion activities. That can be done in many ways...by means of apps, but that can also be done by being critical: well, you have sports equipment...a traditional football and a volleyball, but now there are also a lot of exercise products that are completely separate from a certain sport with the main aim of being fun.... For me, it is actually a very broad term, but to summarize it is about translating game elements elsewhere to that traditional sports world." [Executive NSC]

Skills

- Quote 4: "I'm not constantly thinking about [designing new content for the app]. So that's quite a shame, eh, that we don't all have that [creative skill] from ourselves, but at the same time I don't think you can eh, force something like that. You have to have a little, er, [technical] feeling and I just have that a little less than my colleagues." [Executive NSC]
- Quote 5: "One says, you have to start small.... One of our [executive NSCs] used to promise a lot and was very enthusiastic and creative, but, in the end, they do not manage in 'how are we going to apply that in a good way?' And, in the end, they get frustrated because it doesn't work, so that's from a practice point-of-view." [Coordinating NSC]
- Quote 6: "This is [an effective] tool, at least if promotion eh, support from the municipalities and other parties, when this is assured, then this could be the tool." [Executive NSC]
- Quote 7: "It should actually be ensured that it will just become something normal. And it shouldn't be that they do it because it's new and then they um, again um, ignore it because um, the novelty wears off. So, it has to – they have to be stimulated every time, uh, to use the app, I think." [Executive NSC]

Knowledge

- Quote 8: "In our work it is difficult to determine first, what exactly do we want, where, in which direction do we want to go, do we have enough evidence for this, do we have to find out whether this is already happening elsewhere, in other municipalities or cities, what are their experiences with this and then we already notice that we have lost quite a bit of time just figuring that out before anything really gets off the ground." [Coordinating NSC]

Defining Gamification

The definition of gamification varied among NSCs, and, at the onset of the interview, not all NSCs (8/25, 32%) were acquainted with the term. Many NSCs stated that they found it difficult to define gamification. They defined gamification mostly in terms of applying game elements, either on the web or offline, with the overall aim of promoting behavioral change. Others defined it in terms of the direct translation of web-based games to offline activities (eg, playing Super Mario [Nintendo] in real life), the application of different game elements (eg, competition or leaderboards) to offline traditional sport activities, or playful learning (quote 3, Textbox 1). Some NSCs defined gamification as the trend of increasingly making use of digital games or being increasingly more on the web and stated that they were especially confused by multiple terms, such as gamification, gaming, and esports used interchangeably.

Barriers to, and Facilitators of, Implementing Gamification

Overview

NSCs made a distinction between implementing gamification during their program (eg, integrating gamification on the site) or offering it as an extension to their program (eg, offering a gamified app during holidays, when NSCs are not physically present). In addition, 7 topics relating to barriers to, and facilitators of, gamification implementation were identified, namely, skills, knowledge, expected effects, responsibility, social support, financial costs, and time. Barriers and facilitators were discussed equally often, except for expected effects and responsibility, which were raised more often.

Skills

NSCs with gamification experiences felt very capable of applying this to their program compared with NSCs with no former experience. They felt more capable of applying gamification tools, which they often already applied unconsciously and did not require technical skills. Creative skills also appeared necessary. Some gamified apps offered the possibility to ideate and create new content, for example, challenges. Although this feature supported NSCs to use apps, it sometimes also impeded structural implementation. NSCs noticed that they continuously needed to create new content to keep the app up-to-date. Thinking continuously about new content was considered tiresome and could, together with a lack of creativity to ideate new content, hinder the implementation of a gamified app (quote 4, Textbox 1).

The development of their creative skills was influenced by starting with small steps and their former education (quote 5, Textbox 1). Several requirements need to be met to use the potential of gamification fully, such as a successful promotion of gamified apps (quote 6, Textbox 1). NSCs pointed out that they lacked skills in promoting gamification tools among the target group. In some cases, their employer took over the promotion (eg, public relations department), yet, neither succeeded in attracting (app) users (long term), which eventually resulted in the decision to stop offering a gamified app. Moreover, NSCs confirmed that it was a challenge to keep the target group engaged, especially when the novelty effect wore off (quote 7, Textbox 1).

Knowledge

NSCs indicated that their knowledge of gamification was limited, as keeping track of all trends demanded a lot of time. An overload of apps in different app stores hampered the search for suitable gamified apps. A clear overview of gamification tools (eg, examples from other municipalities), in addition to workshops, would help NSCs acquire new knowledge. Proven effects and shared experiences contributed to the decision to implement gamification (quote 8, Textbox 1).

Expected Effects

NSCs generally believed in the positive effects of gamification. Contributory factors included their own experiences, scientific evidence on the effectiveness of gamification, and the target group experiencing it as fun. However, not having an affinity for gamification, irrespective of experience, hampered implementation (quote 9, Textbox 2). NSCs stated that gamification could contribute to lowering the threshold to engage in their program, to reaching a broad target group that is often perceived as difficult to reach (eg, people who were inactive), and to reaching people from a distance (eg, during holidays). Further, gamification was considered easy to organize, and gamified apps were appreciated in terms of having one place where all information could be bundled, and valuable insights could be found based on user data. However, NSCs also noticed risks with regard to gamified apps. They feared that the target group's (especially children's) total screen time would increase and that the target group would be distracted by other apps or notifications on their phone while using the intended gamified app, and therefore NSCs feared that they would contribute to the digital divide (eg, by excluding people who have no phone or speak another language).

Textbox 2. Numbered quotes, listed per theme expected effects and responsibility.

Expected effects

- Quote 9: "I don't have anything with that at all, not myself, not personally. So um, I'm not into eh, iPads or computers. I don't have anything with that at all. So – I don't like it anyway." [Coordinating neighborhood sport coach; NSC]

Responsibility

- Quote 10: "I also think that as part of the role as NSC we should simply keep up with the times and this is no longer just the standard games as before. But we, uh, have to go along with it." [Executive NSC]
- Quote 11: "I like to move with the times...then I'd rather be too early with something than...too late." [Executive NSC]
- Quote 12: "Um, the question is, whose task is it? Maybe of the municipality? or education? Ehm, and I think that er, that issue has never actually been mentioned and as long as, as long as no one actually raises it, it will actually not be taken up, eh, I think." [Coordinating NSC]
- Quote 13: "To share the knowledge, that would be my task.... So, I should actually acquire the knowledge first and then share it among our [executive NSCs]. So, it should be the responsibility of the [coordinating NSCs] to train." [Coordinating NSC]
- Quote 14: "The [executive NSCs] we employ are responsible for one sport. And they don't yet have enough knowledge to continue with the bigger piece. And they are used to just providing fun sports lessons and being enthusiastic about their sport. But [gamification] requires something completely different – that has a completely different goal than if you try to get the kids excited to start playing [one particular] sports and let them discover what they like .... So, that's quite a big shift for them. So, it requires a different way of looking at how you organize your lessons and a different way of looking at the children you have in your group and that takes quite a lot." [Coordinating NSC]
- Quote 15: "Sometimes you just need to be handed the [gamification] solution to eventually do it." [Executive NSC]
- Quote 16: "Eh, if this is an assignment from the municipality, eh, then I feel that we should just, eh, take it up together." [Executive NSC]
- Quote 17: "Who takes care of it and is also going to implement it and, er, everything that comes with it. So, the implementation of the lessons, the implementation of new elements eh, having knowledge about it, but also the communication of, okay, how are we going to promote it?" [Coordinating NSC]
- Quote 18: "It is good that a number of colleagues are in the lead and are really going to pick that up and roll it out and then maybe give me a concrete assignment of, you now have to come up with a mission and then I will think about it. But I don't quickly start from myself." [Executive NSC]

Responsibility

NSCs viewed the adoption of gamification in their professional responsibilities as a means of keeping up with the latest developments and trends (quote 10 and Q11, Textbox 2). At the same time, NSCs questioned whose responsibility it was to adopt gamification eventually (quote 12, Textbox 2). One coordinating NSC considered knowledge transfer as part of their responsibility (quote 13, Textbox 2). Another coordinating NSC indicated that imposing the responsibility on executive NSCs could be challenging (quote 14, Textbox 2). Appointing a responsible NSC in the team and receiving a direct assignment by the supervisor or policy makers were deemed to promote implementation (quotes 15 and 16, Textbox 2). Yet, many NSCs indicated that it was difficult to find suitable, but indispensable, persons to enthusiastically promote gamification in the field (quote 17, Textbox 2). Reinforcement by colleagues or supervisors was considered essential to keep reminding them to ideate and create new content or updates for a gamified app. A clear assignment by supervisors can support NSCs. However, once they noticed that the gamified app was not or barely used, they felt demotivated to create new content or to promote the gamified app. NSCs clearly designated reinforcement as a responsibility of coordinating NSCs (quote 18, Textbox 2).

Social Support

NSCs received both direct support from their employer and support from external parties, such as schools, municipalities, or commercial sports clubs.

Support by Employer

Overall, NSCs emphasized that they experienced a lot of freedom in performing their job. Freedom was allowed, as long as activities aligned with national or local sports agreements that often reflected the needs of the neighborhood for which they worked (quote 19, Textbox 3). NSCs felt supported by their employer to apply gamification (quote 20, Textbox 3). However, several NSCs asserted that they did not receive active support to apply gamification. They often used social media to get inspiration on how to apply gamification. Sharing this content with their direct colleagues contributed to the implementation of gamification tools.

Textbox 3. Numbered quotes, listed per theme social support, external support, financial costs, and time.

Social support

- Quote 19: “I do have a lot of freedom and that stems, I think, from the vision of, er, our employer, that, er, every neighborhood actually needs customization. So, what works in my neighborhood may not work in another neighborhood for example. So, it is also necessary to have some freedom, to be able to test some things and then also build on that.” [Executive neighborhood sport coach; NSC]
- Quote 20: “I just notice, we are with quite a large team of [executive NSCs] and everyone has their own expertise and their own eh, eh, eh, thing they get energy from. And I do notice that, as far as possible, efforts are made to give people tasks that suit them best.” [Executive NSC]

External support

- Quote 21: “Financial support from the municipality is necessary...what you often see in municipal processes is that they are quite slow and start slowly.” [Executive NSC]
- Quote 22: “If the municipality would focus on that and if the municipality had the vision to do more with it, um, financial support would be desirable.” [Executive NSC]
- Quote 23: “Certain choices that are made higher up and which you as an [NSC] cannot influence.” [Executive NSC]
- Quote 24: “The municipality actually thought the results were too, too small. Eh, while the investment, uh, was quite high. So eh, quite a lot of time and energy eh, went into that, but actually too little [impact], in a structural way.” [Coordinating NSC]
- Quote 25: “I am getting more and more budget....I looked for more eh, partners who eh, who have similar initiatives in the city. So uh, I looked at what we, what we can uh, do together. Eh, little by little, the more the municipality and other parties experience it, the more enthusiastic they become, the more eh, the more budget is available and the more they are convinced that it can really be an added value eh, for the city.” [Coordinating NSC]

Financial costs

- Quote 26: “Money is always an uh, an obstacle, let me put it this way. So, you do indeed have to have good considerations and good arguments that it has an eh, that it has added value and there is always – you always have to argue that very well and eh, eh, have that information before the choice would indeed be taken.” [Coordinating NSC]

Time

- Quote 27: “To shape it into an uh, a concept idea that is realistic, feasible, and that really, uh, can take place, also needs time.” [Coordinating NSC]
- Quote 28: “That will take a lot of time. Ehm, so if that threshold is lowered, it will become easier to apply it in practice.” [Executive NSC]
- Quote 29: “Of course, you have to make time to find out how it works. How are you going to apply it? But on the other hand, it also eases your work sometimes, as the app takes over a lot.” [Executive NSC]
- Quote 30: “The goal was to, uh, actually get people moving. Eh, without having to be physically present and that was eh, ideally eh, a very nice goal for us. Only in practice it turned out differently. Eh, mainly because it just takes a lot of time to prepare. Ehm, but once it’s in practice, it’s done.” [Executive NSC]
- Quote 31: “I think it’s not so much a question of do you have time, but more of, do you make time?” [Executive NSC]

External Support

Support from outside the organization was rarely mentioned, except for NSCs who indicated that support from external parties (eg, schools and commercial sports clubs) contributed to a more sustained implementation of gamification. In general, a strong collaboration and exchange between various stakeholders was considered important. NSCs perceived pressure to adopt and implement gamified apps when they noticed that colleagues outside their organization kept up with gamified app trends. NSCs lacked highly needed support and vision from policy makers (quotes 21 and 22, Textbox 3). Furthermore, NSCs mentioned policy makers’ power, which they could not influence (quotes 23 and 24, Textbox 3). One NSC mentioned the importance of starting small to slowly convince the municipality and other organizations and to build a sustainable gamification network (quote 25, Textbox 3). NSCs stated that it was part of their role to create sustained networks in the neighborhood. However, making external organizations enthusiastic about gamification, or providing more information about the gamification concept or the added value of gamified apps, were perceived as barriers.

Financial Costs

Challenges with regard to costs included (1) high product costs (eg, expensive physical gamification tools or subscriptions to gamified apps), (2) difficulties finding suitable subsidies to afford gamification tools or gamified apps, and (3) substantiating the budget plan to policy makers with persuasive arguments. Regarding high product costs, it was mentioned that it was financially challenging for smaller municipalities to subscribe to a gamified app compared to larger municipalities, as often the same price was applied and was not linked to the number of inhabitants and therefore (possible) users. In addition to subscription costs, expensive on-site products (eg, smart playgrounds) that required subsidies from municipalities were mentioned as a barrier. Furthermore, some NSCs experienced difficulty finding suitable subsidies or finding different municipal subsidies to combine, whereas others seemed to know all the ins and outs of finding the right subsidies. Justification or argumentation was important to receive the subsidy eventually. An adequate level of knowledge about gamification tools and gamified apps (eg, including scientific effects) and expected outcomes was required to justify their subsidy plans (quote 26, Textbox 3).

Time

For small gamification adaptations, time was not identified as an impeding factor, as this felt more intuitive for NSCs to align with their current way of working and their perceived freedom to plan their activities. However, especially preparatory work, such as justifying costs, finding scientific evidence or suitable applications, and good practices was identified as challenging (quotes 27 and 28, Textbox 3). NSCs indicated that the time investment could be worth it (quotes 29 and 30, Textbox 3). NSCs framed it as the willingness to prioritize time, which was seen as being crucial to the sustainable provision of an up-to-date app (quote 31, Textbox 3).

Discussion

Principal Findings

This study aimed to better understand the perceived barriers to, and facilitators of, NSCs' use of gamification. The results reveal that NSCs have limited experience with gamification; this is in line with earlier research showing that only 4% of lifestyle professionals in the Netherlands use gamified tools [37]. In comparison with their use of gamified apps, NSCs apply gamification (without technology) more naturally in their daily work. Perceived barriers and facilitators related to skills, knowledge, expected effects, responsibility, social support, financial costs, and time. No other research has been done on NSCs' implementation of gamification, but we can compare our results with a study focusing on the implementation of digital health tools by lifestyle professionals in the Netherlands [37]. That study found that lifestyle professionals were motivated to use digital coaching tools but required more content-specific exchanges with colleagues and a complete overview of available tools. They lacked training during their education and experienced a high financial burden when implementing digital coaching tools. All these factors align with the results of this study in terms of expected effects, knowledge, social support, skills, and financial costs [37]. Regarding responsibility and time, our findings in this study are different. Lifestyle professionals perceived the successful implementation of digital coaching as time-intensive, and hours spent could not be claimed from health insurance. In comparison, the NSCs in this study did not consider time to be a barrier. Rather, their perceived freedom in their work contributed to implementing gamification. However, both executive and coordinating NSCs felt uncertain about which of the 2 should take responsibility for implementing gamification, whereas this did not appear as a barrier for lifestyle professionals [37].

Former research among intermediaries, such as physiotherapists, physicians, and nurse practitioners, investigated the implementation of health interventions in primary care. Structurally implementing existing health guidelines (ie, PA guidelines), exercise referral schemes, or interventions (ie, mobile health, eHealth, and combined lifestyle interventions) in organizational structures and regulations were identified as a common barrier to successful implementation [30,37-44]. Besides NSCs, school teachers are important intermediaries and implementers of gamification to promote PA. They already have experience with implementing gamification with the aim of promoting PA in schools [45,46]. Schools are often considered as a place where a diverse population, including students that are often considered hard to reach, can be addressed in a structural manner (eg, integration in the curriculum). This is the first study to focus on the implementation of gamification by NSCs as intermediaries. Offering interventions via intermediaries may contribute to the successful implementation of interventions. Intermediaries can reach populations that may benefit the most from the interventions and are often considered hard to reach. Our research reveals that NSCs perceive advantages in implementing gamification (ie, expected effects) for hard-to-reach populations that are often known to engage in lower levels of PA, to reach people from a distance, and to lower the threshold to engage in their program. Expected effects emerged as one of the most frequently mentioned themes in our study. Furthermore, NSCs perceive that they have the freedom to choose suitable activities themselves and are in charge of allocating their time to a range of diverse activities. This suggests that NSCs form a specific group of intermediaries that hold promise to integrate gamification in their current work.

Coordinating and executive NSCs' roles and responsibilities regarding the adoption and implementation of gamification appear to be unclear; this emerged as a frequently mentioned theme. This is in line with earlier research indicating that a blueprint for NSCs' function is not in place [47] and at the same time not desired, as the freedom to implement PA activities and allocate their time independently is part of the function description [48]. In general, both coordinating and executive NSCs see gamification implementation as part of their function but experience difficulties in successfully establishing collaborations with other parties to structurally implement gamification. Moreover, a clear vision on the part of municipal policy makers could facilitate the structural implementation of gamification. As a result, strategic implementation at a systemic level is often lacking [24]. A responsible person who is enthusiastically involved and reinforces colleagues often seems to be lacking, although such a person is considered crucial for the strategic implementation of gamification in the long term to prevent NSCs from implementing gamification only sporadically. Clear assignments by the manager or municipal policy makers may support successful implementation in the future [43,44,49].

Research has shown that scaling up digital health innovations is often hindered by barriers, such as a lack of evidence or long-term financial funding, and therefore often depends on the acceptance of different individual stakeholders working in different domains or different levels of systems. Research suggests that a higher level of coordination and knowledge sharing across digital health projects is needed [50]. A promising pathway to facilitate scaling up is to incorporate systems thinking in the implementation process. Systems thinking helps to elucidate how situations and experiences are connected and influence implementations. Viewing gamification through the conceptual lens of systems thinking is deemed to have the potential to create harmonized solutions that are relevant to individuals and take the various mutually influencing levels of the system into account [51]. Systems thinking is in line with our study, as different levels in the system, such as research, policy, and practice, do not seem to align well during the implementation phase. For example, policy needs to be informed by researchers to support the scientific effects of gamification (eg, on health outcomes). In contrast, practice (eg, NSCs) lacks subsidization and policy makers' vision. A better alignment among the 3 different levels of the system together with end users is crucial [52] to ensure adoption, implementation, and sustainability and therefore effects on health outcomes (ie, increased PA). Implementation science can facilitate this need to translate research into practice and contribute to system change. Including different stakeholders from different system levels through the development of the gamification tool is crucial. Integrating a horizontal systems approach (ie, sports clubs, schools and PA organizations) with a vertical systems approach (ie, local-, regional-, national-level governments) may strengthen synergetic effects among a varied group of stakeholders and further strengthen sustainable implementation and even scale up of gamification and gamified apps [53,54].

Strengths and Limitations

Several limitations of this research should be considered. First, no generalizations can be made beyond the study sample. NSCs seemed to focus on children and adolescents during the interviews, although this age group was not the sole focus of this study. Possibly, NSCs relate the concept of gamification and gaming to a younger age population. As we recruited via a contact person in each organization, it might also be possible that the contact person made a preselection beforehand. According to the national NSC monitor, many NSCs do focus on children or adolescents, therefore our sample seems to be in line with the general NSC profile [55]. In addition, as the NSCs' job description varies between municipalities, it is difficult to determine whether our sample reflects the diversity of NSC profiles currently present across municipalities. In the Netherlands, municipalities can employ executive or coordinating NSCs, or NSCs that fulfill both functions. However, some municipalities employ only coordinating NSCs or only executive NSCs, who may either lack vision, usually developed by coordinating NSCs, or lack executive activities performed by executive NSCs [48]. In our study, the interviewed NSCs in the 5 municipalities are all employed as either coordinating or executive NSCs.

Several strengths of this study can be identified too. We conducted a substantial number of qualitative interviews that led to data saturation. In addition, in terms of method, we managed to inform our analysis based on well-established theoretical models, the COM-B and TDF, that cover the wide spectrum of behavioral determinants. Coding the interviews according to the COM-B model and TDF shows that both are relevant in implementation science and can identify factors facilitating or hindering the implementation of innovation [30,38-41,56]. Moreover, the analysis was performed by 2 coders, thereby further strengthening the credibility of the study. In terms of gamification experience, we reached a diverse sample including nonexperienced, inexperienced, and experienced gamification NSCs. As this was the first study to focus on intermediaries promoting PA in the neighborhood via gamification in their daily work, we managed to contribute to a novel and relevant scientific and societal topic.

Implications

For Practice

By linking the identified facilitators and barriers to the behavior change wheel [29], we can identify several relevant intervention functions, such as education (eg, providing information), training (eg, practicing technical and creative skills in workshops), modeling (eg, providing best practices and clear implementation guidelines), and environmental restructuring (eg, support a restructuring of the organization where NSCs are employed). From the results, we formulated 5 recommendations for policy makers, industry, intermediaries, and professionals (Textbox 4).

First, NSCs indicated that they lacked not only sufficient skills training to implement gamification but also education about the scientific effects and an overview of the different tools. Hence, it is recommended to provide a clear overview of tools and best practices, for example, via education and workshops. Second, NSCs felt supported by their employer who granted them a lot of freedom to design activities in line with local needs and wishes. However, NSCs indicated a lack of exchange with colleagues and external parties. A network within but also outside the organization or municipality appears to be important in terms of sharing content, best practices, and keeping up with the latest trends and developments. This relates to intersectoral collaborations [37,43,44,47,57], which might be needed among NSCs in the Netherlands to learn how to better collaborate with other parties. Here, NSCs' professional organizations could play an important role in facilitating a network [37]. Third, coordinating and executive NSCs were uncertain about their responsibility to adopt gamification, a frequently mentioned theme. Clarifying NSCs' responsibility, as indicated by NSCs during the interviews, for example, by appointing a dedicated gamification NSC to provide a clear assignment to executive NSCs or to allocate hours to NSCs, may help the sustainable implementation of gamification. Fourth, NSCs perceive large barriers with regard to financial costs, as product costs are high and suitable subsidies are lacking. A clear overview of subsidy options may help to better connect and scale up initiatives and integrate gamification and gamified apps sustainably. Fifth, NSCs indicated that they lacked sufficient promotion skills that they considered necessary for the successful implementation of gamification and gamified apps. Supporting NSCs to successfully embed the gamification tool in a larger promotion campaign might impact the implementation of an intervention.

Textbox 4. Actions that should be taken to enable all relevant stakeholders, policy makers, industry, intermediaries, and professionals to exploit the potential of gamification.

Recommendations

- Establish a sustainable *network* of organizations to strengthen the offer within one municipality and to exchange ideas and best practices beyond municipalities
- Clarify the *responsibility* of neighborhood sport coaches (NSCs), for example, appoint a dedicated NSC to provide a clear assignment or to allocate hours
- Provide a clear *subsidy* overview that helps NSCs to connect initiatives and integrate gamification and gamified apps sustainably
- Offer guidance on how to *promote* gamification and gamified apps among the target group to secure book long-lasting engagement

For Research

One of the barriers identified by intermediaries was the lack of a clear overview of gamification that has proven effective. Research is often struggling with conclusive results on the effects of gamification and often lacks rigorous study designs [13,22,58]. A white paper that translates research results in a comprehensive manner could support intermediaries in the future to understand and refer to the proven effects, for example, to justify financial subsidies. Furthermore, more quantitative research is necessary to identify the determinants that significantly explain NSCs' implementation of gamification across the Netherlands and how these determinants influence different stages of the implementation phase (ie, adoption, implementation, and sustainability) [59]. As differences between NSC profiles depend on the municipality, it is difficult to assess which resources contribute to the implementation of gamification, for example, small versus large municipalities' offering. Moreover, to understand the implementation of gamification at the systems level, it is important to better align research with practice and policy, as all 3 subsystems depend on one another. Future studies may want to consider integrating the focus on different levels of the system and how they can be better aligned to implement long-lasting gamification interventions with effects on health outcomes.

Conclusions

This study aimed to assess the barriers and facilitators that NSCs perceive toward using gamification in their daily work. NSCs highly valued the application of gamification to stimulate PA in the neighborhood. Different barriers and facilitators relating to 7 themes emerged, namely, skills, knowledge, expected effects, responsibility, social support, financial costs, and time. As this is one of the first studies to focus on NSCs that implement gamification, different practice-based and research-based implications can be considered. Providing a clear overview of tools and best practices, a clear overview of subsidies available to help NSCs connect initiatives and integrate gamification sustainably, and guidance on how to promote gamification among the target group are important actions that can be undertaken by means of written documentation. Certain restructuring policies can assist the establishment of a sustainable network of organizations to strengthen the offer within and beyond the municipality and to clarify NSCs' responsibilities. Further research is needed to measure the effects of gamification and needs to be approached by the triangle of research, practice, and police to implement gamification in a systems approach.

Acknowledgments

This work was supported by the mVital@2040 project, which is a part of the Vitality Academy and connected to the EWUU (Eindhoven University of Technology, Wageningen University & Research, Utrecht University, and University Medical Centre Utrecht) Alliance. The project and AS were funded by Wageningen University & Research. MS initiated and acquired the funding for the mVital@2040 project. The authors would like to thank the company that was responsible for writing out the transcript.

Authors' Contributions

AS was responsible for leading all stages of this study. AS and MS developed the study design. AS conducted the data collection and executed the coding. KV served as the second coder. AS analyzed the data and drafted the manuscript, which was written together with KV, MS, and EdV. The manuscript was eventually reviewed and adapted by KV, MS, and EdV. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 23 KB - games_v12i1e52991_app1.docx]

References

1. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013 Sep 08;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
2. Jeon CY, Lokken RP, Hu FB, van Dam RM. Physical activity of moderate intensity and risk of type 2 diabetes: a systematic review. Diabetes Care 2007 Mar;30(3):744-752. [doi: 10.2337/dc06-1842] [Medline: 17327354]
3. Li J, Loerbroks A, Angerer P. Physical activity and risk of cardiovascular disease: what does the new epidemiological evidence show? Curr Opin Cardiol 2013 Sep;28(5):575-583. [doi: 10.1097/HCO.0b013e328364289c] [Medline: 23928923]
4. Steinbeck KS. The importance of physical activity in the prevention of overweight and obesity in childhood: a review and an opinion. Obes Rev 2001 May 07;2(2):117-130. [doi: 10.1046/j.1467-789x.2001.00033.x] [Medline: 12119663]
5. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health 2018 Oct;6(10):e1077-e1086. [doi: 10.1016/s2214-109x(18)30357-7]
6. WHO guidelines on physical activity and sedentary behaviour: web annex: evidence profiles. World Health Organization. 2020. URL: https://iris.who.int/bitstream/handle/10665/336657/9789240015111-eng.pdf?sequence=1 [accessed 2024-02-18]
7. Beweegrichtlijnen: het aandeel van de Nederlandse bevolking van 4 jaar en ouder dat voldoet aan de beweegrichtlijnen. Rijksinstituut voor Volksgezondheid en Milieu. Rijksinstituut voor Volksgezondheid en Milieu. 2023 Mar 17. URL: https://www.sportenbewegenincijfers.nl/kernindicatoren/beweegrichtlijnen [accessed 2023-03-20]
8. Dummer E, van den Berg L. Sport en bewegen in de wijk: houd het laagdrempelig en vrijblijvend. Utrecht: Kenniscentrum Sport & Bewegen. 2021 Feb 24. URL: https://www.allesoversport.nl/thema/meedoen-door-sport-en-bewegen/sport-en-bewegen-in-de-wijk-houd-het-laagdrempelig-en-vrijblijvend/ [accessed 2021-02-24]
9. van Dijk ML, Te Loo LM, Vrijsen J, van den Akker-Scheek I, Westerveld S, Annema M, et al. LOFIT (lifestyle front office for integrating lifestyle medicine in the Treatment of patients): a novel care model towards community-based options for lifestyle change-study protocol. Trials 2023 Feb 17;24(1):114 [FREE Full text] [doi: 10.1186/s13063-022-06960-z] [Medline: 36803271]
10. van Stam W, Pulles I. De buurtsportcoach: sport, bewegen en gezondheid verbonden. TSG 2021 Jul 09;99(Suppl 1):11-14 [FREE Full text] [doi: 10.1007/s12508-021-00303-0] [Medline: 34257514]
11. Inspiratiesheet over esports, gaming en gamification. Kenniscentrum Sport & Bewegen. 2021 Sep. URL: https://www.kennisbanksportenbewegen.nl/?file=10484&m=1633592047&action=file.download [accessed 2023-03-20]
12. Dummer E, Bulkens D. Gaming en gamification toepassen in de praktijk: successleutels en tips. alles over sport. 2021 Nov 17. URL: hhttps://tinyurl.com/2fuz2c8s [accessed 2023-03-20]
13. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Jan 04;24(1):e26779 [FREE Full text] [doi: 10.2196/26779] [Medline: 34982715]
14. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland p. 9-15 URL: https://dl.acm.org/doi/10.1145/2181037.2181040 [doi: 10.1145/2181037.2181040]
15. Krath J, Schürmann L, von Korflesch HF. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comput Human Behav 2021 Dec;125:106963. [doi: 10.1016/j.chb.2021.106963]
16. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Hum Comput Interact 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471] [Medline: 26281194]
17. Sailer M, Hense JU, Mayr SK, Mandl H. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput Human Behav 2017 Apr;69:371-380. [doi: 10.1016/j.chb.2016.12.033]
18. Maher CA, Olds T, Vandelanotte C, Plotnikoff R, Edney SM, Ryan JC, et al. Gamification in a physical activity app: what gamification features are being used, by whom, and does it make a difference? Games Health J 2022 Jun 01;11(3):193-199. [doi: 10.1089/g4h.2021.0207] [Medline: 35501981]
19. Hamari J, Koivisto J, Sarsa H. Does gamification work?: a literature review of empirical studies on gamification. In: Proceedings of the 47th Hawaii International Conference on System Science. 2014 Presented at: HICSS '14; January 6-9, 2014; Waikoloa, HI p. 3025-3034 URL: https://ieeexplore.ieee.org/document/6758978 [doi: 10.1109/hicss.2014.377]
20. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
21. Yang Y, Hu H, Koenigstorfer J. Effects of gamified smartphone applications on physical activity: a systematic review and meta-analysis. Am J Prev Med 2022 Apr;62(4):602-613. [doi: 10.1016/j.amepre.2021.10.005] [Medline: 34893387]
22. Koivisto J, Hamari J. Gamification of physical activity: a systematic literature review of comparison studies. In: Proceedings of the 2019 Annual International GamiFIN Conference. 2019 Presented at: GamiFIN '19; April 8-10, 2019; Levi, Finland p. 8-10 URL: https://ceur-ws.org/Vol-2359/paper10.pdf

23. Harris MA, Crone D. Motivations and barriers to engagement with a technology-enabled community wide physical activity intervention. PLoS One 2020 Jun 26;15(6):e0232317 [FREE Full text] [doi: 10.1371/journal.pone.0232317] [Medline: 32589658]
24. Cheng VW. Recommendations for implementing gamification for mental health and wellbeing. Front Psychol 2020 Dec 7;11:586379 [FREE Full text] [doi: 10.3389/fpsyg.2020.586379] [Medline: 33365001]
25. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? Psychiatry Res 2020 Jan;283:112376 [FREE Full text] [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
26. Michie S. Implementation science: understanding behaviour change and maintenance. BMC Health Serv Res 2014 Jul 7;14(S2):O9. [doi: 10.1186/1472-6963-14-s2-o9]
27. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
28. Vital@2040. Vital Academy. 2024. URL: https://vitalityacademy.eu/project/vital-2040/ [accessed 2024-04-15]
29. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. London, UK: Silverback Publishing; 2014.
30. Allin L, Haighton C, Dalkin S, Das J, Allen G. Understanding the challenges and impact of training on referral of postnatal women to a community physical activity programme by health professionals: a qualitative study using the COM-B model. Midwifery 2023 Jan;116:103516. [doi: 10.1016/j.midw.2022.103516] [Medline: 36283295]
31. Smith CA, McNeill A, Kock L, Shahab L. Exploring mental health professionals' practice in relation to smoke-free policy within a mental health trust: a qualitative study using the COM-B model of behaviour. BMC Psychiatry 2019 Feb 04;19(1):54 [FREE Full text] [doi: 10.1186/s12888-019-2029-3] [Medline: 30717722]
32. Pugh G, Petrella A, Pabary A, Cross A, Hough R, Soanes L, et al. Health professionals' views on lifestyle advice delivery to teenage and young adult cancer patients: a qualitative study. Cancer Nurs 2022 Nov 30;45(1):E238-E245. [doi: 10.1097/NCC.0000000000000906] [Medline: 33252408]
33. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 08;320(7227):114-116 [FREE Full text] [doi: 10.1136/bmj.320.7227.114] [Medline: 10625273]
34. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the theoretical domains framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
35. Braun V, Clarke V. Thematic Analysis. Washington, DC: American Psychological Association; 2012.
36. Neale J, Miller P, West R. Reporting quantitative information in qualitative research: guidance for authors and reviewers. Addiction 2014 Feb;109(2):175-176. [doi: 10.1111/add.12408] [Medline: 24422609]
37. Bak M, Bossen D, Braam K, Holla J, Visser B, Dallinga J. [Experiences and support needs of lifestyle professionals in the use of digital coaching tools for clients with overweight]. TSG 2023;101(2):38-45 [FREE Full text] [doi: 10.1007/s12508-023-00379-w] [Medline: 37206640]
38. Alexander KE, Brijnath B, Mazza D. Barriers and enablers to delivery of the Healthy Kids Check: an analysis informed by the Theoretical Domains Framework and COM-B model. Implement Sci 2014 May 23;9:60 [FREE Full text] [doi: 10.1186/1748-5908-9-60] [Medline: 24886520]
39. Rosário F, Santos MI, Angus K, Pas L, Ribeiro C, Fitzgerald N. Factors influencing the implementation of screening and brief interventions for alcohol use in primary care practices: a systematic review using the COM-B system and theoretical domains framework. Implement Sci 2021 Jan 07;16(1):6 [FREE Full text] [doi: 10.1186/s13012-020-01073-0] [Medline: 33413487]
40. Clark RE, McArthur C, Papaioannou A, Cheung AM, Laprade J, Lee L, et al. "I do not have time. Is there a handout I can use?": combining physicians' needs and behavior change theory to put physical activity evidence into practice. Osteoporos Int 2017 Jun 17;28(6):1953-1963. [doi: 10.1007/s00198-017-3975-6] [Medline: 28413842]
41. Downey J, Shearn K, Brown N, Wadey R, Breckon J. Behaviour change practices in exercise referral schemes: developing realist programme theory of implementation. BMC Health Serv Res 2021 Apr 13;21(1):335 [FREE Full text] [doi: 10.1186/s12913-021-06349-9] [Medline: 33849541]
42. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of eHealth services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197 [FREE Full text] [doi: 10.2196/14197] [Medline: 31755869]
43. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
44. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
45. Fernandez-Rio J, de las Heras E, González T, Trillo V, Palomares J. Gamification and physical education. Viability and preliminary views from students and teachers. Phys Educ Sport Pedagogy 2020 Mar 19;25(5):509-524. [doi: 10.1080/17408989.2020.1743253]

46. Simons M, Opdam L, van Empelen P. Feasibility of an active game program in a Dutch pre-vocational high school setting. Games Health J 2013 Dec;2(6):332-340. [doi: 10.1089/g4h.2013.0047] [Medline: 26197074]
47. Smit E, Leenaars KE, Wagemakers A, van der Velden K, Molleman GR. Perceptions of care sport connectors' tasks for strengthening the connection between primary care, sports and physical activity: a Delphi study. Int J Integr Care 2020 Apr 01;20(1):13 [FREE Full text] [doi: 10.5334/ijic.4789] [Medline: 32292311]
48. van Stam DW. Lessen uit vijftien jaar buurtsportcoach. Kenniscentrum Sport & Bewegen. 2022. URL: https://www.kennisbanksportenbewegen.nl/?file=11010&m=1668074256&action=file.download [accessed 2024-04-20]
49. de Vries H, Kloek C, Bossen D, Veenhof C. Waarom wordt e-health (niet) gebruikt?: fysiotherapeutisch gebruik van een blended e-health-interventie. FysioPraxis. 2016. URL: https://research.hanze.nl/ws/portalfiles/portal/24470285/Fysiopraxis_3_2016_pg35_37.pdf [accessed 2024-05-01]
50. Capponi G, Corrocher N. Patterns of collaboration in mHealth: a network analysis. Technol Forecast Soc Change 2022 Feb;175:121366. [doi: 10.1016/j.techfore.2021.121366]
51. AlSaad FM, Durugbo CM. Gamification-as-innovation: a review. Int J Innov Technol Manag 2021 Aug 19;18(05):2130002. [doi: 10.1142/S0219877021300020]
52. Swinkels IC, Huygens MW, Schoenmakers TM, Oude Nijeweme-D'Hollosy W, van Velsen L, Vermeulen J, et al. Lessons learned from a living lab on the broad adoption of eHealth in primary health care. J Med Internet Res 2018 Mar 29;20(3):e83 [FREE Full text] [doi: 10.2196/jmir.9110] [Medline: 29599108]
53. Estabrooks PA, Brownson RC, Pronk NP. Dissemination and implementation science for public health professionals: an overview and call to action. Prev Chronic Dis 2018 Dec 20;15:E162 [FREE Full text] [doi: 10.5888/pcd15.180525] [Medline: 30576272]
54. Maclean LM, Clinton K, Edwards N, Garrard M, Ashley L, Hansen-Ketchum P, et al. Unpacking vertical and horizontal integration: childhood overweight/obesity programs and planning, a Canadian perspective. Implement Sci 2010 May 17;5:36 [FREE Full text] [doi: 10.1186/1748-5908-5-36] [Medline: 20478054]
55. Heijnen E, van Stam W, Berns AG, Vrieswijk S. Monitor brede regeling combinatiefuncties. Mulier Instituu. 2022. URL: https://www.kennisbanksportenbewegen.nl/?file=11012&m=1668605293&action=file.download [accessed 2024-05-01]
56. Ziebart C, McArthur C, Lee L, Papaioannou A, Laprade J, Cheung AM, et al. "Left to my own devices, I don't know": using theory and patient-reported barriers to move from physical activity recommendations to practice. Osteoporos Int 2018 May 13;29(5):1081-1091. [doi: 10.1007/s00198-018-4390-3] [Medline: 29441402]
57. von Heijden A, Kunen M, Vos S. Van alle markten thuis?: een analyse van de competenties en de ondersteuningsbehoefte van buurtsportcoaches. Fontys Sporthogeschool. 2020. URL: https://tinyurl.com/2uu9r2ze [accessed 2024-05-01]
58. Xu L, Shi H, Shen M, Ni Y, Zhang X, Pang Y, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. JMIR Mhealth Uhealth 2022 Feb 03;10(2):e27794 [FREE Full text] [doi: 10.2196/27794] [Medline: 35113034]
59. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol 2008 Jun;41(3-4):327-350. [doi: 10.1007/s10464-008-9165-0] [Medline: 18322790]

Abbreviations

COM-B: capability, opportunity, motivation, behavior
NSC: neighborhood sport coach
PA: physical activity
TDF: theoretical domains framework

Edited by T Leung, T de Azevedo Cardoso; submitted 21.09.23; peer-reviewed by B Bock, T Ferguson; comments to author 29.02.24; revised version received 29.05.24; accepted 03.07.24; published 28.08.24.

Please cite as:
Schwarz A, Verkooijen K, de Vet E, Simons M
Perceived Barriers and Facilitators Regarding the Implementation of Gamification to Promote Physical Activity in the Neighborhood: Interview Study Among Intermediaries
JMIR Serious Games 2024;12:e52991
URL: https://games.jmir.org/2024/1/e52991
doi:10.2196/52991
PMID:39196618

Original Paper

Design of Virtual Reality Exergames for Upper Limb Stroke Rehabilitation Following Iterative Design Methods: Usability Study

Julian Felipe Villada Castillo[1], MSc; Maria Fernanda Montoya Vega[2], MSc; John Edison Muñoz Cardona[3], PhD; David Lopez[4], MSc; Leonardo Quiñones[4], MSc; Oscar Alberto Henao Gallo[5], PhD; Jose Fernando Lopez[5], MD

[1]Faculty of Basic Sciences, Department of Physics, Universidad Tecnologica de Pereira, Pereira, Colombia
[2]Department of Human-Centered Computing, Exertion Games Lab, Monash University, Melbourne, Australia
[3]Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada
[4]Engineering Faculty, Universidad Tecnologica de Pereira, Pereira, Colombia
[5]Medicine Faculty, Universidad Tecnologica de Pereira, Pereira, Colombia

Corresponding Author:
Julian Felipe Villada Castillo, MSc
Faculty of Basic Sciences
Department of Physics
Universidad Tecnologica de Pereira
Carrera 27 #10-02 Álamos
Pereira, 660003
Colombia
Phone: 57 606 3137300
Fax: 57 606 3213206
Email: jfvillada@utp.edu.co

Abstract

Background: Since the early 2000s, there has been a growing interest in using exercise video games (exergames) and virtual reality (VR)–based interventions as innovative methods to enhance physical rehabilitation for individuals with multiple disabilities. Over the past decade, researchers and exercise professionals have focused on developing specialized immersive exercise video games for various populations, including those who have experienced a stroke, revealing tangible benefits for upper limb rehabilitation. However, it is necessary to develop highly engaging, personalized games that can facilitate the creation of experiences aligned with the preferences, motivations, and challenges communicated by people who have had an episode of stroke.

Objective: This study seeks to explore the customization potential of an exergame for individuals who have undergone a stroke, concurrently evaluating its usability as a technological tool in the realm of physical therapy and rehabilitation.

Methods: We introduce a playtest methodology to enhance the design of a VR exergame developed using a user-centered approach for upper limb rehabilitation in stroke survivors. Over 4 playtesting sessions, stroke survivors interacted with initial game versions using VR headsets, providing essential feedback for refining game content and mechanics. Additionally, a pilot study involving 10 stroke survivors collected data through VR-related questionnaires to assess game design aspects such as mechanics, assistance, experience, motion sickness, and immersion.

Results: The playtest methodology was beneficial for improving the exergame to align with user needs, consistently incorporating their perspectives and achieving noteworthy results. The pilot study revealed that users had a positive response. In the first scenario, a carpenter presents a game based on the flexion-extension movement of the elbow; the second scenario includes a *tejo* game (a traditional Colombian throwing game) designed around game mechanics related to the flexion-extension movement of the shoulder; and in the third scenario, a farmer challenges the player to perform a movement combining elbow flexion and extension with internal and external rotation of the shoulder. These findings suggest the potential of the studied exergame as a tool for the upper limb rehabilitation of individuals who have experienced a stroke.

Conclusions: The inclusion of exergames in rehabilitation for stroke-induced hemiparesis has significantly benefited the recovery process by focusing on essential shoulder and elbow movements. These interactive games play a crucial role in helping users regain mobility and restore practical use of affected limbs. They also serve as valuable data sources for researchers, improving the system's responsiveness. This iterative approach enhances game design and markedly boosts user satisfaction, suggesting exergames have promising potential as adjunctive elements in traditional therapeutic approaches.

(JMIR Serious Games 2024;12:e48900) doi:10.2196/48900

KEYWORDS

stroke; user-centered design; exergame; design; virtual reality; playtest; upper limb rehabilitation

Introduction

Background

In recent years, technological advances have influenced motor rehabilitation interventions for survivors of stroke, with the introduction of exergames, known as "serious games for health," which help motivate individuals in their rehabilitation [1-4]. However, the development of such exergames needs to consider users' needs and rehabilitation goals [5].

Virtual reality (VR) immersive [6] systems have become increasingly popular in rehabilitation, as they offer immersive and engaging activities, improving motivation and skill acquisition [7]. Nevertheless, systematic reviews have noted that most VR apps primarily focus on balance and gait, with limited attention to upper extremity rehabilitation [8,9].

Efforts have been made to design exergames tailored for survivors of stroke, but challenges remain, including limited user involvement and lack of immersive VR integration [10,11].

This study aims to address these challenges by designing a VR-based upper limb rehabilitation exergame using a user-centered approach, involving survivors of stroke in the design process and conducting playtests with an immersive VR setup [12]. The methodology aims to improve interdisciplinary collaboration and facilitate the involvement of clinicians in the design process [13,14]. The primary objectives are to provide personalized upper arm physiotherapy for survivors of stroke through an improved VR exergame and to assess its usability through user feedback [15]. This work encourages collaboration among clinicians, researchers, and designers to create an engaging rehabilitation exercise that complements the recovery process for survivors of stroke, ultimately enhancing their quality of life.

Related Work

VR-Based Physical Rehabilitation for Stroke

Experts in rehabilitation, kinesiology, and neuroscience are integrating VR systems with exergames to enhance the appeal and effectiveness of rehabilitation processes [16]. Early studies, such as those by Henrique et al [17] and Burke et al [18], demonstrated the positive impact of exergames on balance, gait, and upper limb motor function in patients with stroke, highlighting improved therapy adherence [14,16-20]. However, systematic reviews have indicated that most VR apps for after-stroke therapy primarily focus on balance and gait, with limited attention to upper extremity rehabilitation [3,21,22]. To address this gap, we aim to evaluate the potential of an exergame for upper limb rehabilitation using immersive VR systems [22].

In addition, prior research has shown that complementing or replacing standard rehabilitation with VR-based rehabilitation can result in significant improvements in gait speed, balance, and mobility in patients with stroke [3,17,21-23]. Our work aims to contribute to the development of guidelines for using VR-based rehabilitation in conjunction with conventional therapy, with a focus on upper limb rehabilitation.

Although some researchers, such as Reis et al [10], Leung et al [11], and Horsham et al [24], have proposed methodologies for developing specific exergames for stroke rehabilitation, there is still limited knowledge regarding immersive VR-based designs targeting upper limb rehabilitation [10,11,24]. Therefore, we intend to involve survivors of stroke in an iterative playtesting process to develop an upper limb VR-based rehabilitation system and bridge this gap.

Playtesting as an Iterative Design for Stroke

This section covers research related to the use of playtesting as an iterative user-centered design (UCD) methodology. UCD has played a significant role in the development of games for rehabilitation and overall health [25,26], as it is a methodology that allows active participation of the target population in the system's prototyping process. UCD, applied in game design, often advocates for an interactive and participative methodology that includes multiple playtests with end players. Playtesting is an activity carried out with potential users or players who interact with game prototypes developed in the early stages, making it easier to gather individual opinions and ideas that contribute to improving the gameplay aspects of exergames during their development [27,28]. Playtesting is a key and standardized methodology used in game studios to iterate and systematically improve games before they are released to the public [29].

A relevant example is the work of Duval et al [30], who conducted a collaborative study with 14 clinicians, focusing on therapeutically validating the game based on their opinions rather than those of users. Duval et al [30] obtained significant findings by addressing the adoption of therapy and personalizing it according to the characteristics valued by medical professionals. In contrast, other UCD works, such as the study by Aguilar et al [31], have not used playtesting but have used usability tests involving scales and flow state questionnaires. Findings from 3 years of experience with exergames developed for older adults using UCD methods concluded that devoting the key to engaging with end users and considering feedback and opinions can be considered the best practice guide for the development of therapeutic games [32]. We believe that playtesting can be beneficial for the design process of games for health, as it is strongly recommended to involve the target audience during the game design and development processes. By doing so, developers increase the likelihood of creating games that consider the specific preferences, motives, and characteristics of survivors of stroke in need of physical therapy [21]. By including survivors of stroke in interactive playtesting and, consequently, in enhancing a VR exergame, we begin to understand how this design methodology affects the subsequent use of the exergame as a therapeutic tool.

Methods

In this section, we introduce the interdisciplinary team that worked on the improvement of the VR exergame we used in the playtesting session and the pilot study, as well as the description of the VR exergame. Furthermore, we present the playtesting methodology and the pilot study methodology.

Interdisciplinary Team

The structured design team was composed of an expert clinical physiatrist who advised the movements that users with stroke are likely to perform from a clinical viewpoint; a physiotherapist who provided permanent follow-up in all sessions with the users; a designer of exergames who helped implement the UCD methodology to have clear game mechanics; an expert in biomechanics who analyzed ranges of movement, postures, and gestures; a user experience researcher who organized all sessions with the users; 2 professional game programmers who created the game prototypes; and 2 users who experienced stroke episodes and interacted with the system and based on the answers they gave us an improved exergame. For 2 months, this group convened weekly to discuss the exergames' requirements, technologies, and overall scope of the project. The discussions were centered on defining the activities in the internet-based environment and strategies for the recruitment of potential users. At the end of the design process, the group of game developers with programming experience used the Unity game engine (Unity Technologies) to materialize the ideas.

The main topics addressed by the interdisciplinary team were (1) the definition of the main objectives and roles of the project; for example, project management was assumed by the exergames designer, and the user experience researcher assumed the role of project manager and conducted most of the fieldwork; (2) socialization of playtesting activities, including user recruitment and experimental protocol; and (3) reconsideration and further adjustment of game design elements, such as game mechanics and their mapping with therapeutic objectives.

Design of the VR Exergame

Prior Design of the VR Exergame

We performed a rapid contextual design based on a previous study, where user profiles were defined using user personas [33]. In this study, we characterized 4 persona roles that distinguish them as gamers: apathetic, empathetic, beginner, and experienced. Specifically, we used the results of the user modeling process to define certain game elements. For example, we found that users showed interest in sports games. Hence, body interaction familiar with certain sports games was an important requirement to be integrated into the exergames. In addition, users were comfortable with game content related to their daily lives. Therefore, incorporating cultural activities into the exergames could be a promising approach. In addition, in this prior study, we considered the following clinical requirements when developing the VR exergame [34,35].

Population specificity: according to previous studies [36], most users who have experienced a stroke are older than 40 years, and few are younger than 30. These studies also showed that the older population has little experience with VR. In contrast, therapist experience suggested that ranges of motion vary across users who have experienced a stroke. Therefore, designers should be careful when adapting internet-based therapy to a wide range of capabilities [37].

Motor learning: different principles of motor recovery should be considered in the creation of the activities to be performed within the exergames, such as a meaningful task, intensive and repetitive practice, movements close to the normal range, muscle activation that drives the practice of movement, and variability and progression of training [10,38,39].

Rehabilitation movements: the movements suggested by clinicians to perform upper limb physical therapy in paretic users are listed in Table 1. These are shoulder flexion and extension, elbow flexion and extension, and Kabat diagonals. Kabat diagonals are internal and external rotation movements of the shoulder. According to Della Tommasina et al [40], a repetitive process of these movements is necessary to perform physiotherapy, from which a more effective range of motion recovery will be obtained [41]. The elbow and shoulder are the upper limbs' main joints, articulating the arm's largest segments. These joints require a greater range of motion in flexion and extension and often affect and limit arm motion when a stroke episode occurs. Hence, we decided that users in a seated position should perform different arm movements while playing the VR-based rehabilitation exergame, targeting multiple possible physical rehabilitation needs in the upper limbs of people with stroke.

Table 1. Rehabilitation movements proposed for the exergame.

	Movement	Application	Action in the exergame
Elbow	Flexion and extension	Improvement in the width of movement in daily life activities	Destroying blocks with a hammer
Shoulder	Flexion and extension	Improvement in the width of movement in daily life activities	Throwing *tejos*
Elbow	Interior and exterior extension	Improved range of motion in the daily life activities	Cutting grass with a machete

Complementary therapy: we propose a therapy that uses the VR-based rehabilitation exergame to complement the rehabilitation process instead of replacing the traditional one, such as the one proposed by Goncalves et al [42]. We are confident that users who have experienced a stroke will play an exergame with engaging activities because they are developed based on their needs and motivations. Therefore, the exergames will allow a disruptive experience different from conventional therapies, generating interest, excitement, and willingness to carry out their rehabilitation process without neglecting the conventional therapy recommended by clinical specialists [38]. Considering the rapid contextual design and the clinical requirements named in this study, in a prior study, we developed a VR exergame following the well-known game design methodologies by Schell [29], as we contextualize in the following subsection.

Motion Health VR: VR Exergame for Stroke Rehabilitation

In a prior study, we established a game design concept and game mechanics using the methodology proposed by Schell [29] and a complete contextual design conducted with users with stroke. This study covers a more systematic and complete description of the playtesting sessions conducted with players with stroke to iterate and improve the game based on the initial concept. Knowing the preferences, ages, and profiles of potential users, we decided to explore a design concept for the cultural regions of Colombia (the Caribbean, Pacific, Andean, Orinoco, and Amazon). We discovered that older adults are inclined to engage in activities in the countryside and typical and authentic Colombian games. We discussed this exergame concept with the clinicians of our design team, gathering feedback regarding the potential movements to be performed, particularly in the context of mapping them for stroke rehabilitation therapies. Considering the scope of the project and previous research conducted with local users, development capabilities, and timelines, we decided to start by developing a game design concept related to the activities of the Colombian Andean region. The Andean region is the central region of Colombia and has crosscutting activities that are representative of the entire country and run throughout the central Andes. The population of this region practices sports, such as *sapo* and *tejo* (throwing games), rowing, and other more well-known sports, such as basketball and boxing. They also engage in other daily activities, such as fruit picking, horseback riding, and bush cutting [43]. Aligned with these cultural activities, we designed game scenarios that focused on a local setting using Colombian games. The design team analyzed existing VR games to establish game mechanics that could involve desired rehabilitation movements, always considering the player's motivators and needs. This analysis facilitated communication between the team of clinicians and specialists and the design and development team while also helping to specify the activities that would be familiar and engaging to the users. Therefore, we called the VR exergame "Motion Health VR." The exergame comprises 3 main scenes in which players must develop 3 different activities: hammering, throwing a metal disk (a traditional Colombian game called *tejo*), and cutting bushes while riding a horse.

The exergame presents 3 meticulously crafted scenarios, each aligned with its unique reference to Figure 1. In the "carpenter" scenario (Figure 1A), players engage in a dynamic elbow flexion and extension challenge, wielding a hammer to systematically crush boxes that vary in color and size, demanding specific ranges of motion. Players must skillfully adjust their proximity to the boxes to adapt to this diverse challenge, seamlessly weaving in back-and-forth and crossbody-reaching motions. Between the box-smashing activity, a captivating puzzle gradually unveils itself, featuring distinct Andean wildlife. Upon completing the activity, players earn the gratifying experience of visualizing the completed animal puzzle. As players enhance their hammering skills, the game dynamically escalates in difficulty either by increasing hammering frequency or by reducing box sizes. In the second scenario, inspired by Colombia's traditional *tejo* game (Figure 1B), players embark on a shoulder-focused flexion and extension adventure, mirroring the popular sport played nationwide. Throwing a metal disk toward an explosive target known as a *mecha* on a clay court, players must adjust their shoulder movements according to the target's distance, finetuning their range of motion for precise throws and aiming to maximize target hits with minimal repetitions. In the third scenario, the "farmer" (Figure 1C), players are challenged with a multifaceted movement that combines elbow flexion, extension, and internal and external rotation of the shoulder, akin to Kabat diagonals. In this rural setting, players ride an internet-based horse while wielding a machete, a staple tool in the Colombian countryside, tasked with clearing the obstructive bushes that appear on both sides of the road. With one arm gripping the machete and the other resting on the horse's rein, players face escalating challenges as the game progresses.

Figure 1. The presented scenarios of the Motion Health VR exergame with (A) a carpenter, (B) a throwing activity, and (C) a farmer, based on the movement of the Kabat diagonals.

Playtesting Sessions

Iterative game design involves playtesting sessions with end users, who will shape the game features before its final deployment. The main objective of playtesting sessions is to iterate different playable prototypes to improve the overall playability of the game, thus increasing the likelihood of adoption. In addition, playtesting allows the researcher to assess the ability of potential players to perform the proposed activities and understand game feedback. We developed playtesting sessions using the VR exergame designed in a prior work.

The playtest sessions were guided by a researcher accompanied by a clinical specialist who helped contact different users who had experienced a stroke. Owing to the COVID-19 pandemic, visits to each user were scheduled in such a way that biosafety protocols were maintained (eg, distancing, constant use of masks, and hand and footwear disinfection). Upon arrival at the agreed location, stakeholders performed the recommended distancing protocols. Then, the researcher prepared the experimental protocol, which consisted of setting up a table, a chair without a hand rest, and verifying the internet connection. Prototypes of the exergames were developed before the playtesting sessions and ported to the VR headsets (Oculus Rift in the first 4 iterations and Oculus Quest in the final version). Disposable headset protectors and cleaners were used to maintain biosafety measures. Table 2 presents the structure of the playtest sessions.

Table 2. Protocol for conducting playtests.

Actions	Considerations	Time (min)
Brief introduction of the dynamics of the session and interaction between the researcher and the user with stroke to obtain informed consent	—[a]	10
A quick explanation of how the VR[b] system works and what to expect from the activity	Preparation and arrangement	5
System implementation (HMD[c], headphones, and controls)	—	—
Free play or natural interaction with the system	Manifestation of difficulties, in real time if necessary	5
Receive feedback or explore ideas while users play	Formulate the questions established for the session	10
Conclude the session with questions about the experience	Questions	10

[a]Not available.

[b]VR: virtual reality.

[c]HMD: head-mounted display.

A total of 4 playtesting sessions involving 9 end users who had experienced a stroke were conducted. Each session was performed with a minimum of 2 users chosen considering their availability (Multimedia Appendix 1 lists the users who participated in each session and their demographic information). After playtesting, we recorded a video summarizing the session and documented a brief analysis that was subsequently discussed with the research and design teams. After each playtest, the team held a general meeting where all the discussions were presented. The subsequent playtest was scheduled after the implementation of the suggested game changes.

Pilot Study: Evaluation of the Game Experience and Usability of the Exergames

After conducting game playtests and completing a playable prototype of the VR exergame (4 iterations), we decided to carry out a pilot study to evaluate the usability of the game with a group of users who had experienced a stroke, in which 2 users who participated in one of the 4 iterations were part of the pilot study group.

We conducted a 20-minute session that was part of the rehabilitation therapy in which users who had experienced a stroke played the iterated version of the Motion Health VR exergame. We ported the final version of the game to the standalone VR headset, Oculus Quest 2, as it has several advantages, such as being wireless, comfortable, and having a high image resolution.

Users

This usability study was developed with 10 users who had experienced a stroke contacted through the clinician and therapist of the design team. We chose this sample size conveniently, considering the availability of users, which was very limited. In contrast, the small sample size allowed us to follow the biosafety protocols required for the COVID-19 pandemic, which was still ongoing in Colombia at the time of the study. The inclusion criteria for the study were being aged >50 years, having experienced a stroke and having hemiparesis or monoparesis, being able to read and write, not having serious vision problems (eg, strabismus), and not having diagnosed cognitive disabilities (eg, dementia).

Ethical Considerations

The bioethics committee of the local university approved this study, which was also approved by the bioethics committee of a local rehabilitation center (52–050623). Users volunteered for this study and agreed to participate by signing an informed consent form.

Usability Study

Two questionnaires (instruments) were used to assess the game user experience immediately after interacting with the immersive game.

Virtual Reality Neuroscience Questionnaire

The Virtual Reality Neuroscientific Questionnaire (VRNQ) measures the quality of user experience, game mechanics, and in-game assistance. It comprises 20 questions, each scored on a Likert-type scale ranging from 0 to 5 [44]. The advantage of using this questionnaire is that it provides the limits to assess the suitability of the software in VR [44]. VRNQ produces a total score that reflects the overall quality of the VR software and 4 categories as follows: (1) game experience, where the level of immersion and pleasure of the experience are evaluated; (2) game mechanics, where user interaction in the internet-based environment is evaluated; (3) game assistance, where the exergame instructions, indications, arrows, and labels are evaluated; and (4) motion sickness, which evaluates whether you experience nausea, disorientation, fatigue, and instability.

Immersive Tendencies Questionnaire

The Immersive Tendencies Questionnaire (ITQ) determines the differences in an individual's tendencies to experience immersion and presence after interacting with a VR scenario. ITQ comprises 18 questions rated on a Likert scale from 1 to 7, resulting in a possible score ranging from 18 to 126. In the original study, the mean score of the samples was 76.66 [45]. This questionnaire is considered a standard in VR research and has been widely used in different applications [22]. A user with a positive immersive tendency based on the ITQ score is likely to experience higher levels of VR presence, which has been associated with better task performance [22,41,46]. This questionnaire is useful because it allows the evaluation of immersion in a way that does not depend on the specific internet-based environment, making it possible to determine independently if an internet-based environment performs poorly or if the statistical sample has low immersion trends. Some ITQ questions are as follows:

- Does it often happen that while daydreaming, you forget what is happening around you?
- Does it happen to you that you are so engrossed in a movie that you forget what is happening around you?
- Do you identify with television characters?
- When you use an exergame, does it occur to you that you feel like you are inside the game instead of sitting down using the controller?
- Do you stay scared for a while after watching a scary movie?

Experimental Setup

The researcher and the physical therapist held the interaction session at each participant's home, where they chose a comfortable space to set up the VR system. The setup consisted of a chair in which the user with stroke was seated with the Oculus Quest 2 wearable headset and its respective wireless controllers. The researchers were able to see what the players were doing via the official Meta Quest app using an electronic tablet in real time.

Protocol

The users began the pilot study session seated, using the headset and holding the controllers. We conducted the session in the following order. The user who had experienced a stroke performed an upper body warm up for 5 minutes, guided by the physical therapist. The researcher and physical therapist prepared the user for the game by helping them put on the VR system. The researcher started the exergame, which was presented throughout the session, accompanied by a physiotherapist. After the interaction, the user who had experienced a stroke completed the 2 proposed questionnaires. Given the biosafety regulations established by the Colombian government because of the COVID-19 pandemic, all those involved in the sessions always wore masks. In addition, the researcher cleaned all VR and mounting elements each time they were used.

Data Analysis

The questionnaires were scored following the instructions of previous works. Descriptive statistics, such as mean and SD, were calculated and reported [47,48].

Results

Overview

The results are presented in 2 subsections. The first subsection details the transformation process of the game after conducting 3 playtesting sessions and iterations involving the end users and the interdisciplinary design team. The second subsection presents the preliminary results of a pilot study evaluating the user experience of the final game with a group of 10 players with stroke.

Playtesting

The following subsections detail the results of the iterative design process of the Motion Health VR exergame, reporting the details of the playtesting sessions and the modifications made to each iteration. The overall objectives of the playtest were to validate the acceptance and playability of users with stroke and to explore whether they were able to perform the activities proposed in the VR scenarios and game mechanics. In addition, analysis of errors and optimization of game mechanics were crucial to improve playability. A total of 9 users with stroke were involved throughout the 4 playtesting sessions conducted (Multimedia Appendix 2), focusing on certain game elements (eg, mechanics and esthetics) via playable prototypes and reporting back to the design team.

Analysis of Playtest 1

This version of the exergame was created to test the first 2 scenarios. Players 1 and 2 (U1 and U2) participated in this session following the protocol in Table 2. We used a VR-ready laptop and an Oculus Rift headset with controllers. The playtesting goals were to evaluate the appropriateness of the proposed range of movement for hammering and throwing the disk and to explore button combinations for performing the activity using the controllers. We found that (1) the game should consider different scales of spasticity to provide a more adaptive experience [49]; (2) the buttons should be suspended from their functions to avoid triggering involuntary functions; and (3) a rest period should be granted to the user because, as recommended by the physiotherapist, long periods of exercise generate symptoms of fatigue.

Analysis of Playtest 2

The objective of these playtests was to test the modifications introduced in the first 2 scenarios based on the considerations in the Analysis of Playtest 1 subsection. Players U3, U4, and U5 participated in this playtest following the protocol in Table 2. A VR-ready laptop and an Oculus Rift headset with controllers were used. Figure 2 shows the evolution of the 2 scenarios after the first playtest, showing the improvement in content according to the real scenarios where these activities were performed.

Figure 2. The final version of each scenario of the Motion Health VR exergame: (A), the carpenter scenario, (B) the tejo scenario, and (C) the farmer scenario.

Users with stroke reported an improvement in the simplicity of the interaction, as they found it much easier to perform the movement owing to an initial calibration of the position added to the 2 scenarios, which adjusts the player's position concerning the internet-based surroundings, ensuring objects are at a reachable distance. In addition, we found that (1) the objects in each scenario should be in a static position; thus, people can avoid unnecessary displacements within the internet-based space that can generate dizziness; (2) we need to improve the auditory feedback of scenarios to create an immersive experience; and (3) we need to improve the calibration scene to allow players with low mobility to perform the tasks.

Analysis of Playtest 3

The third scenario was prototyped, and the game mechanics were ready for playtesting. The objective of this playtest was to test whether players could easily understand and interact in the farmer scenario by performing the proposed movements, that is, riding the horse while holding the rein with the unaffected arm and cutting bunches with the machete using the affected arm. Players U6 and U7 (Multimedia Appendix 3) participated in this test following the protocol in Table 2. We used a VR-ready laptop and an Oculus Rift headset with controllers. We found that (1) it would be useful to place the avatar on the horse from the beginning and (2) the game should allow cutting bushes to be performed using both arms and provide adequate time to switch the game controller between hands because the players become tired after certain repetitions.

Analysis of Playtest 4

After iterating each scenario and exploring potential pitfalls and interaction errors, a final playtesting session was scheduled to test the overall functioning of the integrated system. This prototype of the exergame presented an embellishment of the contents (Figure 2), which was an improvement in the overall esthetics of the game. In addition, a structured exercise session was recommended by the clinical rehabilitation experts, following a 15-minute session (similar to other studies of the same nature [50]). Therefore, the 3 scenarios were presented in sequence and switched after 5 minutes (approximately). Players U3, U4, and U5 participated in this test following the protocol in Table 2. We used a VR-ready laptop and an Oculus Rift headset with controllers.

We found that (1) scene transitions should be smoothed and (2) the game should implement rest periods between each scenario, as users still manifested mild fatigue from performing so many repetitions while preparing for the next mechanic.

Finally, we integrated the above recommendations into the scenarios and developed the final prototype of the Motion Health VR exergame.

Evaluation of Game User Experience

This section presents the results of evaluating the game user experience of the co-designed Motion Health VR exergame involving 10 users with stroke (Multimedia Appendix 4). For this part, the game was modified to a more portable, standalone, and easy-to-use headset, the Oculus Quest 2. Only the final deployment platform was changed (from wired to wireless VR), and no other changes were made. The questionnaires were administered at the end of the session, asking users to rate their experience in a wide range of aspects following the ITQ and VRNQ.

Virtual Reality Neuroscientific Questionnaire

The results of administering VRNQ are reported as average values with SD (Table 3). Each category had a maximum of 35 points. Gaming experience was rated at a mean of 24.8 (SD 4.5), game mechanics mean 23.8 (SD 5.5), game assistance mean 23.9 (SD 5.5), and motion sickness (inversely proportional), mean 31.1 (SD 5.6). Consequently, the maximum possible general score for this test was 140, in which the exergame Motion Health VR obtained a mean of 103.6 (SD 19.4). This level of quality is considered more than adequate as it exceeded 100 points. From this, it can be concluded that the users experienced a high level of immersion during their video game experience, and the quality of the Motion Health VR exergame obtained a general average of mean 103.6 (SD 19.4), which is considered an adequate quality because it exceeded 100 points [47]. On the basis of this result, we observed that users had a high immersion index; the experience with the exergames was very pleasant; and the quality of the graphics, sound, and technology, in general, was perceived as very positive. Finally, the system showed the best results in the motion sickness index, which shows that exergames did not cause major side effects associated with cybersickness or nausea [47,51].

Table 3. Virtual Reality Neuroscientific Questionnaire (VRNQ) categories.

VRNQ categories	Score, mean (SD)
Game experience	24.8 (4.5)
Game mechanics	23.8 (5.5)
Game attendance	23.9 (5.5)
Motion sickness	31.1 (5.6)

Immersive Tendencies Questionnaire

ITQ was used to evaluate users' immersion experience and presence following their engagement with the exercise. The overall ITQ score averaged 60.8 (SD 11.6), signifying that participants with a history of stroke perceived relatively low to moderate levels of immersion and enjoyment (Table 4). In terms of concentration (mean 22.3, SD 2.9, with a maximum score of 35), users consistently achieved high scores, indicating that the game effectively captured their attention.

Regarding immersion (mean 17.9, SD 4.7, with a maximum score of 35), users reported a sense of engagement with the game. In terms of emotions (mean 14.6, SD 5.1, with a maximum score of 28), the findings suggest that users developed a strong emotional connection with the game [33,51,52].

Table 4. Immersive Tendencies Questionnaire (ITQ) categories.

ITQ categories	Score, mean (SD)
Concentration	22.3 (2.9)
Immersion	17.9 (4.7)
Emotion	14.6 (5.1)
Enjoyment	6.1 (2.2)

Discussion

Principal Findings

This study summarizes our efforts to report methodological approaches extensively used in game design that have a significant value when used to design VR exergames for stroke. We also showed the results of a preliminary usability test conducted involving 10 users with stroke who played the Motion Health VR exergame after completing 4 iterations using playtesting sessions. Overall, the VR exergame exhibited medium-to-high levels of game user experience and low levels of perceived symptoms associated with VR, such as nausea and dizziness. Moreover, regarding enjoyment, users expressed a high willingness to participate in therapies and continue with the sessions. The user experience questionnaires showed that users experienced increased immersion, emotional connection, and enjoyment with the VR exergame, although the concentration remained consistent. These results are consistent with previous research using immersive VR in older adults [53] and people with stroke [54,55]. Furthermore, we have carefully reported the methodological aspects related to playtesting sessions with users with stroke and specific procedures to conduct such sessions. From the playtesting sessions, we can extract the value of evaluating prototypes in the early stages of the game design process because this prevents researchers from struggling with complex interactivity and usability issues later during the trials. As reported in the study by Toro et al [56], early involvement of end users in VR systems for exercise promotion is a desired practice, and it is not commonly used among those creating custom-made exergames for older adults [57-59].

Playtesting as a Tool for Iterative Design

Playtesting is a part of the iterative design methodology used in different UCD approaches [51,60]. In our case, we performed playtest with several users who had experienced a stroke, which allowed us to improve the content, playability index, and game mechanics from an ergonomic approach to provide greater user comfort. Although prior designs of physical rehabilitation games for stroke have involved UCD [61], involving users with stroke in playtesting and follow-up sessions is not common [50]. We suggest that designers consider including playtesting with users with stroke because, in terms of rehabilitation therapy performance, playtesting revealed important details, such as the importance of performing a calibration stage or removing the buttons and other interactions with the VR equipment. We found that this stage provides exergames with the characteristics to adapt to the physical needs of each user, such as the range of motion and spasticity scale of each user [62]. In the context of serious games, the importance of adaptive games has increased, as every user has different requirements [62].

Furthermore, the playtesting methodology allowed us to strengthen relationships with all stakeholders, from developers to clinicians. This aligns with the findings of previous studies that have emphasized how involving multiple stakeholders in the design process leads to a more suitable and user-centered prototype [63]. The importance of maintaining close relationships with stakeholders has also been underscored, and this study reaffirms the relevance of this challenge in the successful implementation of playtesting. Effective collaboration between developers and clinicians, driven by the willingness and availability of both parties, has been a recurring theme in the literature [62]. Previous studies have pointed out the pivotal role of this relationship in the success of VR rehabilitation programs following stroke.

In particular, we noticed that playtesting allowed a stronger relationship with patients with stroke, increasing their willingness to participate in future studies. Although other researchers have reported difficulty finding specific populations to be involved in studies [64,65], considering our final results, we highly recommend that research teams plan to conduct multiple playtests before conducting studies.

Finally, we consider that after performing 4 playtests, based on the observations of the users who had experienced a stroke and the recommendations of the clinicians, an optimal version of the exergame was obtained, which could also be used by designers to facilitate piloting and prevent errors during data collection. As mentioned in this study, the inclusion of playtests and the collaboration of specialized professionals in programming and design align with the best practices recommended by previous research [66].

Usability of UCD VR Games in Patients With Stroke

We observed that ITQ and VRNQ scores were below the expected mean, as reported in the Results section. Similar results have been reported previously because there were some concerns about the usability of VR in older adults, including those who have had strokes. A systematic review of clinical and research applications of VR in older people identified usability issues, such as discomfort, cybersickness, and difficulty with the equipment [67]. Therefore, based on the current results, the design team must improve the gameplay mechanics and usability of the VR exergame to achieve better results in gameplay experience metrics before the subsequent trial. Nevertheless, the system scored high in usability, as its overall score was higher than expected (100 points). Furthermore, as our results showed that the VRNQ category with the highest score was motion sickness, meaning that users who experienced a stroke felt little nausea, our work aligns with studies using similar VR apps [8,20,47]. These usability results can guide other exergame designers to adjust their apps to suit older adults.

Use of Interactive Technology for Telehealth Care

The use of portable and autonomous technologies, such as the Oculus Quest 2 headset, during the pandemic has been an innovative response supported by this research [68]. This reinforces the notion that virtualization of health care, driven by technology, is becoming increasingly important. A recent review describing the promising landscape of telerehabilitation tools aided by serious games for upper limb stroke rehabilitation highlights the evidence of efficacy, the need for further research in this area, and the promise of digitally connected games to complement conventional rehabilitation [69]. Nevertheless, although VR has never been more accessible before, the reality is that the cost-effectiveness of its use in telehealth programs in both rural and urban areas in Latin America is still very limited [70]. In summary, our study contributes to emerging efforts in which interdisciplinary collaboration and the use of innovative technology during times of crisis, such as the pandemic, continue to draw a research pathway in this continually evolving field [63,71,72].

Limitations

We developed this study between August 2020 and June 2021, when most rehabilitation centers were closed owing to COVID-19 pandemic restrictions. Therefore, accessing users with stroke was a difficult task that we overcame with the help of the therapists. They provided us with the contact list of their former users, and we contacted them personally. Notably, although we took all safety precautions, people with stroke feared contagion and only a couple of them participated in the playtesting sessions. We acknowledge the lack of homogeneity in the sample of users who participated in the playtesting because stroke is a condition caused by several factors and affects both sexes, and users who have experienced a stroke tend to have a wide range of ages and ethnicities. The small sample size may limit the generalizability of our findings. Moreover, this limitation was difficult to address because of the COVID-19 pandemic restrictions that were under regulation when we developed this study. That is why, for the pilot study, we limited the users' age to >50 years. VR is a technology that is constantly changing and improving, in the sense that we started playtesting with the Oculus Rift headset and then moved on to using the Oculus Quest 2 for the pilot study because of its portability advantages. We overcame these technological changes because of the cross-platform features offered by the game engine used (Unity).

However, VR content development is a challenge when designing deployable solutions. Nonetheless, despite the pandemic situations in which this study was developed, the iterative design and preliminary study were carried out owing to the implementation of portable and autonomous tools, such as the Oculus Quest 2 VR headset, which are becoming increasingly important in the virtualization of health care delivery.

The duration of the usability study was very short, and users only interacted with the final game in a single session lasting approximately 15 minutes. A short-term study may not adequately capture the long-term benefits or challenges of using VR exergames for stroke rehabilitation. We plan to extend this initial pilot study and conduct a single-arm longitudinal study involving a similar group of users for 12 sessions for 3 months. Furthermore, this study did not include a control group for comparison. Without a control group undergoing traditional rehabilitation methods, it is challenging to conclusively attribute improvements to the VR exergame alone. Finally, although the study emphasizes the iterative design process and user feedback, it is difficult to know how these design changes directly affect the rehabilitation outcomes. Future studies with this game should include rehabilitation outcomes such as upper limb range of motion and spasticity levels.

Conclusions

Our research has conclusively demonstrated that creating VR exergames for stroke rehabilitation by involving end users early in the design stages brings advantages such as reducing interaction errors and unnecessary game design elements that do not contribute to the therapy. UCD is highly recommended as a design methodology for creating games specifically tailored to the rehabilitation of users who have experienced strokes. The results of this study support the effectiveness of multiple playtesting in producing therapeutic games that align with the needs and abilities of users with stroke, such as creating familiar internet-based environments and activities and removing unnecessary motion that could lead to motion sickness. This conclusion underscores the importance of adopting a patient-centered approach in the development of medical apps and technologies for rehabilitation. Our findings have yielded promising results regarding the use of immersive VR in the context of upper limb stroke rehabilitation. Users who immersed themselves in internet-based environments using custom-built exergames showed good levels of immersion and enjoyment and reduced levels of perceived nausea or dizziness. These results suggest that VR technology holds potential as a therapeutic tool for the treatment of users with stroke-related impairments, especially for at-home therapies. However, further research and long-term follow-up are required to fully understand the scope and limitations of this technology in the rehabilitation of this user group. The use of playtesting as an iterative tool for enhancing video game design enables comprehensive interaction with the user. This interaction allows for genuine customization of the therapy, leading to the development of a video game tailored specifically for users who have experienced stroke.

Acknowledgments

The authors would like to thank Felipe Gomez for his contribution as a physical therapist to the design team. The authors also thank game programmer Ricardo Arango, who contributed to the construction of the exergames. Finally, the authors would like to thank the Comfamiliar Hospital and all the users who helped with testing the exergames.

Data Availability

The data sets generated during or analyzed during this study are available from the corresponding author upon request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Users in the playtests.
[DOCX File , 15 KB - games_v12i1e48900_app1.docx]

Multimedia Appendix 2
Demographic data of the user’s study pilot.
[DOCX File , 15 KB - games_v12i1e48900_app2.docx]

Multimedia Appendix 3
Changes made throughout the iteration process of the playtests.
[DOCX File , 16 KB - games_v12i1e48900_app3.docx]

Multimedia Appendix 4
Scenario sketches.
[DOCX File , 60 KB - games_v12i1e48900_app4.docx]

References

1. Koutsiana E, Ladakis I, Fotopoulos D, Chytas A, Kilintzis V, Chouvarda I. Serious gaming technology in upper extremity rehabilitation: scoping review. JMIR Serious Games 2020 Dec 11;8(4):e19071 [FREE Full text] [doi: 10.2196/19071] [Medline: 33306029]
2. Fidopiastis CM, Rizzo AA, Rolland JP. User-centered virtual environment design for virtual rehabilitation. J Neuroeng Rehabil 2010 Feb 19;7(1):11 [FREE Full text] [doi: 10.1186/1743-0003-7-11] [Medline: 20170519]
3. Domínguez-Téllez P, Moral-Muñoz JA, Salazar A, Casado-Fernández E, Lucena-Antón D. Game-based virtual reality interventions to improve upper limb motor function and quality of life after stroke: systematic review and meta-analysis. Games Health J 2020 Feb 01;9(1):1-10. [doi: 10.1089/g4h.2019.0043] [Medline: 32027185]
4. Vieira C, Ferreira da Silva Pais-Vieira C, Novais J, Perrotta A. Serious game design and clinical improvement in physical rehabilitation: systematic review. JMIR Serious Games 2021 Sep 23;9(3):e20066 [FREE Full text] [doi: 10.2196/20066] [Medline: 34554102]
5. Tatla SK, Shirzad N, Lohse KR, Virji-Babul N, Hoens AM, Holsti L, et al. Therapists' perceptions of social media and video game technologies in upper limb rehabilitation. JMIR Serious Games 2015 Mar 10;3(1):e2 [FREE Full text] [doi: 10.2196/games.3401] [Medline: 25759148]

6. Patsaki I, Dimitriadi N, Despoti A, Tzoumi D, Leventakis N, Roussou G, et al. The effectiveness of immersive virtual reality in physical recovery of stroke patients: a systematic review. Front Syst Neurosci 2022 Sep 22;16:880447 [FREE Full text] [doi: 10.3389/fnsys.2022.880447] [Medline: 36211591]
7. Levac D, Glegg S, Colquhoun H, Miller P, Noubary F. Virtual reality and active videogame-based practice, learning needs, and preferences: a cross-Canada survey of physical therapists and occupational therapists. Games Health J 2017 Aug;6(4):217-228. [doi: 10.1089/g4h.2016.0089] [Medline: 28816511]
8. Rose T, Nam CS, Chen KB. Immersion of virtual reality for rehabilitation - review. Appl Ergon 2018 May;69:153-161. [doi: 10.1016/j.apergo.2018.01.009] [Medline: 29477323]
9. Pizzoli SF, Mazzocco K, Triberti S, Monzani D, Alcañiz Raya ML, Pravettoni G. User-centered virtual reality for promoting relaxation: an innovative approach. Front Psychol 2019 Mar 12;10:479 [FREE Full text] [doi: 10.3389/fpsyg.2019.00479] [Medline: 30914996]
10. Reis A, Lains J, Paredes V, Filipe V, Abrantes C, Ferreira F, et al. Developing a system for post-stroke rehabilitation: an exergames approach. In: Proceedings of the 10th International Conference, UAHCI 2016, Held as Part of HCI International 2016 on Universal Access in Human-Computer Interaction, Users and Context Diversity. 2016 Presented at: UAHCI '16; July 17-22, 2016; Toronto, ON p. 403-413 URL: https://link.springer.com/chapter/10.1007/978-3-319-40238-3_39 [doi: 10.1007/978-3-319-40238-3_39]
11. Leung RC, Zhang H, Tao X. Serious game design for stroke rehabilitation. Int J Inf Technol 2017;23(1):1-25 [FREE Full text] [doi: 10.32657/10356/75902]
12. Nelson CR, Gabbard JL, Moats JB, Mehta RK. User-centered design and evaluation of ARTTS: an augmented reality triage tool suite for mass casualty incidents. In: Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality. 2022 Presented at: ISMAR '22; October 17-21, 2022; Singapore, Singapore p. 336-345 URL: https://ieeexplore.ieee.org/document/9994987/ [doi: 10.1109/ismar55827.2022.00049]
13. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
14. Peláez-Vélez FJ, Eckert M, Gacto-Sánchez M, Martínez-Carrasco Á. Use of virtual reality and videogames in the physiotherapy treatment of stroke patients: a pilot randomized controlled trial. Int J Environ Res Public Health 2023 Mar 08;20(6):4747 [FREE Full text] [doi: 10.3390/ijerph20064747] [Medline: 36981652]
15. Montoya MF, Villada JF, Cardona JE, Gallo OA, López JF. Diseño contextual para la creación de videojuego basado en Realidad Virtual usado en terapia de rehabilitación física en personas con accidente cerebrovascular. Rev EIA 2022;19(38):1-30 [FREE Full text] [doi: 10.24050/reia.v19i38.1549]
16. Bui T, King C, Llado A, Lee D, Leong G, Paraparum A, et al. App-based supplemental exercise during inpatient orthopaedic rehabilitation increases activity levels: a pilot randomised control trial. Pilot Feasibility Stud 2019 Mar 16;5(1):47 [FREE Full text] [doi: 10.1186/s40814-019-0430-9] [Medline: 30923628]
17. Henrique PP, Colussi EL, De Marchi AC. Effects of exergame on patients' balance and upper limb motor function after stroke: a randomized controlled trial. J Stroke Cerebrovasc Dis 2019 Aug;28(8):2351-2357. [doi: 10.1016/j.jstrokecerebrovasdis.2019.05.031] [Medline: 31204204]
18. Burke JW, McNeill MD, Charles DK, Morrow PJ, Crosbie JH, McDonough SM. Optimising engagement for stroke rehabilitation using serious games. Vis Comput 2009 Aug 27;25(12):1085-1099. [doi: 10.1007/s00371-009-0387-4]
19. Corbetta D, Imeri F, Gatti R. Rehabilitation that incorporates virtual reality is more effective than standard rehabilitation for improving walking speed, balance and mobility after stroke: a systematic review. J Physiother 2015 Jul;61(3):117-124 [FREE Full text] [doi: 10.1016/j.jphys.2015.05.017] [Medline: 26093805]
20. Charles D, Holmes D, Charles T, McDonough S. Virtual reality design for stroke rehabilitation. In: Rea PM, editor. Biomedical Visualisation. Cham, Switzerland: Springer; 2020:53-87.
21. Duque E, Fonseca G, Vieira H, Gontijo G, Ishitani L. A systematic literature review on user centered design and participatory design with older people. In: Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems. 2019 Presented at: IHC '19; October 22-25, 2019; Vitória Espírito Santo, Brazil p. 1-11 URL: https://dl.acm.org/doi/10.1145/3357155.3358471 [doi: 10.1145/3357155.3358471]
22. Faas D, Bao Q, Frey DD, Yang MC. The influence of immersion and presence in early stage engineering designing and building. Des Comput Cognit 2014 May 16;28(2):139-151 [FREE Full text] [doi: 10.1017/s0890060414000055]
23. Herrlich M, Smeddinck J, Runge N, Malaka R. Applying human-centered design to develop motivating exergames. In: Proceedings of the 2016 Mensch und Computer Workshop. 2016 Presented at: MuC '16; September 4–7, 2016; Aachen, Germany p. 1-5 URL: https://dl.gi.de/server/api/core/bitstreams/edfb8f03-cf13-465b-a3fd-3726f9b63f52/content
24. Horsham C, Dutton-Regester K, Antrobus J, Goldston A, Price H, Ford H, et al. A virtual reality game to change sun protection behavior and prevent cancer: user-centered design approach. JMIR Serious Games 2021 Mar 25;9(1):e24652 [FREE Full text] [doi: 10.2196/24652] [Medline: 33764308]
25. Javaid M, Haleem A. Virtual reality applications toward medical field. Clin Epidemiol Glob Health 2020 Jun;8(2):600-605 [FREE Full text] [doi: 10.1016/j.cegh.2019.12.010]
26. Jerald J. The VR book: Human-Centered Design for Virtual Reality. San Rafael, CA: Morgan & Claypool; 2015.

27. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 2017 Nov 20;11(11):CD008349 [FREE Full text] [doi: 10.1002/14651858.CD008349.pub4] [Medline: 29156493]
28. Langener S, Klaassen R, VanDerNagel J, Heylen D. Immersive virtual reality avatars for embodiment illusions in people with mild to borderline intellectual disability: user-centered development and feasibility study. JMIR Serious Games 2022 Dec 07;10(4):e39966 [FREE Full text] [doi: 10.2196/39966] [Medline: 36476721]
29. Schell J. The Art of Game Design: A Target Book. Boca Raton, FL: CRC Press; 2008.
30. Duval J, Thakkar R, Du D, Chin K, Luo S, Elor A, et al. Designing spellcasters from clinician perspectives: a customizable gesture-based immersive virtual reality game for stroke rehabilitation. ACM Trans Access Comput 2022 Aug 19;15(3):1-25 [FREE Full text] [doi: 10.1145/3530820]
31. Aguilar FA, Bautista DP, Gómez A, Toledo GT, García OS. User-centered virtual environment for poststroke motor rehabilitation. J Med Devices 2022 Jun;16(2):021014 [FREE Full text] [doi: 10.1115/1.4053605]
32. Brox E, Konstantinidis ST, Evertsen G. User-centered design of serious games for older adults following 3 years of experience with exergames for seniors: a study design. JMIR Serious Games 2017 Jan 11;5(1):e2 [FREE Full text] [doi: 10.2196/games.6254] [Medline: 28077348]
33. Soegaard M, Dam RF. The Encyclopedia of Human-Computer Interaction. Aarhus, Denmark: Interaction Design Foundation; 2012.
34. Chu CH, Biss RK, Cooper L, Quan AM, Matulis H. Exergaming platform for older adults residing in long-term care homes: user-centered design, development, and usability study. JMIR Serious Games 2021 Mar 09;9(1):e22370 [FREE Full text] [doi: 10.2196/22370] [Medline: 33687337]
35. Muñoz J, Mehrabi S, Li Y, Basharat A, Middleton LE, Cao S, et al. Immersive virtual reality exergames for persons living with dementia: user-centered design study as a multistakeholder team during the COVID-19 pandemic. JMIR Serious Games 2022 Jan 19;10(1):e29987 [FREE Full text] [doi: 10.2196/29987] [Medline: 35044320]
36. Mirnig AG, Meschtscherjakov A, Wurhofer D, Meneweger T, Tscheligi T. A formal analysis of the ISO 9241-210 definition of user experience. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. 2015 Presented at: CHI EA '15; April 18-23, 2015; Seoul, Republic of Korea p. 437-450 URL: https://dl.acm.org/doi/10.1145/2702613.2732511 [doi: 10.1145/2702613.2732511]
37. Lima Rebêlo F, de Souza Silva LF, Doná F, Sales Barreto A, de Souza Siqueira Quintans J. Immersive virtual reality is effective in the rehabilitation of older adults with balance disorders: a randomized clinical trial. Exp Gerontol 2021 Jul 01;149:111308. [doi: 10.1016/j.exger.2021.111308] [Medline: 33744393]
38. Levin MF, Weiss PL, Keshner EA. Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles. Phys Ther 2015 Mar;95(3):415-425 [FREE Full text] [doi: 10.2522/ptj.20130579] [Medline: 25212522]
39. Arya KN, Pandian S, Verma R, Garg R. Movement therapy induced neural reorganization and motor recovery in stroke: a review. J Bodyw Mov Ther 2011 Oct;15(4):528-537. [doi: 10.1016/j.jbmt.2011.01.023] [Medline: 21943628]
40. Della Tommasina I, Trinidad-Morales A, Martínez-Lozano P, González-de-la-Flor Á, Del-Blanco-Muñiz J. Effects of a dry-land strengthening exercise program with elastic bands following the Kabat D2 diagonal flexion pattern for the prevention of shoulder injuries in swimmers. Front Physiol 2023 Nov;14:1275285 [FREE Full text] [doi: 10.3389/fphys.2023.1275285] [Medline: 38028779]
41. Dix A. Human–computer interaction: a stable discipline, a nascent science, and the growth of the long tail. Interact Comput 2010 Jan;22(1):13-27. [doi: 10.1016/j.intcom.2009.11.007]
42. Gonçalves A, Muñoz J, Cameirão MS, Gouveia É, Sousa H, Bermúdez I Badia S. The benefits of custom exergames for fitness, balance, and health-related quality of life: a randomized controlled trial with community-dwelling older adults. Games Health J 2021 Aug;10(4):245-253. [doi: 10.1089/g4h.2020.0092] [Medline: 34370609]
43. Aguirre JJ, Gómez JA. Del animal laborans al homo ludens: la configuración del juego y la fiesta como voces de resistencia al poder. PACARINA, Revista Latinoamericana. 2012. URL: https://rescoldoredpacarina.com/wp-content/uploads/2020/11/Revista-Pacarina-N3.pdf#page=26 [accessed 2023-12-12]
44. Somrak A, Pogačnik M, Guna J. Suitability and comparison of questionnaires assessing virtual reality-induced symptoms and effects and user experience in virtual environments. Sensors (Basel) 2021 Feb 08;21(4):1185 [FREE Full text] [doi: 10.3390/s21041185] [Medline: 33567570]
45. Lloréns R, Gil-Gómez JA, Alcañiz M, Colomer C, Noé E. Improvement in balance using a virtual reality-based stepping exercise: a randomized controlled trial involving individuals with chronic stroke. Clin Rehabil 2015 Mar 23;29(3):261-268. [doi: 10.1177/0269215514543333] [Medline: 25056999]
46. Yin CW, Sien NY, Ying LA, Chung SF, Tan May Leng D. Virtual reality for upper extremity rehabilitation in early stroke: a pilot randomized controlled trial. Clin Rehabil 2014 Nov;28(11):1107-1114. [doi: 10.1177/0269215514532851] [Medline: 24803644]
47. Kourtesis P, Linnell J, Amir R, Argelaguet F, MacPherson SE. Cybersickness in virtual reality questionnaire (CSQ-VR): a validation and comparison against SSQ and VRSQ. Virtual Worlds 2023 Jan 29;2(1):16-35 [FREE Full text] [doi: 10.3390/virtualworlds2010002]

48. Kazemi R, Lee SC. Human Factors/Ergonomics (HFE) evaluation in the virtual reality environment: a systematic review. Int J Hum Comput Interact 2023 Jun 29:1-17 [FREE Full text] [doi: 10.1080/10447318.2023.2227835]
49. Platz T, Eickhof C, Nuyens G, Vuadens P. Clinical scales for the assessment of spasticity, associated phenomena, and function: a systematic review of the literature. Disabil Rehabil 2005 Jul 07;27(1-2):7-18. [doi: 10.1080/09638280400014634] [Medline: 15799141]
50. Cikajlo I, Rudolf M, Mainetti R, Borghese NA. Multi-exergames to set targets and supplement the intensified conventional balance training in patients with stroke: a randomized pilot trial. Front Psychol 2020 Nov;11(6):572-590 [FREE Full text] [doi: 10.3389/fpsyg.2020.00572] [Medline: 32300321]
51. Holtzblatt K, Beyer H. Contextual Design: Defining Customer-Centered Systems. Burlington, MA: Morgan Kaufmann; 1997.
52. Saltiveri TG. MPIu+a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares. Universitat de Lleida. 2007. URL: https://www.tdx.cat/handle/10803/8120#page=1 [accessed 2023-12-11]
53. Saeedi S, Ghazisaeedi M, Rezayi S. Applying game-based approaches for physical rehabilitation of poststroke patients: a systematic review. J Healthc Eng 2021:3 [FREE Full text]
54. Hung YX, Huang PC, Chen KT, Chu WC. What do stroke patients look for in game-based rehabilitation: a survey study. Medicine (Baltimore) 2016 Mar;95(11):e3032 [FREE Full text] [doi: 10.1097/MD.0000000000003032] [Medline: 26986120]
55. Bedar K, Bubanovich C, Rosemore J, Radford K, Taylor KL. Virtual reality intervention and its impact on upper extremity function in the stroke population: a scoping review. Games Health J 2023 Dec;12(6):429-439. [doi: 10.1089/g4h.2022.0239] [Medline: 37948555]
56. Carrión-Toro M, Santorum M, Acosta-Vargas P, Aguilar J, Pérez M. iPlus a user-centered methodology for serious games design. Appl Sci 2020 Dec 17;10(24):9007 [FREE Full text] [doi: 10.3390/app10249007]
57. Rojo A, Castrillo A, López C, Perea L, Alnajjar F, Moreno JC, et al. PedaleoVR: usability study of a virtual reality application for cycling exercise in patients with lower limb disorders and elderly people. PLoS One 2023 Feb 22;18(2):e0280743 [FREE Full text] [doi: 10.1371/journal.pone.0280743] [Medline: 36812248]
58. Wu J, Zeng A, Chen Z, Wei Y, Huang K, Chen J, et al. Effects of virtual reality training on upper limb function and balance in stroke patients: systematic review and meta-meta-analysis. J Med Internet Res 2021 Oct 12;23(10):e31051 [FREE Full text] [doi: 10.2196/31051] [Medline: 34636735]
59. Muñoz JE, Montoya MF, Boger J. From exergames to immersive virtual reality systems: serious games for supporting older adults. In: Choukou MA, Syed-Abdul S, editors. Smart Home Technologies and Services for Geriatric Rehabilitation. Cambridge, MA: Academic Press; 2022:141-204.
60. Hussain N. Virtual reality based kinematics for assessment of post-stroke upper limb function. University of Gothenburg. 2020. URL: https://gupea.ub.gu.se/handle/2077/63288 [accessed 2023-12-11]
61. Santoso HB, Schrepp M, Kartono Isal RY, Yudha Utom AY, Priyogi B. Measuring user experience of the student-centered e-learning environment. J Educ Online 2016 Jan;13(1):1-22 [FREE Full text] [doi: 10.9743/jeo.2016.1.5]
62. Proffitt R, Glegg S, Levac D, Lange B. End-user involvement in rehabilitation virtual reality implementation research. J Enabling Technol 2019;13(2):92-100 [FREE Full text] [doi: 10.1108/JET-10-2018-0050] [Medline: 31663039]
63. Krishnan S, Mandala MA, Wolf SL, Howard A, Kesar TM. Perceptions of stroke survivors regarding factors affecting adoption of technology and exergames for rehabilitation. PM R 2023 Nov 12;15(11):1403-1410. [doi: 10.1002/pmrj.12963] [Medline: 36787167]
64. Montoya MF, Muñoz J, Henao O. Design of an upper limbs rehabilitation videogame with sEMG and biocybernetic adaptation. In: Proceedings of the 5th Workshop on ICTs for improving Patients Rehabilitation Research Techniques. 2019 Presented at: REHAB '19; September 11-13, 2019; Popayan, Columbia p. 152-155 URL: https://dl.acm.org/doi/abs/10.1145/3364138.3364170#:~:text=By%20using%20an%20adaptation%20mechanism,followed%20an%20iterative%20design%20process [doi: 10.1145/3364138.3364170]
65. Fregna G, Schincaglia N, Baroni A, Straudi S, Casile A. A novel immersive virtual reality environment for the motor rehabilitation of stroke patients: a feasibility study. Front Robot AI 2022 Aug 29;9:906424 [FREE Full text] [doi: 10.3389/frobt.2022.906424] [Medline: 36105763]
66. Muñoz JE, Dautenhahn K. Robo ludens: a game design taxonomy for multiplayer games using socially interactive robots. ACM Trans Hum Robot Interact 2021 Jul 14;10(4):1-28 [FREE Full text] [doi: 10.1145/3451343]
67. Tuena C, Pedroli E, Trimarchi PD, Gallucci A, Chiappini M, Goulene K, et al. Usability issues of clinical and research applications of virtual reality in older people: a systematic review. Front Hum Neurosci 2020 Apr 8;14:93 [FREE Full text] [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
68. Lau ST, Siah RC, Dzakirin Bin Rusli K, Loh WL, Yap JY, Ang E, et al. Design and evaluation of using head-mounted virtual reality for learning clinical procedures: mixed methods study. JMIR Serious Games 2023 Aug 30;11:e46398. [doi: 10.2196/46398] [Medline: 37647108]
69. Amorim P, Santos BS, Dias P, Silva S, Martins H. Serious games for stroke telerehabilitation of upper limb - a review for future research. Int J Telerehabil 2020 Dec 08;12(2):65-76. [doi: 10.5195/ijt.2020.6326] [Medline: 33520096]

70. Prvu Bettger J, Liu C, Gandhi DB, Sylaja PN, Jayaram N, Pandian JD. Emerging areas of stroke rehabilitation research in low- and middle-income countries: a scoping review. Stroke 2019 Nov;50(11):3307-3313. [doi: 10.1161/strokeaha.119.023565]
71. Ruiz-Rodriguez A, Hermens HJ, van Asseldonk EH. HEROES, design of an exergame for balance recovery of stroke patients for a home environment. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI EA '23; April 23-28, 2023; Hamburg, Germany p. 1-4 URL: https://dl.acm.org/doi/10.1145/3544549.3577037 [doi: 10.1145/3544549.3577037]
72. Tao G, Garrett B, Taverner T, Cordingley E, Sun C. Immersive virtual reality health games: a narrative review of game design. J Neuroeng Rehabil 2021 Feb 11;18(1):31 [FREE Full text] [doi: 10.1186/s12984-020-00801-3] [Medline: 33573684]

Abbreviations

ITQ: Immersive Tendencies Questionnaire
UCD: user-centered design
VR: virtual reality
VRNQ: Virtual Reality Neuroscientific Questionnaire

Edited by T Leung, C Prahm; submitted 10.05.23; peer-reviewed by K Stawarz, P André; comments to author 02.08.23; revised version received 28.09.23; accepted 24.11.23; published 11.01.24.

Please cite as:
Villada Castillo JF, Montoya Vega MF, Muñoz Cardona JE, Lopez D, Quiñones L, Henao Gallo OA, Lopez JF
Design of Virtual Reality Exergames for Upper Limb Stroke Rehabilitation Following Iterative Design Methods: Usability Study
JMIR Serious Games 2024;12:e48900
URL: https://games.jmir.org/2024/1/e48900
doi:10.2196/48900
PMID:38206670

Original Paper

Comparing Walking-Related Everyday Life Tasks of Children with Gait Disorders in a Virtual Reality Setup With a Physical Setup: Cross-Sectional Noninferiority Study

Sophia Rhiel[1,2], MSc; Andrina Kläy[1,2], MSc; Urs Keller[1,2], PhD; Hubertus J A van Hedel[1,2], PT, PhD; Corinne Ammann-Reiffer[1,2], PT, PhD

[1]Swiss Children’s Rehab, University Children's Hospital Zurich, University of Zurich, Affoltern am Albis, Switzerland

[2]Children’s Research Center, University Children’s Hospital Zurich, University of Zurich, Zurich, Switzerland

Corresponding Author:
Sophia Rhiel, MSc
Swiss Children’s Rehab, University Children's Hospital Zurich
University of Zurich
Mühlebergstrasse 104
Affoltern am Albis, 8910
Switzerland
Phone: 41 44 762 52 97
Email: sophia.rhiel@kispi.uzh.ch

Abstract

Background: A frequent rehabilitation goal for children with gait disorders is to practice daily-life walking activities. Unfortunately, these are often difficult to practice in a conventional therapeutic setting. Virtual reality (VR) with head-mounted displays (HMDs) could be a promising approach in neurorehabilitation to train such activities in a safe environment. First, however, we must know whether obstacles in VR are indeed mastered as obstacles.

Objective: This study aimed to provide information on whether VR is feasible and motivating to induce and practice movements needed to master real obstacles in children and adolescents with gait disorders. Furthermore, this project aims to evaluate which kinds of everyday walking activities are appropriate to be practiced in VR.

Methods: In this cross-sectional study, participants stepped over a bar, crossed a gap, balanced over a beam, and circumvented stationary obstructions arranged in a course under real physical and virtual conditions wearing a VR HMD. We recorded the respective primary outcomes (step height, step length, step width, and minimal shoulder-obstacle distance) with motion capture. We then calculated the mean differences and 95% CI of the spatiotemporal parameters between the VR and physical setup and later compared them using noninferiority analysis with margins defined a priori by a clinical expert panel. Additionally, the participants responded to a standardized questionnaire while the therapists observed and evaluated their movement performance.

Results: We recruited 20 participants (mean age 12.0, range 6.6-17.8 years) with various diagnoses affecting their walking ability. At 3.77 (95% CI 1.28 to 6.26) cm, the mean difference in step height of the leading foot in the overstepping task did not exceed the predefined margin of –2 cm, thus signifying noninferiority of the VR condition compared to mastering the physical obstacles. The same was true for step length (–1.75, 95% CI –4.91 to 1.41 cm; margin –10 cm), step width (1.05, 95% CI 0.20 to –1.90 cm; margin 3 cm), and the minimal shoulder-obstacle distance (0.25, 95% CI –0.85 to 0.35 cm; margin –2 cm) in the other tasks. Only the trailing foot in the overstepping task yielded inconclusive results.

Conclusions: Children with gait disorders perform everyday walking tasks like overstepping, crossing, balancing, or circumventing similarly in physical and VR environments, suggesting that VR could be a feasible therapeutic tool to practice everyday walking tasks.

(JMIR Serious Games 2024;12:e49550) doi:10.2196/49550

KEYWORDS

adolescent; child; gait; head-mounted display; motion capture; neurological rehabilitation; noninferiority trial; physical therapy; virtual reality; walking

Introduction

In pediatric neurorehabilitation, children and adolescents with congenital or acquired lesions of the sensorimotor system often experience impairments in gait [1,2]. Consequently, recovery of walking ability is a frequent rehabilitation goal in pediatric neurorehabilitation [3]. Thereby, the focus is on promoting everyday life activities and ensuring meaningful participation for the child and their family [4]. Therapies targeting gait encompass a wide variety of therapeutic approaches. In our clinic, Swiss Children's Rehab, these therapies include, for example, conventional physical therapy, including task- and everyday life–oriented training, rehabilitation robots, and sports therapy. Normally, these therapies occur in a conventional therapeutic setting. However, within this setting, many everyday walking tasks, such as, for example, crossing a wide gap to board public transportation or avoiding contact with people or obstacles while navigating through crowded places, cannot be reasonably practiced.

In recent years, immersive virtual reality (VR) has become increasingly popular. Since companies have made the technology more accessible to the community through more affordable and easy-to-use devices, the use of VR has increased, as have the areas of its use [5]. Accordingly, this upswing in VR could be promising for its implementation in neurorehabilitation. Immersive VR puts users directly into virtual scenarios and gives the illusion of a full physical presence, providing rich sensory fidelity (high degree of reliability) [6,7]. To experience immersive VR, head-mounted displays (HMDs) are most suitable and can convey many of the abovementioned impressions [8]. A potential goal of using VR in pediatric neurorehabilitation could be to enhance children's abilities in their daily lives by practicing task-specific activities relevant to their everyday lives while still being in a safe therapeutic environment. Furthermore, its game-like attributes and animations can increase children's motivation and enhance their active participation by minimizing their focus on task repetitions [9,10]. Additionally, as VR is an accessible and affordable technology, it could enable home training. Moreover, a significant advantage of using VR in children aged between 6 and 18 years could be that they experience higher levels of presence and "realness" within a virtual environment compared to adults [11].

Recent studies have already investigated the effectiveness of acquiring different cognitive and motor tasks with VR. In the pediatric field, VR has been mainly used for pain management [6] or educational purposes [12,13], as well as to create relaxing and learning opportunities for children diagnosed with autism spectrum disorder [14,15] or attention deficit hyperactivity disorder [16]. However, the long-term effects of VR on developing children are unknown, and cybersickness or fatigue of the eyes and brain are potential disadvantages [6,17,18]. According to the authors' best knowledge, no evidence exists of using immersive VR as a gait therapy intervention in children with gait disorders. When including results from augmented reality studies, a systematic review showed moderate evidence for improved gait-related outcomes when gait training was enhanced with commercially available videogame systems, such as the Nintendo Wii or Microsoft Xbox Kinect, in children with cerebral palsy (CP) [19]. Furthermore, a systematic review and meta-analysis from Chen et al [20] showed a large effect size of d=0.861 for improved motor function in children with CP when comparing commercially available game systems with conventional therapy or controls (eg, no intervention). However, such systems lack essential aspects of VR since they are usually presented on a 2D screen or as floor projections [8] and, therefore, do not transmit the entire concept of VR, including full physical presence and immersion.

Immersive VR offers many advantages regarding task-specific training, motivation, "realness," and costs [5-7]. Still, it remains uncertain whether the use of VR in children with gait disorders is a feasible approach to inducing and practicing the movements required to perform everyday gait activities. Reasons to assume that VR in children with gait disorders might not be feasible are the lack of visual information of the lower extremities and the difference in the perception of virtual obstacles by the children [6,11,21]. Therefore, a prerequisite for the meaningful use of VR in training everyday gait activities would be that the children master obstacles presented in VR like they master physical obstacles. Thus, this project aims to provide information on whether a VR setup is feasible and motivating to induce and practice movements that are needed to master real obstacles in children and adolescents with gait disorders. Furthermore, this project aims to evaluate which kinds of everyday walking activities are appropriate to be practiced in such a VR setup. To evaluate this, we compare the spatiotemporal parameters of performing certain everyday walking tasks in a virtual and a physical environment using a noninferiority analysis. The noninferiority analysis should indicate that the virtual setup is not unacceptably worse than the physical setup.

Methods

Ethical Considerations

This cross-sectional study took place at the gait laboratory of Swiss Children's Rehab, University Children's Hospital Zurich, during a single 60-minute session. The ethics committee of the Canton of Zurich confirmed through a clarification of responsibility that no approval was needed for this study (Req-2021-00364).

Participants

We included children and adolescents aged between 6 and 18 years with gait disorders undergoing inpatient or outpatient rehabilitation at Swiss Children's Rehab. In line with recommendations for comparative studies, which propose 8 to 25 participants [22], we aimed to include 20 participants. All children who were receiving physiotherapy at the time of recruitment were screened according to the inclusion and exclusion criteria and recruited consecutively within 3 months. To be eligible to participate, they had to be able to walk short indoor distances without assistive devices or with crutches. Additionally, they had to be able to follow simple verbal instructions. Exclusion criteria were a history of seizures, epilepsy, blindness, or inability to use the HMD (eg, cybersickness, open wounds on the head).

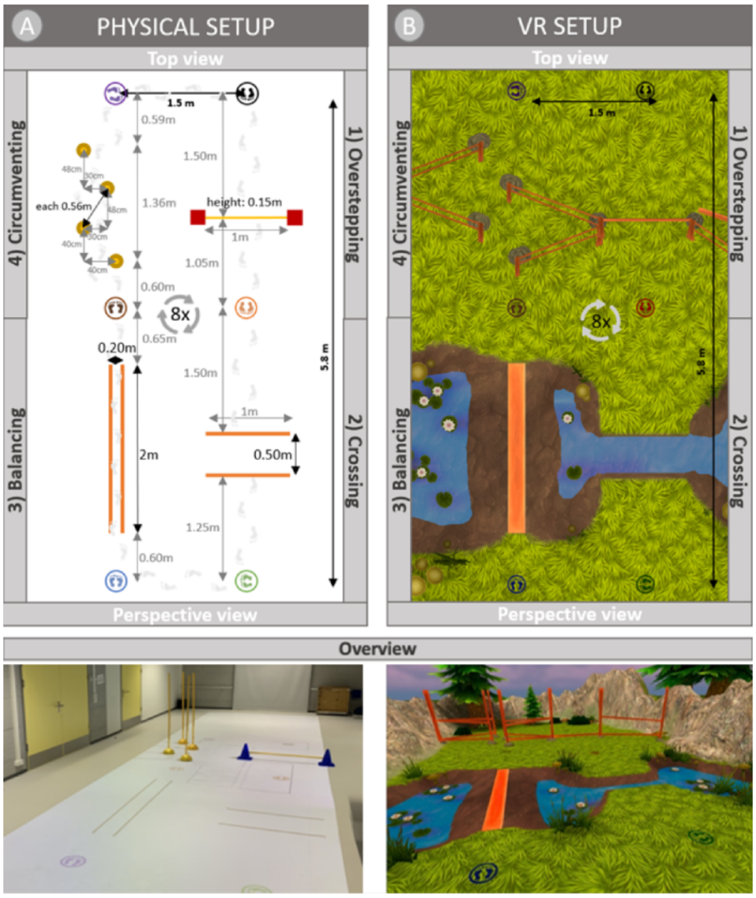
Participants’ characteristics were collected from the patient records. The physiotherapist rated the functional mobility level using 2 performance measures: the Functional Mobility Scale (FMS) and the Gillette Functional Assessment Questionnaire (FAQ) walking scale [23]. The FMS describes the participant’s level of functional mobility by assessing the assistive device used in everyday life over 5 m, 50 m, and 500 m on a scale from 1 (uses a wheelchair) to 6 (independent on any terrain). The FAQ assesses functional walking abilities on a scale from 1 (can not make any steps at all) to 10 (walks, runs, and climbs on even and uneven terrain). Finally, the lower extremity proprioceptive impairments of the participants were rated with the percentage score of the proprioception subsection of the Fugl-Meyer (FM) assessment for the lower extremities [24]. We assessed proprioception at the hip, knee, ankle, and toe joints while the participant was supine and barefoot.

According to good clinical practice standards, we obtained written informed consent from the participants and their legal representatives before participation.

Experimental Setup

The participants had to perform everyday walking tasks in 2 different conditions: physical setup and VR setup. In the physical setup, the participants had to master real (physical) obstacles (Figure 1A). The 4 obstacles, including overstepping, crossing, balancing, and circumventing, were arranged in a course. In the VR setup, the participants had to master the same 4 obstacles virtually. The obstacles were incorporated into an everyday environment (Figure 1B). The VR setup matched the locations and dimensions, but not the appearance of the physical obstacles. This discrepancy was chosen intentionally since we wanted to incorporate the obstacles into an everyday environment as they would appear in future applications. During the development process, it was ensured that the environment was designed as stimulatingly as possible, since interaction and sensorimotor contingencies are crucial contributors to a full VR experience [8]. Nevertheless, to compare the 2 conditions, we also had to keep the VR environment simple to avoid the participants being distracted from their tasks.

Figure 1. (A) Scheme of physical setup, and (B) the appearance of the VR setup when wearing the HMD.



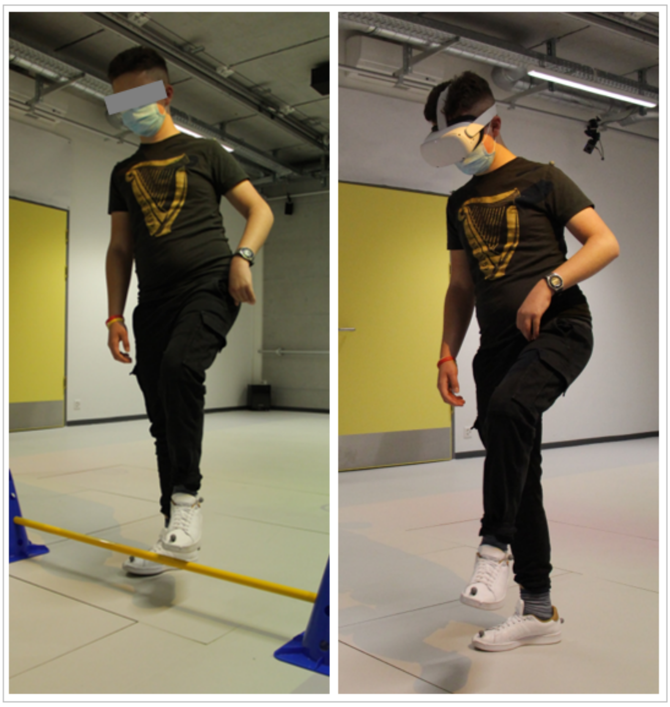
For this setup, the commercially available VR HMD Meta Quest 2 (Meta Platforms) was used. We aligned the coordinate systems of the physical and the virtual world, using the hand-tracking function of the Meta Quest, and scaled and rotated the virtual world based on 2 points. To test the alignment between the 2 conditions, we checked that the scaling coefficient was near 1.0.

To minimize the influence of fatigue, we randomized the sequence of the conditions and the starting position within the obstacle course. We used a minimization method (randomization factor 1), including the factors of gender, age, and functional walking ability defined by the FAQ. During the session, the physiotherapist accompanied the participants to ensure their safety and provide assistance if necessary.

Task Description

For the overstepping task, the participants had to step over a 15-cm-high obstacle, which consisted of a plastic bar mounted on 2 cones (physical setup) or the lower part of a fence (VR setup; Figure 2). In the physical setup, participants had to cross two 3-cm-wide lines projected on the ground with a beamer, whereas they had to cross a small stream in the VR setup. In both setups, the gap was 50 cm, thus exceeding the average step length of children with CP aged between 7 and 14 years (Gross Motor Function Classification System [GMFCS] levels I and II) or traumatic brain injury (TBI) [25-27]. For the balancing task, we instructed the participants to walk between two 2-cm-wide lines projected 20 cm apart on the floor in the physical setup and a 20-cm-wide wooden panel over a pond in the VR setup. Circumventing was performed by walking around 4 plastic poles (physical setup) or fence posts (VR setup). The distance of the poles was 56 cm, corresponding to approximately 1.7 times the average shoulder width of children aged between 6 and 18 years [28,29]. With an estimated protective zone of 30 cm around the obstacle [30], even smaller participants would sidestep, while taller participants could still pass through the obstacles, even when relying on crutches. In addition to the 4 tasks, the participants walked 6.5 m in a straight line without any obstacles, both with the HMD (walking on green grass) and without the HMD.

Figure 2. Execution of the overstepping tasks in the physical and virtual reality setups.



Measurement Procedure

Task execution was recorded with a total of 12 Vicon Vero 2.2 high-speed cameras (Vicon Motion Systems). We placed 9 infrared reflective markers of 16 mm diameter on specific anatomical landmarks at both feet (3 markers each) and shoulders (3 markers). The markers were attached to the shoes as the participants performed the tasks with shoes and orthotics (if needed) as in everyday life.

After measuring the participants' height and shoulder width and attaching the 9 reflective markers to the defined positions, the measurements started with either the physical or the VR condition. The participants first walked 4 times along the 6.5-meter walkway at self-selected walking speeds. Afterward, they performed 2 accommodation rounds of the obstacle course to familiarize themselves with the condition and the tasks. The physiotherapist could provide physical support if the participants had difficulties with any obstacle. Finally, we instructed the participants to always step over the obstacle and cross the gap with the leg they had spontaneously used in the first round.

According to Redekop et al [31], reliability with an interclass correlation coefficient of 0.90 is given for an average of 6 strides when examining discrete gait parameters in children with CP. Therefore, 8 trials per condition were recorded to have 2 spare measurements if any unexpected errors arose while reviewing the recordings. Once the 8 valid attempts per task were recorded, the participants had a short break, during which they answered the first part of the questionnaire. Subsequently, the same procedure was repeated with the second condition, followed by the second part of the participants' questionnaire and the proprioception subsection of the FM assessment performed by the investigator. Meanwhile, the physiotherapist completed the therapist's questionnaire and rated the participant's FMS and FAQ.

Data Processing

Vicon data were processed using Nexus Motion Capture Software (version 7.2; Vicon Inc). Processing of the raw data included visual determination and defining gait events like foot strike, foot off, etc. We analyzed the data from the first 6 valid trials for each condition and task. Then, the data were exported to MATLAB R2021a (version 9.10; MathWorks) to calculate the spatiotemporal parameters. For the spatiotemporal parameters, we calculated the mean of the 6 valid trials per task for each participant and condition individually. A negative mean difference between the VR and physical setup indicated a smaller value in the VR setup.

Outcome Measures

For the 4 tasks, we selected spatiotemporal parameters (Figure 3) in line with the literature [27,32-35]. We calculated the walking speed, step length and width, and double-stance phase during normal walking with and without the VR HMD. Additionally, we recorded the time to master each task and the number of failures, indicating unsuccessful obstacle negotiations.

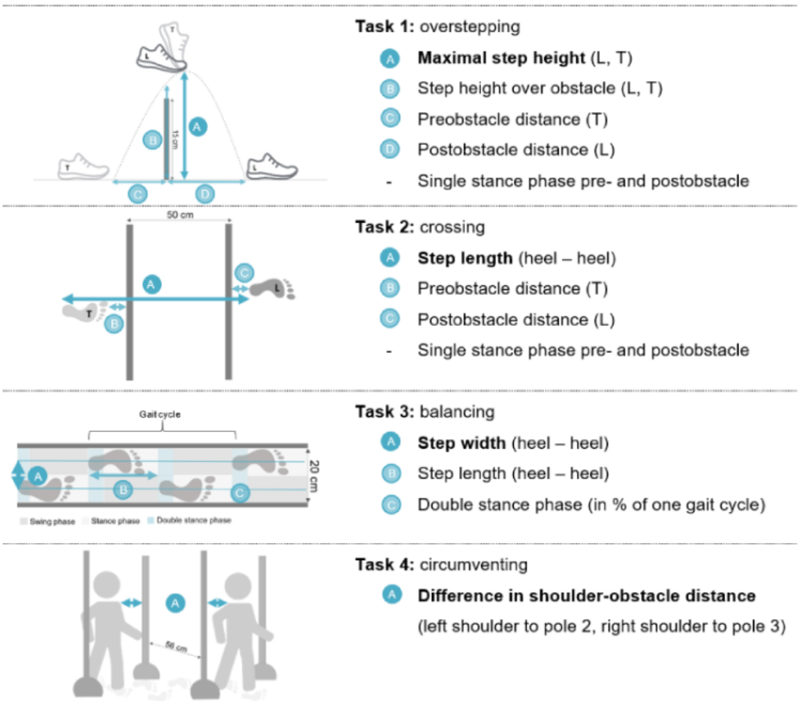


Figure 3. Investigated spatiotemporal parameters for each task. (A) represents the respective primary outcomes. L: leading foot; T: trailing foot.

The participants answered standardized questions covering their movement ability, spatial presence, and enjoyment during task execution on a visual analog scale (VAS). In addition, the physiotherapists rated the participants' movement execution, level of engagement, and meaningful use on a 5-point Likert scale.

Statistical Analysis

Participants' clinical and functional characteristics are presented using descriptive statistics. A normal distribution could be assumed for the differences between the primary outcomes (Shapiro-Wilk test; P>.05). Therefore, the mean differences and their SDs were subsequently calculated. Additionally, the primary spatiotemporal parameters were analyzed using noninferiority testing with 95% CIs and a priori defined margins of noninferiority [36]. The noninferiority margins, which served as boundaries for the 95% CI of the mean differences, were defined for each task by a panel of 15 expert physiotherapists (n=14 women; n=1 men). These margins represent the maximum difference between the VR and the physical setup defined as acceptable while still considering the conditions to be equal [37]. To determine the maximum tolerated deviation, the physiotherapists compared the tasks with everyday life tasks and considered what deviation they would accept in conventional therapy for the respective task. A normal distribution could not be assumed with 15 responses; therefore, we described the margins using nonparametric parameters such as the median and IQR. Descriptive statistics are used to present the participants' and therapists' questionnaire responses. Additionally, to analyze the difference in fun between the 2 conditions, we used the Wilcoxon signed rank test with continuity correction.

Results

Participants

All patients that were examined for eligibility within the recruitment agreed to participate. In total, 7 girls and 13 boys with different gait disorders participated in this study. Their mean age was 12.0 (SD 3.5) years, and their mean height was 1.46 (SD 0.21) meters. All participants were able to follow the instructions and remained compliant during the measurements. None of the participants reported cybersickness. The spectrum of functional mobility was broad, including FMS levels 3-6 for 5 m and 50 m and 1-6 for 500 m, as well as levels 6-10 of the FAQ. However, most participants could walk independently on all surfaces without any walking device, for at least short to medium distances (FMS 5 m and FMS 50 m ≥5 each). Participants' lower extremity proprioception (FM score) ranged from normal to mildly impaired. A total of 9 of the 20 participants had already used a VR HMD at least once before this study. Participants' clinical and functional characteristics are presented in Table 1.

Table 1. Clinical and functional characteristics of the participants.

ID	Sex	Age (years)	Height (cm)	Diagnosis[a]	FMS[b]	FAQ[c]	FM[d]	Mobility aid[e]
1	Male	13.4	155	Unilateral spastic cerebral palsy (I)	6/5/5	9	93%[f]	None
2	Male	13.4	157	Unilateral spastic cerebral palsy (I)	6/6/6	9	100%[f]	None
3	Female	17.0	165	Vasomotor dysregulation with neurological involvement	5/3/3	9	93%[f]	Forearm crutches
4	Male	9.3	135	Brain tumor	6/6/6	9	94%	None[g]
5	Male	14.2	166	Polytrauma	6/6/5	9	94%	None
6	Male	17.8	176	Spinal tumor with neurological involvement	6/5/5	8	94%	None[g]
7	Male	8.0	141	Stroke	6/6/6	10	100%	None
8	Female	16.8	163	Myasthenia gravis	6/6/5	9	94%	None
9	Female	8.0	121	Rhabdomyolysis	6/6/5	9	94%	None
10	Male	6.6	110	Brain tumor	6/6/6	9	88%	None
11	Male	13.6	148	Myelomeningocele	5/3/1	7	100% [f]	Forearm crutches
12	Female	10.9	147	Stroke	6/6/6	9	100%	None
13	Male	15.1	160	Myelomeningocele	3/3/1	6	100%	Forearm crutches
14	Male	14.5	165	Stroke	6/6/6	9	100%	None
15	Female	13.4	171	Ataxia	6/6/5	9	100%	None
16	Male	11.6	145	Bilateral spastic cerebral palsy (I)	6/6/5	9	100%	None
17	Male	7.0	112	Arthrogryposis Multiplex Congenita	5/5/2	7	94%	None[g]
18	Male	9.7	118	Myelomeningocele	5/5/1	9	88%	None
19	Female	8.3	121	Unilateral spastic cerebral palsy (I)	6/6/6	10	100%	None
20	Female	10.9	142	Brain tumor	6/5/5	8	94%	None[g]

[a]In children and adolescents diagnosed with cerebral palsy, the Gross Motor Function Classification System Level is given in parentheses.

[b]FMS: Functional Mobility Scale 5/50/500 m.

[c]FAQ: Gillette Functional Assessment Questionnaire-walking scale.

[d]FM: Fugel-Meyer assessment.

[e]Mobility aid used in both conditions.

[f]Due to restricted movements in certain joints or due to pain, not all movements of the FM could be performed by these participants. Therefore, for these participants, the relative value is not calculated from the maximum score (16 points), but from the individual maximum score (8-14 points).

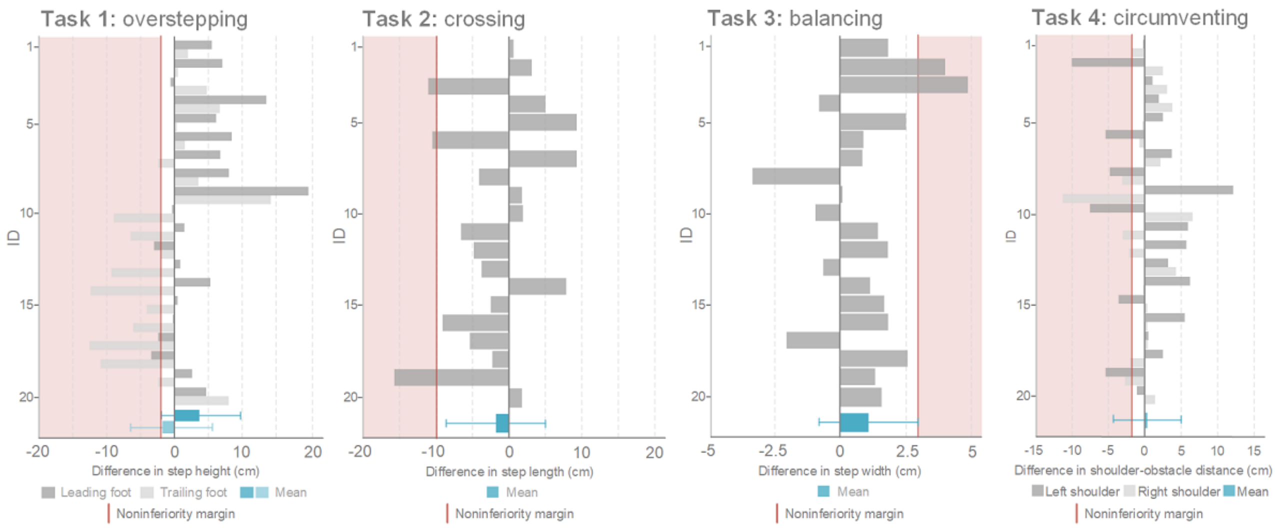
[g]Did not need a mobility aid, but needed close supervision of their physiotherapist.

The participants had to walk the obstacle course from 8 to 16 times to obtain 6 valid trials per task. This resulted in 25-39 recordings per participant for the entire measurement. The most frequent reason why a trial was considered invalid was crossing the obstacle with the wrong leading foot. Furthermore, some attempts were declared invalid when the instructions were not followed or the recording of the markers failed. There were no missing data, except for participant 9 (only 5 valid crossing task trials in the physical setup) and participant 10 (only 5 valid overstepping task trials in the physical setup).

Spatiotemporal Parameters

The differences between the VR and the physical condition varied widely between the participants and tasks (Figure 4).

Figure 4. Differences per participant and task for the primary outcomes. The blue bars represent the mean differences and SD over all participants.



During normal walking, step length and gait speed decreased, and step width slightly increased in the VR condition compared to the physical setup (Table 2). In task 1, participants lifted the leading foot 3.77 cm higher and the trailing foot 1.75 cm lower in the VR setup when overstepping the obstacle. In task 2, they decreased the step length by 1.75 cm in the VR setup. As in normal walking, step width and the double stance phase increased, while step length decreased in the VR setup of the balancing task. For task 4, the distance from the shoulder to the obstacle did not differ between the 2 conditions.

Table 2. Spatiotemporal parameters for the conditions and tasks.

Task and parameter	Physical setup	Virtual reality setup	Difference[a]
Task 0: normal walking, mean (SD)			
Step length (cm)	60.44 (10.22)	54.91 (7.11)	–5.53 (7.14)
Step width (cm)	9.29 (3.92)	9.48 (3.07)	0.19 (2.07)
Gait speed (m/second)	1.10 (0.23)	0.95 (0.20)	–0.15 (0.24)
Double stance phase (%)	24.75 (4.41)	27.48 (3.82)	2.72 (4.21)
Task 1: overstepping			
Max step height (L[b]; cm)[c], mean (SD)	27.53 (4.74)	31.31 (7.21)	3.77 (5.69)
Max step height (T[d]; cm)[c], mean (SD)	28.30 (6.27)	26.55 (8.47)	–1.75 (7.07)
Step height over obstacle (L; cm), mean (SD)	24.77 (5.18)	25.30 (8.29)	0.53 (5.64)
Step height over obstacle (T; cm), mean (SD)	25.32 (5.61)	18.80 (9.06)	–6.52 (8.28)
Preobstacle distance (T; cm), mean (SD)	16.45 (7.66)	10.17 (9.01)	–6.28 (5.60)
Postobstacle distance (L; cm), mean (SD)	19.60 (5.67)	24.45 (6.97)	4.85 (5.58)
Single stance preobstacle (T; seconds), mean (SD)	0.70 (0.17)	0.75 (0.16)	0.05 (0.10)
Single stance postobstacle (L; seconds), mean (SD)	0.62 (0.14)	0.60 (0.10)	–0.01 (0.12)
Total time (seconds), mean (SD)	3.64 (1.49)	4.03 (1.16)	0.39 (0.84)
Total failures max step height <16 cm (L), n (number of children)	1 (1)[e]	4 (2)[e]	3 (1)[e]
Total failures max step height <16 cm (T), n (number of children)	1 (1)[e]	15 (3)[e]	14 (2)[e]
Task 2: crossing			
Step length (cm)[c], mean (SD)	83.81 (7.11)	82.06 (9.32)	–1.75 (7.22)
Preobstacle distance (T; cm), mean (SD)	6.36 (4.55)	–5.91 (8.27)	–12.27 (8.87)
Postobstacle distance (L; cm), mean (SD)	3.29 (5.99)	13.82 (7.50)	10.53 (6.96)
Single stance preobstacle (T; seconds), mean (SD)	0.61 (0.14)	0.69 (0.17)	0.08 (0.15)
Single stance postobstacle (L; seconds), mean (SD)	0.54 (0.08)	0.54 (0.09)	0.00 (0.07)
Total time (seconds), mean (SD)	4.05 (1.26)	4.69 (1.11)	0.64 (0.79)
Total failures step length <51 cm, n (number of children)	14 (7)[e]	31 (10)[e]	17 (3)[e]
Task 3: balancing			
Step width (cm)[c], mean (SD)	5.36 (2.92)	6.41 (2.69)	1.05 (1.93)
Step length (cm), mean (SD)	52.73 (8.51)	47.31 (11.56)	–5.41 (8.45)
Double stance phase (%), mean (SD)	28.58 (5.35)	32.55 (6.05)	3.97 (6.39)
Total time (seconds), mean (SD)	4.44 (1.43)	5.31 (1.72)	0.87 (1.88)
Total failures step width >19 cm, n (number of children)	6 (3)[e]	5 (3)[e]	–1 (0)[e]
Task 4: circumventing			
Minimal shoulder-obstacle distance (cm)[c], mean (SD)	10.66 (3.36)	10.41 (3.77)	0.25 (4.44)
Total time (seconds), mean (SD)	5.25 (2.48)	5.76 (1.98)	0.50 (1.51)
Total failures minimal distance <2 cm, n (number of children)	3 (3)[e]	13 (7)[e]	10 (4)[e]

[a]The differences were calculated by subtracting the value of the physical setup from the value of the virtual reality setup. Consequently, negative differences indicate a lower value for the virtual reality setup.

[b]L: leading foot.

[c]Primary outcomes (also used to define the number of fails).

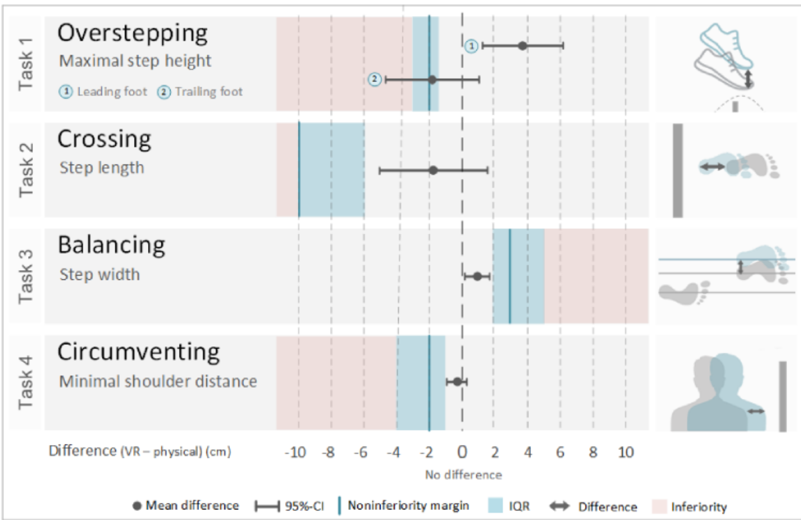
[d]T: trailing foot.

[e]The number of children that made these fails.

Noninferiority Analysis

We applied noninferiority analyses [37] to compare the differences in the primary outcomes between the VR and physical setups for each task according to the a priori defined noninferiority margins. As depicted in Figure 5, the noninferiority analysis revealed noninferiority for the leading foot and was inconclusive for the trailing foot when overstepping the obstacle. For crossing, balancing, and circumventing, the results of the statistical analysis showed noninferiority in all cases.

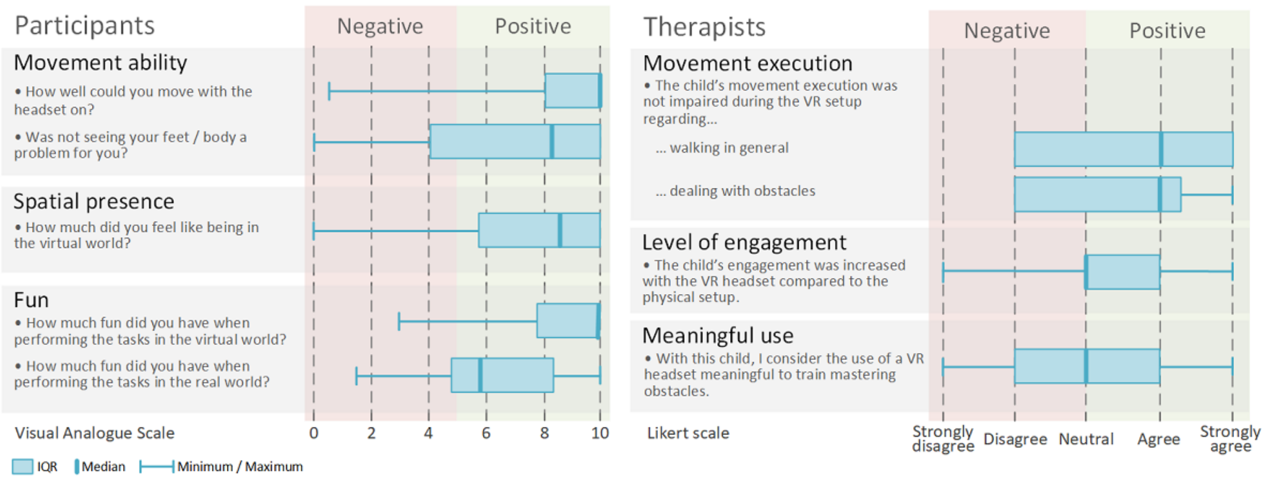
Figure 5. Noninferiority analysis for the primary outcomes. A negative mean difference indicates a smaller value in the virtual reality (VR) setup than in the physical setup. The noninferiority margins in blue represent the maximum difference between the two conditions while still considering the conditions to be equal. As long as the 95% CI of the mean difference does not exceed this margin, the VR setup is noninferior to the physical setup. Inferiority of the VR setup is assumed when the 95% CI touches the red inferiority area and, at the same time, does not cross the line of no difference between the two conditions.



Questionnaires

When asked how well the participants could move around using the HMD, 75% (15/20) of the participants scored ≥8 and did not feel restricted in their ability to move around. Not being able to see their body or feet was no problem (score ≥7) for most (14/20, 70%; Figure 6) participants. Most participants (14/20, 70%) felt physically present in the virtual scenario (score ≥7), even if the environment and the objects did not seem entirely realistic to them. The participants had fun in both conditions; however, the VR setup was rated significantly better (P<.001).

Figure 6. The participants' and therapists' views on the use of the virtual reality (VR) head mounted display (HMD) in physiotherapy.



According to the therapists, movement execution during the VR setup was not impaired in 65% (13/20) of the participants when walking normally or dealing with obstacles. The most common reason why therapists considered mild impairment in movement execution while wearing the HMD was a more cautious and slower gait pattern. The therapists perceived the level of engagement in the VR setup to be lower in 4 participants, similar in 7 participants, and higher in 9 participants. Therapists had ambivalent views regarding the meaningfulness of using VR to train for mastering obstacles. Reasons for considering the application meaningful included increased enthusiasm for movement, the challenge of altered visual control, and, therefore, the increased awareness of the children's bodies. Reduced speed, reduced focus on the given instructions, lack of feeling the edges of the obstacles, and consequences, such as stumbling when not lifting the foot high enough, were reasons against VR being a meaningful application for some participants. The balancing task was the task most often considered meaningful by the therapists.

Discussion

Principal Findings

This study aimed to provide information on whether a VR setup is feasible and motivating to induce and practice movements that are needed to master real obstacles in children and adolescents with gait disorders. Furthermore, this study aimed to evaluate which kinds of everyday walking activities are appropriate to be practiced in such a VR setup. To achieve these goals, a virtual and a physical condition, the latter resembling therapeutic setup, were compared with each other. We chose a noninferiority analysis to quantify the differences between spatiotemporal parameters defined a priori. This noninferiority analysis revealed that for 3 of the 4 walking tasks encountered in everyday life, mastering the virtual obstacles provided by an HMD was noninferior to mastering the physical obstacles. Thus, the results suggest that children and adolescents with gait disorders can practice crossing a wide gap, balancing on a narrow area, and circumventing stationary obstacles appropriately in a virtual setup. Furthermore, overstepping a virtual obstacle with the leading foot was also noninferior; only the results for the trailing foot were inconclusive.

Comparison to Previous Work

Normal Walking

The participants walked slower in the VR compared to the physical setup, which corresponds to the findings of Almajid et al [38] and Horsak et al [39]. Almajid et al [38] found that younger and older adults needed significantly more time to perform the timed up-and-go test when wearing an HMD, even without the projection of a virtual scene. In the study of Horsak et al [39], healthy individuals also demonstrated a slower walking pattern when walking in an overground VR environment compared to a real environment. This effect should be considered when wearing an HMD in rehabilitation. Still, the mean gait speed during normal walking in the VR setup was within the range of self-selected walking speed in children aged between 7 and 14 years with CP (GMFCS levels I and II) and TBI [25-27]. Although the participants' FMS and FAQ values were in the upper range of the scales, their gait speed in the physical setup was still below the average of typically developing youths [40]. The reduced gait speed in the VR setup was accompanied by a decreased step length and a slightly prolonged double stance phase.

The mean step length in both conditions was above the average step length of 50 cm reported for children aged between 7 and 14 years with CP (GMFCS stages I and II) and TBI, but below the average step length of 68 cm reported for typically developing children of the same age [25-27]. The double-stance phase of our participants during normal walking was in both conditions remarkably longer than in typically developing youths aged between 5 and 21 years [40]. Several therapists observed that the movements of their patients were constrained at the beginning of wearing the HMD, especially during normal walking. This could be because more than half of the participants had never worn a VR HMD before participating in this study. However, the difference in double stance time between the VR and physical setup was considerably smaller compared to the difference in double stance time between typically developing individuals and individuals with blindness [41].

Overstepping

First, the noninferiority analysis demonstrated that the maximal step height of the leading foot when stepping over the virtual obstacle was noninferior compared to the physical obstacle. This indicated that participants raised their leading foot to the same height when overstepping the virtual obstacle. However, they lifted their trailing foot considerably less high in the VR condition. This finding is supported by a study by Hagio and Kouzaki [42], in which healthy adults overstepped a virtual and physical obstacle. While the vertical height of the leading foot correlated highly (r=0.77) between the VR and physical condition, the correlation was lower for the trailing foot (r=0.47). As Kim et al [43] describe, an explanation for the difference between the leading and trailing foot in the VR setup could be the missing visual information regarding the height of the foot and, therefore, not being able to correct its height. Further results from Hagio and Kouzaki [42] suggest that visuomotor transformation in the leading leg contributes to a motor plan for trailing limb toe trajectories while stepping over an obstacle.

Crossing

Second, although the primary outcome parameters were mostly comparable between the virtual and physical setups, the movement was slightly displaced when overstepping or crossing the obstacles in the VR setup. Participants stepped too close to the obstacle or even over the edge of the obstacle. In general, however, the steps were almost the same length and height in the VR and physical setups, just at different locations. As the HMD blocks out the physical world, a lack of spatial information about the environment and the body's state relative to the environment could be a reason for the slightly displaced movement execution in the VR condition. However, most participants indicated on the VAS that not seeing their feet or body was not a problem for them. Furthermore, almost half of the participants scored 100% on the FM assessment, which tests the proprioception of the lower extremities. Nevertheless, using a fully immersive VR, Kim et al [43] investigated how visual information about the lower extremities is integrated with information about the environment to facilitate successful obstacle avoidance in healthy young adults. Their study revealed that visual information about the lower extremities promoted more consistent behavior while stepping over an obstacle.

Balancing

Third, in both conditions of the balancing task, the step length was slightly decreased, and the double stance phase increased compared to the corresponding normal walking condition. As reduced step length and prolonged double stance phase are considered indicators of reduced balance [41], we can assume that the participants made a real effort to balance over the physical and virtual obstacles. Although the participants rated this task as rather difficult, they produced only a small number of failures. The step width, which we considered crucial for successfully completing the balance task, was, on average, 1 cm larger in the VR than in the physical setup. However, the noninferiority analysis illustrated that the step width in the VR setup was noninferior to that of the physical condition. Therefore, we assume that the participants successfully balanced over the obstacle in VR and in reality.

Circumventing

Fourth, when moving in public areas, it becomes essential to circumvent stable objects or moving people, have a stable base of support, and balance in a narrow space. Several studies have investigated the critical point (the ratio between aperture width and shoulder width at which a shoulder rotation occurs at the time of crossing) and safety margin (the space that is maintained between the shoulders and the obstacles at the time of crossing) for aperture crossing [30,44]. Whenever the participants had to rotate their shoulders, they maintained a larger safety margin when crossing [44]. For example, the critical point for circumventing poles, calculated from the mean shoulder width of the participants and the distance between the poles, was a ratio of 1.3 [44]. The present study's ratio between the aperture width and the mean shoulder width equaled 1.6. Assuming that participants did not rotate their shoulders at such a ratio, the safety margin was slightly less than the 30 cm observed in the study of Hackney et al [30]. However, the safety margins of 10 cm of the VR and physical setup equal those of young, healthy adults who had to avoid poles with an aperture/shoulder width ratio of 1.3 [44]. The results of the noninferiority analysis suggest that participants successfully circumvented the obstacles in the VR setup. In addition, Hackney et al [45] recently showed that individuals who had to avoid obstacles in a virtual scenario wearing an HMD behave similarly with virtual poles and avatars, indicating generalization to a wide range of applications in VR.

Questionnaires

In summary, the participants were very positive toward training walking tasks in a VR setting. Due to its game-like features, the participants experienced significantly more fun in the VR than in the physical setup. How VR-assisted physical therapy might affect a participant's enjoyment and motivation over time needs to be investigated in the future. The physiotherapists did not observe a difference in the participants' engagement level between the VR and physical setup, indicating that the participants made similar efforts in both conditions. Thus, a comparison between the 2 conditions was feasible.

Limitations

This study has several limitations. First, the group size of 20 participants was rather small. However, it is in line with recommendations [22], as the purpose of this study was to provide information on whether a VR setup is feasible and motivating to induce movements that are needed to master real obstacles and which kinds of everyday walking activities are appropriate to be practiced in such a VR setup. To examine the appropriateness and effectiveness of VR training, more participants would have to be included in the next study. Despite the considerable heterogeneity of this study, noninferiority could still be shown in 3 tasks.

Second, even though the dimensions and locations of the obstacles did match in both conditions, the different visualizations of the physical setup and the VR setup could have impacted the participants' gait. However, this limitation was chosen intentionally, as we wanted the obstacles to look like they would appear in future applications.

Third, a panel of experts decided on specific margins to define noninferiority, as no reliable reference values for the noninferiority analyses existed in the literature. In order to minimize this limitation for a further project, additional external experts could be asked and added to the panel.

Fourth, the gait laboratory is frequently used for clinical gait analysis. Therefore, the Vicon cameras pointed to the middle of the room. Since the recording area for this study was slightly broader, some markers disappeared at times from the measurement volume, which is one reason why some participants had to complete more than 8 rounds to record sufficient valid trials. Consequently, the high number of repetitions might have bored and fatigued some participants, which might have decreased their concentration toward the end. With verbal input for the participants and breaks between the trials if needed, we tried to keep the number of trials and the fatigue of the participants as low as possible.

Fifth, a slight misalignment between the real and virtual setups might have introduced an unknown error in calculating the parameters. We calibrated the alignment immediately before putting the HMD on the participant's head to minimize this error.

Sixth, the feet were not visible to the participants in the VR condition. We assume that a lack of spatial information rather than impairments in proprioception might have caused failures such as stepping over the edge, as the FM assessment did not indicate major lower limb proprioception impairments in the participants. A further study investigating the influence of foot projection in VR could provide further information regarding the influence of the visability of the feet.

Conclusions

This is the first study showing that children and adolescents with gait disorders master various obstacle tasks, such as overstepping a bar, crossing a wide gap, balancing on a narrow area, and circumventing stationary obstacles, similarly in VR and physical conditions. Only the results for the trailing foot in the overstepping task were inconclusive. Therefore, we conclude that using a VR setup to practice mastering obstacles with children and adolescents with gait disorders is feasible and motivates them to practice everyday walking tasks. In the long run, the feasibility of using HMDs in a clinical therapy setting, patient motivation over a longer period of time, the appropriateness and effectiveness of such VR interventions, and identifying potential responders to such interventions require further investigations.

Acknowledgments

The authors would like to thank the participants and their physiotherapists for the time and effort they have dedicated to this project. The project was financially supported by the J&K Wonderland Foundation (Steinhausen, Switzerland) and the Promedica Foundation (Chur, Switzerland).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors contributed to the conceptualization and methodology of the study. CA-R obtained ethical approval. SR performed participant recruitment. SR, AK, and CA-R assisted with the measurements. SR and AK were involved in the data analysis. SR provided the figures. SR and CA-R were responsible for writing the first draft. All authors critically reviewed the manuscript.

Conflicts of Interest

None declared.

References

1. Katz-Leurer M, Rotem H, Keren O, Meyer S. Balance abilities and gait characteristics in post-traumatic brain injury, cerebral palsy and typically developed children. Dev Neurorehabil 2009;12(2):100-105. [doi: 10.1080/17518420902800928] [Medline: 19340662]
2. Frisch D, Msall ME. Health, functioning, and participation of adolescents and adults with cerebral palsy: a review of outcomes research. Dev Disabil Res Rev 2013;18(1):84-94. [doi: 10.1002/ddrr.1131] [Medline: 23949832]
3. Rast FM, Labruyère R. ICF mobility and self-care goals of children in inpatient rehabilitation. Dev Med Child Neurol 2020;62(4):483-488 [FREE Full text] [doi: 10.1111/dmcn.14471] [Medline: 31984500]
4. Rosenbaum P, Gorter JW. The 'F-words' in childhood disability: I swear this is how we should think!. Child Care Health Dev 2012;38(4):457-463 [FREE Full text] [doi: 10.1111/j.1365-2214.2011.01338.x] [Medline: 22040377]
5. Cipresso P, Giglioli IAC, Raya MA, Riva G. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. Front Psychol 2018;9:2086 [FREE Full text] [doi: 10.3389/fpsyg.2018.02086] [Medline: 30459681]
6. Bailey JO, Bailenson JN. Chapter 9—Immersive virtual reality and the developing child. In: Cognitive Development in Digital Contexts. Cambridge, MA: Academic Press; 2017:181-200.
7. Bainbridge WS. The scientific research potential of virtual worlds. Science 2007;317(5837):472-476. [doi: 10.1126/science.1146930] [Medline: 17656715]
8. Perez-Marcos D. Virtual reality experiences, embodiment, videogames and their dimensions in neurorehabilitation. J Neuroeng Rehabil 2018;15(1):113 [FREE Full text] [doi: 10.1186/s12984-018-0461-0] [Medline: 30477527]
9. Levac D, Rivard L, Missiuna C. Defining the active ingredients of interactive computer play interventions for children with neuromotor impairments: a scoping review. Res Dev Disabil 2012;33(1):214-223. [doi: 10.1016/j.ridd.2011.09.007] [Medline: 22093667]
10. Granic I, Lobel A, Engels RCME. The benefits of playing video games. Am Psychol 2014;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
11. Sharar SR, Carrougher GJ, Nakamura D, Hoffman HG, Blough DK, Patterson DR. Factors influencing the efficacy of virtual reality distraction analgesia during postburn physical therapy: preliminary results from 3 ongoing studies. Arch Phys Med Rehabil 2007;88(12 Suppl 2):S43-S49. [doi: 10.1016/j.apmr.2007.09.004] [Medline: 18036981]
12. Morrongiello BA, Corbett M, Beer J, Koutsoulianos S. A pilot randomized controlled trial testing the effectiveness of a pedestrian training program that teaches children where and how to cross the street safely. J Pediatr Psychol 2018;43(10):1147-1159 [FREE Full text] [doi: 10.1093/jpepsy/jsy056] [Medline: 30113643]
13. Adamo-Villani N, Wilbur R, Wasburn M. Gender differences in usability and enjoyment of VR educational games: a study of SMILE™. 2008 Presented at: 2008 International Conference Visualisation; July 09-11, 2008; London, UK p. 114-119. [doi: 10.1109/vis.2008.10]
14. Newbutt N, Bradley R, Conley I. Using virtual reality head-mounted displays in schools with autistic children: views, experiences, and future directions. Cyberpsychol Behav Soc Netw 2020;23(1):23-33. [doi: 10.1089/cyber.2019.0206] [Medline: 31502866]
15. Parsons S. Learning to work together: designing a multi-user virtual reality game for social collaboration and perspective-taking for children with autism. Int J Child Comput Interact 2015;6:28-38. [doi: 10.1016/j.ijcci.2015.12.002]

16. Shema-Shiratzky S, Brozgol M, Cornejo-Thumm P, Geva-Dayan K, Rotstein M, Leitner Y, et al. Virtual reality training to enhance behavior and cognitive function among children with attention-deficit/hyperactivity disorder: brief report. Dev Neurorehabil 2019;22(6):431-436. [doi: 10.1080/17518423.2018.1476602] [Medline: 29771624]
17. Mon-Williams M. Is virtual reality bad for our health? The risks and opportunities of a technology revolution. Medium. 2017. URL: https://medium.com/university-of-leeds/is-virtual-reality-bad-for-our-health-the-risks-and-opportunities-of-a-technology-revolution-31520e50820a [accessed 2021-02-18]
18. Tychsen L, Foeller P. Effects of immersive virtual reality headset viewing on young children: visuomotor function, postural stability, and motion sickness. Am J Ophthalmol 2020;209:151-159. [doi: 10.1016/j.ajo.2019.07.020] [Medline: 31377280]
19. Ravi DK, Kumar N, Singhi P. Effectiveness of virtual reality rehabilitation for children and adolescents with cerebral palsy: an updated evidence-based systematic review. Physiotherapy 2017;103(3):245-258. [doi: 10.1016/j.physio.2016.08.004] [Medline: 28109566]
20. Chen Y, Fanchiang HD, Howard A. Effectiveness of virtual reality in children with cerebral palsy: a systematic review and meta-analysis of randomized controlled trials. Phys Ther 2018;98(1):63-77 [FREE Full text] [doi: 10.1093/ptj/pzx107] [Medline: 29088476]
21. Ammann-Reiffer C, Kläy A, Keller U. Virtual reality as a therapy tool for walking activities in pediatric neurorehabilitation: usability and user experience evaluation. JMIR Serious Games 2022;10(3):e38509 [FREE Full text] [doi: 10.2196/38509] [Medline: 35834316]
22. Macefield R. How to specify the participants groups size for usability studies: a practitioner's guide. J Usability Stud 2009;5(1):34-45 [FREE Full text]
23. Ammann-Reiffer C, Bastiaenen CHG, Klöti C, van Hedel HJA. Concurrent validity of two gait performance measures in children with neuromotor disorders. Phys Occup Ther Pediatr 2019;39(2):181-192. [doi: 10.1080/01942638.2017.1420003] [Medline: 29336642]
24. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehabil Med 1975;7(1):13-31. [Medline: 1135616]
25. Fosdahl MA, Jahnsen R, Kvalheim K, Holm I. Medicina (Kaunas) 2019;55(6):250 [FREE Full text] [doi: 10.3390/medicina55060250] [Medline: 31174397]
26. Katz-Leurer M, Rotem H, Keren O, Meyer S. The relationship between step variability, muscle strength and functional walking performance in children with post-traumatic brain injury. Gait Posture 2009;29(1):154-157. [doi: 10.1016/j.gaitpost.2008.06.013] [Medline: 18701292]
27. Kimoto M, Okada K, Sakamoto H, Kondou T. The association between the maximum step length test and the walking efficiency in children with cerebral palsy. J Phys Ther Sci 2017;29(5):822-827 [FREE Full text] [doi: 10.1589/jpts.2017.822] [Medline: 28603353]
28. Weiten J. Auxiologische Untersuchung bei Kindern mit Wachstumshormonmangel vor und unter der Substitutionstherapie mit Wachstumshormon. Humboldt-Universität zu Berlin. [Doctoral thesis]. 2004. URL: https://edoc.hu-berlin.de/handle/18452/15703 [accessed 2024-02-27]
29. Spranger J, Ochsenfarth A, Kock HP, Henke J. Anthropometrische Normdaten im Kindesalter [Anthropometric standards in childhood]. Z Kinderheilkd 1968;103(1):1-12. [Medline: 5708736]
30. Hackney AL, Vallis LA, Cinelli ME. Action strategies of individuals during aperture crossing in nonconfined space. Q J Exp Psychol (Hove) 2013;66(6):1104-1112. [doi: 10.1080/17470218.2012.730532] [Medline: 23718874]
31. Redekop S, Andrysek J, Wright V. Single-session reliability of discrete gait parameters in ambulatory children with cerebral palsy based on GMFCS level. Gait Posture 2008;28(4):627-633. [doi: 10.1016/j.gaitpost.2008.04.008] [Medline: 18514523]
32. Creem-Regehr SH, Gill DM, Pointon GD, Bodenheimer B, Stefanucci JK. Mind the gap: gap affordance judgments of children, teens, and adults in an immersive virtual environment. Front Robot AI 2019;6:96 [FREE Full text] [doi: 10.3389/frobt.2019.00096] [Medline: 33501111]
33. Cappellini G, Sylos-Labini F, MacLellan MJ, Assenza C, Libernini L, Morelli D, et al. Locomotor patterns during obstacle avoidance in children with cerebral palsy. J Neurophysiol 2020;124(2):574-590 [FREE Full text] [doi: 10.1152/jn.00163.2020] [Medline: 32667246]
34. Berard JR, Vallis LA. Characteristics of single and double obstacle avoidance strategies: a comparison between adults and children. Exp Brain Res 2006;175(1):21-31. [doi: 10.1007/s00221-006-0529-0] [Medline: 16761138]
35. Chang MC, Lee BJ, Joo NY, Park D. The parameters of gait analysis related to ambulatory and balance functions in hemiplegic stroke patients: a gait analysis study. BMC Neurol 2021;21(1):38 [FREE Full text] [doi: 10.1186/s12883-021-02072-4] [Medline: 33504334]
36. Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. Trials 2011;12:106 [FREE Full text] [doi: 10.1186/1745-6215-12-106] [Medline: 21539749]
37. Stefanos R, Graziella DA, Giovanni T. Methodological aspects of superiority, equivalence, and non-inferiority trials. Intern Emerg Med 2020;15(6):1085-1091. [doi: 10.1007/s11739-020-02450-9] [Medline: 32705494]
38. Almajid R, Tucker C, Keshner E, Vasudevan E, Wright WG. Effects of wearing a head-mounted display during a standard clinical test of dynamic balance. Gait Posture 2021;85:78-83. [doi: 10.1016/j.gaitpost.2021.01.020] [Medline: 33517040]

39. Horsak B, Simonlehner M, Schöffer L, Dumphart B, Jalaeefar A, Husinsky M. Overground walking in a fully immersive virtual reality: a comprehensive study on the effects on full-body walking biomechanics. Front Bioeng Biotechnol 2021;9:780314 [FREE Full text] [doi: 10.3389/fbioe.2021.780314] [Medline: 34957075]
40. Voss S, Joyce J, Biskis A, Parulekar M, Armijo N, Zampieri C, et al. Normative database of spatiotemporal gait parameters using inertial sensors in typically developing children and young adults. Gait Posture 2020;80:206-213 [FREE Full text] [doi: 10.1016/j.gaitpost.2020.05.010] [Medline: 32531757]
41. Hallemans A, Ortibus E, Truijen S, Meire F. Development of independent locomotion in children with a severe visual impairment. Res Dev Disabil 2011;32(6):2069-2074. [doi: 10.1016/j.ridd.2011.08.017] [Medline: 21985990]
42. Hagio S, Kouzaki M. Visuomotor transformation for the lead leg affects trail leg trajectories during visually guided crossing over a virtual obstacle in humans. Front Neurosci 2020;14:357 [FREE Full text] [doi: 10.3389/fnins.2020.00357] [Medline: 32390793]
43. Kim A, Kretch KS, Zhou Z, Finley JM. The quality of visual information about the lower extremities influences visuomotor coordination during virtual obstacle negotiation. J Neurophysiol 2018;120(2):839-847 [FREE Full text] [doi: 10.1152/jn.00931.2017] [Medline: 29742030]
44. Hackney AL, Cinelli ME, Frank JS. Does the passability of apertures change when walking through human versus pole obstacles? Acta Psychol (Amst) 2015;162:62-68. [doi: 10.1016/j.actpsy.2015.10.007] [Medline: 26529484]
45. Hackney AL, Cinelli ME, Warren WH, Frank JS. Are avatars treated like human obstacles during aperture crossing in virtual environments? Gait Posture 2020;80:74-76 [FREE Full text] [doi: 10.1016/j.gaitpost.2020.05.028] [Medline: 32492623]

Abbreviations

CP: cerebral palsy
FAQ: Functional Assessment Questionnaire
FM: Fugl-Meyer
FMS: Functional Mobility Scale
GMFCS: Gross Motor Function Classification System
HMD: head-mounted display
TBI: traumatic brain injury
VAS: visual analog scale
VR: virtual reality

Edited by G Eysenbach, T Leung; submitted 01.06.23; peer-reviewed by J Finley, B Sindelar, S Hwang; comments to author 23.07.23; revised version received 13.08.23; accepted 02.02.24; published 18.03.24.

Please cite as:
Rhiel S, Kläy A, Keller U, van Hedel HJA, Ammann-Reiffer C
Comparing Walking-Related Everyday Life Tasks of Children with Gait Disorders in a Virtual Reality Setup With a Physical Setup: Cross-Sectional Noninferiority Study
JMIR Serious Games 2024;12:e49550
URL: https://games.jmir.org/2024/1/e49550
doi:10.2196/49550
PMID:38498048

Original Paper

Addressing Data Absenteeism and Technology Chauvinism in the Use of Gamified Wearable Gloves Among Older Adults: Moderated Usability Study

Edmund W J Lee[1,2], PhD; Warrick W Tan[1]; Ben Tan Phat Pham[3], MBA; Ariffin Kawaja[4,5], PhD; Yin-Leng Theng[1,3], PhD

[1]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore, Singapore

[2]Centre for Information Integrity and the Internet, Nanyang Technological University, Singapore, Singapore, Singapore

[3]Ageing Research Institute for Society and Education, Nanyang Technological University, Singapore, Singapore, Singapore

[4]StretchSkin Technologies Pte Ltd, Singapore, Singapore

[5]SingHealth Polyclinics, Singapore, Singapore

Corresponding Author:
Edmund W J Lee, PhD
Wee Kim Wee School of Communication and Information
Nanyang Technological University, Singapore
31 Nanyang Link
WKWSCI Building, #02-06, Singapore 639798
Singapore, 639798
Singapore
Phone: 65 6908 3429
Email: elwj88@gmail.com

Abstract

Background: Digital health technologies have the potential to improve health outcomes for older adults, especially for those recovering from stroke. However, there are challenges to developing these technologies, such as *data absenteeism* (where older adults' views are often underrepresented in research and development) and *technology chauvinism* (the belief that sophisticated technology alone is the panacea to addressing health problems), which hinder their effectiveness.

Objective: In this study, we aimed to address these challenges by developing a wearable glove integrated with culturally relevant exergames to motivate older adults to exercise and, for those recovering from stroke, to adhere to rehabilitation.

Methods: We conducted a moderated usability study with 19 older adults, of which 11 (58%) had a history of stroke. Our participants engaged in a 30-minute gameplay session with the wearable glove integrated with exergames, followed by a quantitative survey and an in-depth interview. We used descriptive analysis to compare responses to the System Usability Scale between those who had a history of stroke and those who did not. In addition, we analyzed the qualitative interviews using a bottom-up thematic analysis to identify key themes related to the motivations and barriers regarding the use of wearable gloves for rehabilitation and exercise.

Results: Our study generated several key insights. First, making the exergames exciting and challenging could improve exercise and rehabilitation motivation, but it could also have a boomerang effect, where participants may become demotivated if the games were very challenging. Second, the comfort and ease of use of the wearable gloves were important for older adults, regardless of their stroke history. Third, for older adults with a history of stroke, the functionality and purpose of the wearable glove were important in helping them with specific exercise movements.

Conclusions: Our findings highlight the importance of providing contextual support for the effective use of digital technologies, particularly for older adults recovering from stroke. In addition to technology and usability factors, other contextual factors such as gamification and social support (from occupational therapists or caregivers) should be considered to provide a comprehensive approach to addressing health problems. To overcome data absenteeism and technology chauvinism, it is important to develop digital health technologies that are tailored to the needs of underserved communities. Our study provides valuable insights for the development of digital health technologies that can motivate older adults recovering from stroke to exercise and adhere to rehabilitation.

(JMIR Serious Games 2024;12:e47600) doi:10.2196/47600

KEYWORDS
wearables; exergames; older adults; active aging; rehabilitation; stroke

Introduction

Background

There has been an increasing trend in the development and implementation of digital health technologies for older adults, known as *gerontechnologies*. However, there are challenges to developing these technologies, such as *data absenteeism* (where older adults' views are often underrepresented in research and development) and *technology chauvinism* (the belief that sophisticated technology alone is the panacea to addressing health problems) [1]. Gerontechnologies aim to assist older adults in healthy aging through the promotion of physical exercise and empowering them to maintain a certain level of functional independence throughout their later years and delay the onset of frailty [2,3]. For instance, examples of such technologies are the use of *exergames* to motivate physical activity or virtual reality technology for the purposes of pain management and therapy [4,5]. Other types of health technologies, such as wearables, are also becoming common among older adults. Existing commercial products such as FitBit (Google, Inc) and Apple Watch make it easy for older adults to track and monitor their health [6]. To date, such digital health technologies targeted at older adults have received substantial attention and investments, ranging from US $1.1 billion to US $3.1 billion between 2019 and 2020, and it is expected to grow in the next few years [7,8].

While there are several studies documenting the positive impact of digital health technologies on older adults' physical and mental health, there are 2 key gaps in existing public health literature [9,10]. First, very few studies have explicitly addressed the problems of *data absenteeism* and *technology chauvinism*. Second, as many of the existing health technologies are developed in a clinical setting for data collection, few researchers have examined how to develop technology that is fun, relatable, and equitable for the older adults [11]. The overall objective of the study was to address the problem of data absenteeism and technology chauvinism in the development of digital health technologies—a stretchable wearable glove integrated with exergames—by examining the motivations and barriers of older adults without a history of stroke and those recovering from stroke in the use of such technologies through the lens of equity.

Theoretical Framework: Data Absenteeism and Technology Chauvinism

While the momentum to build and develop different digital health technologies by academia or industry is noteworthy, it is important to adopt the lens of equity in building such technologies for older adults. Older adults are often disadvantaged as compared to the general and younger population in terms of their access, use, attention, and processing of health information from digital health technologies [12]. Not giving attention to the context in which such digital health technologies are developed and introduced may result in the unintentional consequence of exacerbating the health disparities between those who are well resourced and those with less resources [13]. In advocating for technology and big data use through an equitable lens, Lee and Viswanath [1] argued for the need for health communication and informatics scholars and technology developers to pay attention to the 2 perennial problems: data absenteeism and technology chauvinism.

Data absenteeism refers to a situation where data from underserved populations are not represented. For example, recent studies of large-scale national programs that use wearables to boost physical activity, such as Singapore's National Steps Challenge, which offers free wearables and gamified mobile apps to encourage participants to walk 10,000 steps, show that older adults are often underrepresented in the data. This gap aligns with previous findings indicating that users of such technology tend to be younger and more educated and possess a higher degree of eHealth literacy [14,15].

In contrast, *technology chauvinism*, refers to the blind faith in big data systems or technology platforms in addressing health disparities. One of the most infamous cases was the use of Google Flu Trends to predict the outbreak of influenza, where search trends overestimated the prevalence of influenza as compared to official sources [16,17]. To address the problems of data absenteeism and technology chauvinism in the use of digital health technologies to improve physical activity among healthy older adults and those recovering from stroke, it is vital to involve them from the start of the research process and engage them to codevelop wearable gloves and exergames. This is consistent with studies of the principles of user-centered design, where it is crucial to be intentionally inclusive in the process by incorporating the target group as full partners in the decision and design groups as part of the research process. This ensures that their participation and input are not merely symbolic or exploitative, but rather meaningful and beneficial for them as a whole [18].

Context of the Study: Motivating Rehabilitation and Exercise Using Wearable Gloves and Exergames

In recent years, there has been a surge in the exploration and creation of wearable glove technologies aimed at enhancing exercise motivation and aiding stroke recovery. The smart glove industry is projected to reach a value of US $3.9 billion by 2028, with an annual growth rate of 10% [19]. These innovative gloves are developed in 2 main styles: rigid hand exoskeletons and soft assistive gloves. The latest soft rehabilitation gloves are being designed to support bending, straightening, and spreading or closing of each finger to address the difficulties some patients face in performing hand grabbing motion [20].

There are several existing studies that have documented the efficacy of the use of wearable gloves for patients with stroke in improving upper limb movement across several metrics. For instance, Yurkewich et al [21] tested the Hand Extension Robot

Orthosis (HERO) Grip Glove among 11 participants with difficulty in finger extension in their poststroke journey and found that the glove significantly improved their water bottle grasp and index finger movement and extension and enabled individuals who are lacking grip strength to handle blocks, use a fork, and write with a pen. Wang et al [22] conducted a study where 69 patients with severe upper limb impairment following a stroke were divided into three treatment groups: (1) repetitive transcranial magnetic stimulation, (2) soft robotic glove use, and (3) standard treatment, and the results showed that the group assigned to use the robotic gloves achieved better upper extremity scores compared to those in the standard treatment group.

Given that the use of wearable gloves is comparable to neurostimulation or neuromodulation techniques in improving movements for patients with stroke, other research teams have incorporated the use of other forms of stimulation (ie, tactile sensations) in wearable gloves to improve its efficacy. Seim et al [23] conducted a pilot study involving 16 patients with chronic stroke, where participants were randomly placed into 2 groups for an 8-week period: one group received a vibrotactile stimulation glove, and the other group received a similar glove without vibration (acting as the control condition). The outcomes demonstrated that those using the vibrotactile stimulation glove experienced notable enhancements in finger mobility and improvements in the range of motion of their elbows and shoulders compared to the control group.

While these studies have found improvements in upper limb movements through the use of wearable gloves in their respective sample, it was unclear why or how they were effective. In a systematic review of wearable technology for improving activity in adult patients with stroke, Parker et al [24] found that, overall, very few studies have found evidence for the use of wearable gloves to improve rehabilitation. Thus, a significant oversight in many recent advancements in these state-of-the-art wearable gloves is the lack of consideration for how their target user group might engage with these technologies and the probability of their adoption as consumer products. Examples of these gaps in knowledge would be ideas about what could motivate a potential consumer to purchase the wearable glove product and incorporate it into their daily rehabilitation routine, and ultimately, what could be the barriers that prevent them from making the gloves a staple part of their rehabilitation journey. Examining how specific design features affect the perceived usefulness and perceived ease of use of the product could also affect their decision to adopt these fancy wearable gloves into their routines [25].

Addressing the Gaps in Wearable Gloves Research: Integrating Wearable Gloves With Exergames

One of the ways to improve rehabilitation through wearable gloves among older adults is through the gamification of the rehabilitation process. This is done through the use of specialized video games that simulate exercise, known colloquially as *exergames*. It is an area of health technology that has gained interest in recent years, especially among older adults. According to Harrington et al [26], the use of exergames may address barriers to older adults being physically active and exercising. These video games demand that the player physically moves their body to advance within the game or program [27]. Therefore, the implications surrounding the movement-centric nature of the gaming technology means that it has the potential to be used in rehabilitative health care practices also.

Existing systems such as the Nintendo Wii Fit, a low-cost commercial gaming system, have been found to be effective in improving clinical measures of balance in older adult patients [28,29]. Thus, exergames are able to transform exercise, through the process of gamification, by introducing alternative motivators such as entertainment and encouragement into activities that would otherwise be considered as physically strenuous [30]. A study by Yu et al [31] found that exergames as an intervention using the Xbox Kinect significantly improved the physical activity level, leg strength, and cardiopulmonary endurance of healthy older adults. Although exergames make physical exercise and rehabilitation more accessible, there are still obstacles to the uptake of these innovations in health technology among older adults.

One key aspect in which exergames have been shown to aid the rehabilitation process is adherence to exercise. Research by Oesch et al [32] on the difference between conventional self-regulated exercise and exergames revealed that exergames showed heightened levels of adherence to rehabilitation exercise routines within the first 2 weeks of introduction. However, the same study also showed that the motivation levels reversed after the first 2 weeks.

Understanding the Motivations and Barriers

One of the main goals of using a gamified wearable glove is to further motivate patients and older adults who are undergoing hand rehabilitation by increasing the movement of their hands through gameplay using wearable gloves and, subsequently, reduce the barriers to use of technology through seamless incorporation of the gaming medium. In addition, the hand movement data that the gloves are capable of collecting can further assist health care professionals in providing focused rehabilitation activities to their individual patients.

As such products are aimed to be used by patients and older adults, one must also consider motivations and barriers to adoption. This is because new technologies (ie, the wearable gloves) may present potential problems from the perspective of the consumers (older adults and health care professionals), such as being incompatible with existing products or technologies and individuals' needs [33]. In this regard, health care technology has never really seen the introduction of gloves with the technological capability of collecting vast amounts of data from each exergaming session. Traditionally known as "haptic gloves," these wearable devices offer force feedback and originated in the virtual reality gaming industry. They have the potential for gathering data for video game technology developers, although their application in rehabilitation and health care is quite rare [34]. Thus, to understand older adults' needs, it is paramount to investigate their motivations and barriers regarding the use of wearable gloves and exergames and whether their past experiences with rehabilitation have any effect on their perception of these products.

Study Objectives

This study aims to answer the following research questions (RQs) regarding the general usability of these wearable gloves in the context of the rehabilitation process:

- RQ1: What are the motivations and barriers toward the use of an integrated wearable glove system with exergames?
- RQ2: What are the differences between older adults without a history of stroke and older adults with a history of stroke regarding their perceptions about an integrated wearable glove system with exergames?

Methods

Wearable Glove and Exergame Development

Before the study, the research team worked with a start-up company that specializes in the development of rehabilitation gloves to develop a wearable data glove that could be integrated with exergames. Figure 1 shows the glove prototype that was developed for both survivors of stroke and older adults, specifically to capture hand motion for gesture recognition and for data visualization with the help of the multiple sensors integrated within the gloves. The bendable and flexible sensors embedded in the glove cover the joints on each individual finger and capture bending data signals from finger movement. This technology supports the training of specific wrist, hand, or finger movement and finger joints mobilization activities that is commonly found in the rehabilitation process, as seen in Figure 2.

The gloves were designed to be integrated with existing in-house exergames developed by the research team at the university. The exergame system consists of exercises and games that could be personalized for each individual user to suit different durations and number of repetitions for a particular exercise [35,36]. These exergames were specifically created to suit older adults to promote successful aging by motivating exercise, where they aim to improve physical and cognitive functional capacity through easy-to-follow actions and interface and culturally relevant game themes [37].

While there were a series of 9 exergames developed to date, the research team integrated the wearable glove with 1 exergame called “Chinatown Race” (Figure 3), where players are required to dodge barriers while aiming to collect coins to score points while running down a road in an online Chinatown setting. The avatar is controlled through rotation of hand and finger movements that would be detected by the wearable glove. As shown in Figure 2, participants had to perform movement 5 and movement 8 to slide the avatar toward left and right, respectively. To catch the lantern power-ups in the game, participants had to flex their pointer finger, as seen in movement 2.

Figure 1. Prototype of the wearable glove.

Figure 2. Specific hand motions for rehabilitation: (A) idle, (B) finger trigger, (C) up, (D) down, (E) right side, (F) right, (G) left, and (H) left side.

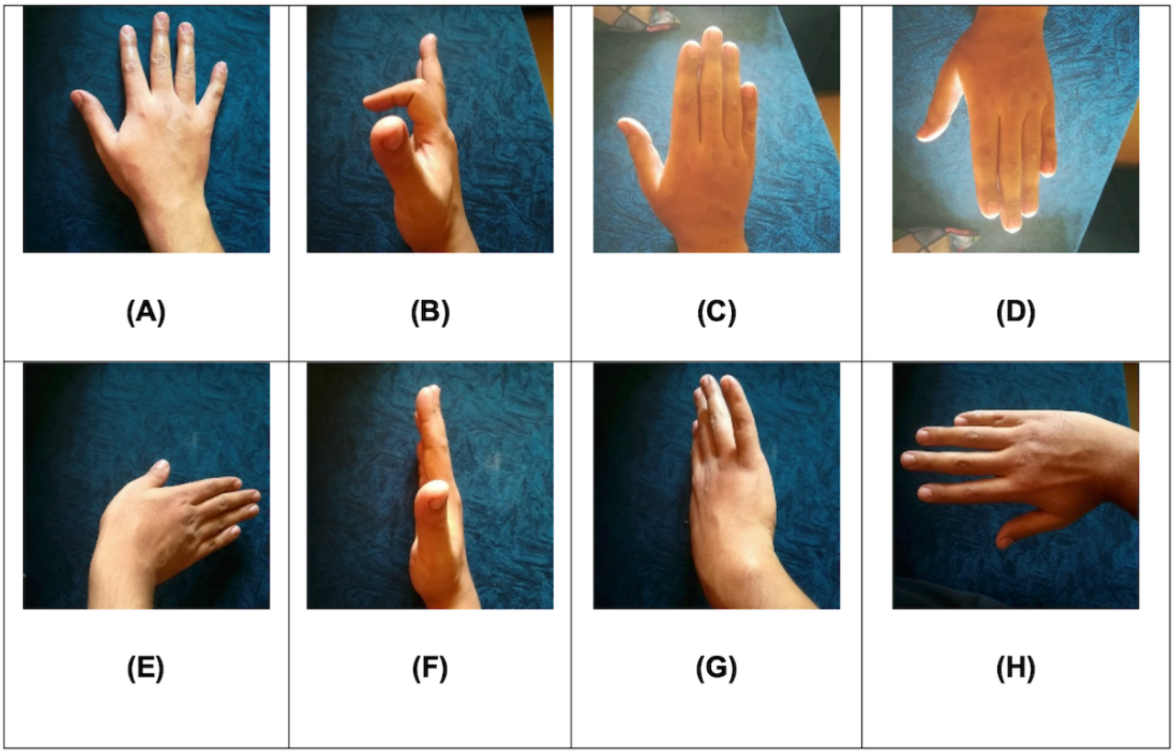


Figure 3. Screenshot of the Chinatown Race gameplay.

Participants and Recruitment

After the development of the wearable glove and integration with the exergame, we conducted a moderated usability study with 19 participants aged ≥50 years, recruited from an older adult activity center with a convenient sample of individuals with and without a history of stroke (n=11, 58% had a history of stroke and n=8, 42% did not). The average age of the participants was 66.8 (SD 4.5) years, and 53% (10/19) were women and 47% (9/19) were men. Of the 19 participants, 13 (68%) belonged to the Chinese ethnic group, while 4 (21%) belonged to the Malay ethnic group. There was also 5% (1/19) Burmese and 5% (1/19) Singh individuals who participated in the study. Overall, 74% (14/19) of the participants were considered to have lower levels of education below an “A-Level” certification. The inclusion criteria for the study were that participants had to be (1) aged at least 50 years and (2) willing and able to use the wearable glove to play the designated exergames and participate in a qualitative interview and survey. The moderated usability sessions were conducted in either English or Chinese depending on the language preference of the participants.

Ethical Considerations

Before the study, we obtained approval from the institutional review board of from the institutional review board of Nanyang Technological University, Singapore (IRB-2022-405). Written informed consent was also obtained from the participants. They were briefed about the procedures involved in the use of the wearable glove and exergames; they were also informed that the risks were minimal in the gameplay and that they could exit the study without any penalty. Upon successful completion of all the tasks in the moderated usability study, the participants were given a voucher worth SGD 50 (US $37) as incentive. To safeguard participants' privacy, the data were de-identified before analysis.

Study Design and Procedure

The moderated usability study session was designed to be a 1-hour long session, which consisted of a series of tasks involving participants' use of the wearable glove and exergame, led by 1 moderator and 1 technical specialist. The moderator guided the participants through the required activities and conducted the interview and survey upon the completion of the tasks. The technical specialist ensured that the devices were working as intended and handled all the technical difficulties that arose during the session.

In the session, participants were required to complete the following tasks:

1. Navigation of wearable glove and exergames: Participants were taught the basics about how to use the wearable gloves and control functions on the exergame dashboard.
2. Chinatown Race gameplay: Participants had to use the wearable glove to engage in Chinatown Race gameplay, where they had to move their avatars to avoid barriers and collect lanterns to score points by using hand and finger rotation.
3. Qualitative interview: This is a qualitative interview through which the moderator obtained feedback about participants' attitude and perceptions regarding the wearable glove and exergame.
4. System Usability Scale (SUS) survey: The moderator administered a short SUS scale, which is a 10-item measure of the usability of systems.

Measures of SUS

The SUS scale (Cronbach α=.95) was adapted from Chu et al [37], and participants provided their responses to the following items measured using a 5-point scale (1=strongly disagree and 5=strongly agree): (1) I think that I would like to use the glove frequently, (2) I found the glove unnecessarily complex, (3) I thought the glove was easy to use, (4) I think that I would need the support of a technical person to be able to use the glove, (5) I found the various functions of the glove to be well integrated with the game, (6) I thought there was too much inconsistency with the glove, (7) I would imagine that most people would learn to use the glove very quickly, (8) I found the glove very cumbersome to use, (9) I felt very confident using the glove, and (10) I needed to learn a lot of things before I could get going with the glove.

The qualitative interviews were analyzed using a bottom-up thematic analysis to identify key themes regarding the motivations and barriers toward the use of wearable gloves for rehabilitation and exercise among older adults who had a history of stroke and those who did not. The thematic analysis was performed in accordance with the steps described by Proulx et al [38] regarding usability testing of wearable gloves: (1) the recordings were transcribed, (2) the transcript was first read by a member of the research team to develop the coding frame, and (3) the coding frame was refined to identify different types of motivators and barriers regarding the use of wearable gloves and exergames. Next, descriptive analysis was conducted to compare the differences in responses to SUS items between older adults who had a history of stroke and those who did not. This is consistent with the studies by Tong et al [39] and Casterlé et al [40].

Results

Competition as a Motivation

RQ1 involves the motivations and barriers toward the use of a wearable glove integrated with exergames. Multimedia Appendix 1 shows the summary of the mean scores of participants' responses to SUS items. From the qualitative interview, one of the key motivations identified was that a certain *degree of competition* was required to ensure that the older adults are engaged and motivated. For instance, the inclusion of features such as collectible coins, lanterns, and a scoring system presented a form of incentive for the participants to continue persevering to complete the game. These features made the game fun, as there was a tangible goal associated with each movement that they made within the game. Some even noted that it spurred them to strive to get better at the game, as a participant mentioned the following:

> *To score the highest score, that is very exciting.* [Participant 10; with a history of stroke]

As such, the competitive element of the gameplay objectives plays a key role in motivating older adults to continue playing and, in essence, adhere better to their rehabilitation practices. However, they need to be familiar enough with the game systems and movements for them to attribute poor performance with their own lack of skill rather than an external device, which would otherwise deter them from continuing to play the game.

Helplessness as a Barrier

Regarding RQ1, we found that a significant barrier to using the integrated wearable glove is the feeling of helplessness, particularly if it is in the context of technical difficulties. While the participants mostly agreed that the exergame navigation was easy to understand because of the simple nature of the menu layout and user-friendly, large, and visible buttons, the older adults were not familiar with the movement system of the wearable gloves. Therefore, while they had very clear intentions regarding navigating the menu, the disconnect between intention and execution made the process a lot more difficult than intended for the participants. This became even more apparent when participants were asked to describe if anything was confusing about the gameplay movements of Chinatown Race, to which many expressed the following general sentiment:

> *Not confusing, only unfamiliar with the controls [in reference to glove].* [Participant 2; without a history of stroke]

Therefore, it was evident to a certain degree that the overall experience of playing the game was hindered because the gameplay affordances provided by the glove apparatus were not intuitive to the participants. This, in turn, had a demotivating effect on the participants as their ultimately poor gameplay would then be attributed to an external factor such as the glove and the game system, making them feel somewhat helpless as they struggled to competently dodge all the obstacles and collect the harder-to-collect lantern items that required them to hyperextend their thumb. This is exemplified when looking at "necessity of support" (ie, question 4 of the SUS scale), where the participants were asked to rate how much technical support they would require to use the gloves. The patients with a history of stroke were more likely to want to get more help (mean 2.45, SD 1.51) as compared to those who had no history of stroke (mean 3.45, SD 1.37). This could be attributed to how older adults with a history of stroke may have experienced feelings of helplessness using novel rehabilitation technologies in the past and therefore might show an aversion to new technologies.

RQ2 deals with the differences between older adults without a history of stroke and those with a history of stroke regarding their perceptions about the glove integrated with exergame. Our findings indicated that the rehabilitation history influenced the level of critique from user groups when engaging with the integrated glove. The comfort level of the glove was a major theme; in the SUS questions, participants who did not have a history of stroke indicated that comfort level was very important (mean 3.78, SD 0.97) as compared to those with a history of stroke (mean 4.18, SD 0.60).

When asked about what could be improved to make the overall experience of playing Chinatown Race using the wearable gloves better, there was a noticeable difference regarding the category of improvement suggested by both groups. The comments by participants with a history of stroke tended to pertain toward making the controls more intuitive:

> *Improve the controls, like grabbing action for the lantern instead of thumb extension.* [Participant 4; with a history of stroke]

In contrast, comments by participants who did not have a history of stroke were more likely related to areas such as game design and esthetics (collectible items or general color contrast):

> *Barrier needs to be bigger in size and have more contrasting colours so that it can be seen from the distance.* [Participant 19; without a history of stroke]

This difference could suggest that older adults with a history of stroke were more focused on the movements associated with the gloves and what their actual bodies were doing in relation to the game than whether the game was appealing enough to be played effectively. A possible attribution for this difference could be how the older adults with stroke have an acute awareness that the exergame is a tool to be used in a long and tedious process that they have experienced previously, and therefore, they have deeper awareness of what is essentially important in the actual process of rehabilitation. In contrast, the older adults without a history of stroke could be looking at the exergame and the wearable gloves as just another gaming device, and therefore, their criticism would be directed toward the exergames as a game than as a rehabilitation tool.

Finally, for RQ2, our results showed that the valence of anticipation regarding using a new device varies based on their history with stroke. This is seen when looking at "intention to frequently use," which refers to question 1 of the SUS scale used in the survey. It must be noted that the group without a history of stroke indicated more eagerness to use the glove with exergame (mean 4.38, SD 0.71) as compared to those with a history of stroke (mean 3.18, SD 1.47). It could be because participants who have had stroke before maintained a level of cynicism toward such novel methods for stroke rehabilitation as they have been through the rehabilitation process previously. Furthermore, it might be due to barriers related to perceptions of the ease of use. Participants without a history of stroke indicated that the glove was easy to use (mean 4, SD 1.50), and the score was slightly higher than those with a history of stroke (mean 3.64, SD 1.50). It might also be due to the visual esthetics, as there is a "box" attached to the glove that contains the wiring. A participant noted the following:

> *The box is big and is not needed...The extra electronic devices seem delicate and can break.* [Participant 9; with a history of stroke]

While the box itself is safely secured and presents no actual hindrance to the gameplay experience, the apparently excessive amount of gadgetry can scare older adults who are not familiar with the durability that most present-day technology have and, at worst, can incite a certain fear response within them to abstain from handling something they think can break any minute.

Discussion

Principal Findings

Several key findings were generated from this study. First, this study addresses the problem of data absenteeism and technology chauvinism by engaging older adults— those with and those without a history of stroke—to provide insights into the potential motivations and barriers regarding the use of integrated wearable glove with exergame solution for rehabilitation and exercise at the early stages of the development process. Most notably, while technologies such as wearable gloves and exergames play a pivotal role in the rehabilitation process, it is reductionistic to assume that the technology itself would be the panacea to addressing health issues. For instance, although older adults with and those without a history of stroke indicated that the glove and exergames were well integrated, we found that individuals who have a history of stroke were more likely to indicate that they would still require help in operating the wearable glove compared to those who have not experienced stroke, which was perceived to be complicated. This finding is supported by the "blind faith" aspect of technology chauvinism research: if technology is offered as a solution in isolation without paying attention to the larger social context of the participants, it would not be effective [13]. This is supported by our findings where older adults with a history of stroke indicated that they would need the support of a technical person to be able to use it effectively and that it could be unnecessarily complex. Thus, solutions that embrace the development and implementation of wearable gloves with exergames need to consider designing the technology such that it can be used in a community setting by tapping into the social and support networks of older adults as co-users to improve adherence and uptake. This is consistent with the study by Proulx et al [38], where they examined occupational therapists' perceived usability and utility of a similar wearable glove for rehabilitation. In their study, while the researchers found that the occupational therapists rated the usability of such gloves ranging from "moderate to good" on the SUS, they shared that the gloves would be challenging for patients if they did not have the assistance of a therapist, owing to the physical and cognitive deficits of patients with stroke. The therapists also suggested that the development of wearable gloves would need to account for different contexts, such as using the gloves with therapists or with the assistance of a caregiver.

Second, it was noted that although older adults indicated that the gloves were relatively easy to use and that they intend to use them frequently, the comfort level while using the glove is an important factor for older adults with a history of stroke. This is corroborated by existing studies of technology acceptance and usability, which suggests that the practicality of use is a fundamental cornerstone in the acceptance and integration of technology into a daily routine. Our finding is consistent with that of the study conducted by Yurkewich et al [21], where they designed the HERO Grip Glove to help patients with stroke to perform activities of daily living and finger movement. While participants reported that they were relatively satisfied with the glove in terms of safety, security, and general ease of use, they had the lowest satisfaction regarding the ease of wearing the glove.

Third, we found that it is crucial to consider how gameplay could motivate or demotivate older adults from using wearable gloves. Existing studies of commercial device-based hand rehabilitation for patients with stroke have shown that game-based training using wearable gloves was generally positive in improving hand function and that they would be received favorably as they would be perceived as entertaining [41]. However, our study showed that the design of gameplay would need to aim for a fine balance in managing the difficulty level for a diverse group of patients and players, such that players would find the game challenging enough to sustain their interest, but it would not be very complex to demotivate them. This is important because our results showed that some older adults could be easily demotivated by some of the gameplay scenarios, especially when they feel that they cannot achieve the objectives (ie, collecting lanterns and points and avoiding barriers) or when their perceived expectations of their personal performance do not match with their scores. This results in some of them being demotivated, not enjoying the gameplay, and feeling helpless as they felt that there was nothing they could do to improve their scores.

Finally, it could be observed that individuals with a history of stroke show a stronger aversion to using rehabilitative technology and would thus require more assistance and technical support. This is an interesting as it suggests that those with previous experience with such novel rehabilitative technology may have a predisposed resignation that they would not be able to fully use these new technological methods in their recovery and would thus rather heavily rely on some form of instruction or expertise in conducting the rehabilitation exercises. Patients with a history of stroke have experienced rehabilitation in the past and thus could experience resistance to new rehabilitation technology due to a bias or preference to use what they are already familiar with [42].

In the case of patients with a history of stroke, they might prefer to use something that worked for them in the past because there is precedence of that working. Thus, getting used to a completely new device in stroke rehabilitation presents a level of psychological uncertainty or risk perception associated with the new method and would thus cause them to feel like they may not be able to adapt or manage it efficiently without proper assistance [43].

Implications

The theoretical implication of the findings from this study has shown that the motivation levels toward new technology among a homogenous group of people can differ depending on their personal experiences related to the purpose of the proposed technology. For instance, in this study, it was demonstrated that even among older adult patients, the mere experience of a stroke altered their perception toward the wearable gloves by a significant degree compared to those who did not have a history of stroke. Therefore, it is crucial that future studies consider medical conditions or individuals' experiences when designing health technologies.

The practical implication of the findings of this study illustrates how future exergames can be properly and suitably designed for an aging audience, for example, ensuring that visual elements can be differentiated from one another in a very clear manner and that the video game's difficulty level is attuned not to be extremely difficult but moderately challenging to both prevent the demotivation of the players and encourage continued gameplay, which basically means that the patients continue the rehabilitation process through the exergame medium. In addition, digital health technologies must be designed to have clear affordances without visibly looking like it would be very difficult or complicated to use, as this has an effect on the users' perception about the apparatus, which would inadvertently affect their motivation to buy and use the product.

Limitations

Similar to all studies, there are several limitations in our research. The first limitation of this study would be the lack of a substantial sample size. The sample size of 19 is very small to draw concrete quantitative conclusions and comparison between the focus groups, and therefore, any findings and results that were gleaned from the study are educated guesses and conjecture. However, it should be noted that the opinions shared by the older adults during the study are ideas that people in the target group actually have, but they need to be carefully examined in future studies. Second, our participants only used the wearable gloves and games in 1 session with 2 games. We are cognizant that our participants who willingly participated in our research may be qualitatively different from those who did not join. Third, we are mindful that older adults who did not have a history of stroke might have other health conditions that the research team is not aware of that might influence their perceptions about the wearable glove and gameplay.

Conclusions

In summary, there is tremendous potential in the use of digital health technologies such as wearable gloves and exergames to motivate older adults to exercise and, for patients recovering from stroke, to adhere to rehabilitation exercises. While we recognize the benefits of such digital health technologies, without representation from older adults in such studies, any technology development and implementation may face the problem of data absenteeism and technology chauvinism. Thus, to achieve a more equitable and inclusive use of digital health technologies, researchers need to consider both the individuals and the contexts in which the technologies are used.

Acknowledgments

This study was supported by the Singapore Health Technologies Consortium (HealthTEC) under the National Research Foundation (021293-00001), Ministry of Education AcRF Tier 1 grant (RG38/22), and Social Science and Humanities Research Fellowship (SSRC2022-SSHR-006). The authors would like to acknowledge StretchSkin Technologies Pte Ltd, particularly Mr Mayank Rajput and the engineers, for their contributions in the wearable glove product development. The authors would also like to acknowledge the work of Mr Ming Kiang Chua, Mr Nigel Tan, and Mr Xue Wei Tan, who were student interns at NTU for contributing to the development of Virtual Exercise Therapy exergames and the implementation of the usability study. Finally, the authors are grateful to Ms Polly Lim from the Home Nursing Foundation for helping them recruit older adults for the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mean scores of the System Usability Scale items.
[PNG File , 81 KB - games_v12i1e47600_app1.png]

References

1. Lee EW, Viswanath K. Big data in context: addressing the twin perils of data absenteeism and chauvinism in the context of health disparities research. J Med Internet Res 2020 Jan 07;22(1):e16377 [FREE Full text] [doi: 10.2196/16377] [Medline: 31909724]
2. Ienca M, Schneble C, Kressig RW, Wangmo T. Digital health interventions for healthy ageing: a qualitative user evaluation and ethical assessment. BMC Geriatr 2021 Jul 02;21(1):412 [FREE Full text] [doi: 10.1186/s12877-021-02338-z] [Medline: 34215209]
3. Helbostad J, Vereijken B, Becker C, Todd C, Taraldsen K, Pijnappels M, et al. Mobile health applications to promote active and healthy ageing. Sensors (Basel) 2017 Mar 18;17(3):622 [FREE Full text] [doi: 10.3390/s17030622] [Medline: 28335475]
4. Macvean A, Robertson J. Understanding exergame users' physical activity, motivation and behavior over time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13; April 27-May 2, 2013; Paris France URL: https://doi.org/10.1145/2470654.2466163 [doi: 10.1145/2470654.2466163]
5. Godman H. Virtual reality for chronic pain relief. Harvard Health Publishing. 2022 Apr 1. URL: https://www.health.harvard.edu/pain/virtual-reality-for-chronic-pain-relief [accessed 2022-08-23]
6. Chandrasekaran R, Katthula V, Moustakas E. Too old for technology? use of wearable healthcare devices by older adults and their willingness to share health data with providers. Health Informatics J 2021;27(4):14604582211058073 [FREE Full text] [doi: 10.1177/14604582211058073] [Medline: 34802315]
7. Gupta D. What is fueling investor excitement in Healthtech? Appinventiv. 2022 Jul 22. URL: https://appinventiv.com/blog/increasing-investment-in-healthcare-technology/ [accessed 2022-08-23]
8. Landi H. 2022 forecast: investors will double down on these hot digital health markets. Fierce Healthcare. 2021 Dec 22. URL: https://www.fiercehealthcare.com/digital-health/2022-forecast-investors-will-double-down-these-hot-digital-health-markets [accessed 2024-04-16]
9. Maceira-Elvira P, Popa T, Schmid AC, Hummel FC. Wearable technology in stroke rehabilitation: towards improved diagnosis and treatment of upper-limb motor impairment. J Neuroeng Rehabil 2019 Nov 19;16(1):142 [FREE Full text] [doi: 10.1186/s12984-019-0612-y] [Medline: 31744553]

10. Macoir J, Lavoie M, Routhier S, Bier N. Key factors for the success of self-administered treatments of poststroke aphasia using technologies. Telemed J E Health 2019 Aug;25(8):663-670 [FREE Full text] [doi: 10.1089/tmj.2018.0116] [Medline: 30142032]
11. Wang S, Bolling K, Mao W, Reichstadt J, Jeste D, Kim HC, et al. Technology to support aging in place: older adults' perspectives. Healthcare (Basel) 2019 Apr 10;7(2):60 [FREE Full text] [doi: 10.3390/healthcare7020060] [Medline: 30974780]
12. Hayashi H, Tan AS, Kawachi I, Ishikawa Y, Kondo K, Kondo N, et al. Interpersonal diffusion of health information: health information Mavenism among people age 65 and over in Japan. Health Commun 2020 Jun 23;35(7):804-814 [FREE Full text] [doi: 10.1080/10410236.2019.1593078] [Medline: 31122068]
13. Lee EW, McCloud RF, Viswanath K. Designing effective ehealth interventions for underserved groups: five lessons from a decade of eHealth intervention design and deployment. J Med Internet Res 2022 Jan 07;24(1):e25419 [FREE Full text] [doi: 10.2196/25419] [Medline: 34994700]
14. Yao J, Lim N, Tan J, Matthias Müller A, Martinus van Dam R, Chen C, et al. Evaluation of a population - wide mobile health physical activity program in 696 907 adults in Singapore. J Am Heart Assoc 2022 Jun 21;11(12):e022508 [FREE Full text] [doi: 10.1161/jaha.121.022508]
15. Bol N, Helberger N, Weert JC. Differences in mobile health app use: a source of new digital inequalities? Inf Soc 2018 Apr 26;34(3):183-193 [FREE Full text] [doi: 10.1080/01972243.2018.1438550]
16. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
17. Lee EW, Yee AZ. Toward data sense-making in digital health communication research: why theory matters in the age of big data. Front Commun 2020 Feb 27;5 [FREE Full text] [doi: 10.3389/fcomm.2020.00011]
18. Leave no one behind. Principles for Digital Development. URL: https://digitalprinciples.org/resource/leave-no-one-behind-lnob/ [accessed 2024-04-16]
19. The worldwide smart glove industry is expected to reach $3.9 billion by 2028. GlobeNewswire. 2022 Jun 17. URL: https://tinyurl.com/yc2mprmn [accessed 2024-04-16]
20. Zhu Y, Gong W, Chu K, Wang X, Hu Z, Su H. A novel wearable soft glove for hand rehabilitation and assistive grasping. Sensors (Basel) 2022 Aug 21;22(16):6294 [FREE Full text] [doi: 10.3390/s22166294] [Medline: 36016055]
21. Yurkewich A, Kozak IJ, Hebert D, Wang RH, Mihailidis A. Hand Extension Robot Orthosis (HERO) Grip Glove: enabling independence amongst persons with severe hand impairments after stroke. J Neuroeng Rehabil 2020 Feb 26;17(1):33 [FREE Full text] [doi: 10.1186/s12984-020-00659-5] [Medline: 32102668]
22. Wang T, Liu Z, Gu J, Tan J, Hu T. Effectiveness of soft robotic glove versus repetitive transcranial magnetic stimulation in post-stroke patients with severe upper limb dysfunction: a randomised controlled trial. Front Neurol 2022 Jan 11;13:887205 [FREE Full text] [doi: 10.3389/fneur.2022.887205] [Medline: 36712422]
23. Seim CE, Wolf SL, Starner TE. Wearable vibrotactile stimulation for upper extremity rehabilitation in chronic stroke: clinical feasibility trial using the VTS Glove. J Neuroeng Rehabil 2021 Jan 23;18(1):14 [FREE Full text] [doi: 10.1186/s12984-021-00813-7] [Medline: 33485371]
24. Parker J, Powell L, Mawson S. Effectiveness of upper limb wearable technology for improving activity and participation in adult stroke survivors: systematic review. J Med Internet Res 2020 Jan 08;22(1):e15981 [FREE Full text] [doi: 10.2196/15981] [Medline: 31913131]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319 [FREE Full text] [doi: 10.2307/249008]
26. Harrington CN, Koon LM, Rogers WA. Chapter 17 - Design of health information and communication technologies for older adults. In: Sethumadhavan A, Sasangohar F, editors. Design for Health: Applications of Human Factors. Cambridge, MA: Academic Press; 2020.
27. Carney R, Firth J. Chapter 12 - mHealth and physical activity interventions among people with mental illness. In: Stubbs B, Rosenbaum S, editors. Exercise-Based Interventions for Mental Illness: Physical Activity as Part of Clinical Treatment. Cambridge, MA: Academic Press; 2018:217-242.
28. Brown R, Sugarman H, Burstin A. Use of the Nintendo Wii Fit for the treatment of balance problems in an elderly patient with stroke: a case report. Int J Rehab Res 2009;32:S109-S110 [FREE Full text] [doi: 10.1097/00004356-200908001-00144]
29. Bieryla K, Dold N. Feasibility of Wii Fit training to improve clinical measures of balance in older adults. Clin Interv Aging 2013 Jun;8:775 [FREE Full text] [doi: 10.2147/cia.s46164]
30. Zhao Z, Ali Etemad S, Arya A. Gamification of exercise and fitness using wearable activity trackers. In: Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS). 2015 Presented at: ISCSS 2015; September 9-11, 2015; Loughborough, UK URL: https://doi.org/10.1007/978-3-319-24560-7_30 [doi: 10.1007/978-3-319-24560-7_30]
31. Yu T, Chiang CH, Wu PT, Wu WL, Chu IH. Effects of exergames on physical fitness in middle-aged and older adults in Taiwan. Int J Environ Res Public Health 2020 Apr 08;17(7):2565 [FREE Full text] [doi: 10.3390/ijerph17072565] [Medline: 32276515]

32. Oesch P, Kool J, Fernandez-Luque L, Brox E, Evertsen G, Civit A, et al. Exergames versus self-regulated exercises with instruction leaflets to improve adherence during geriatric rehabilitation: a randomized controlled trial. BMC Geriatr 2017 Mar 23;17(1):77 [FREE Full text] [doi: 10.1186/s12877-017-0467-7] [Medline: 28330455]
33. Aggarwal P, Cha T, Wilemon D. Barriers to the adoption of really - new products and the role of surrogate buyers. J Consum Mark 1998;15(4):358-371 [FREE Full text] [doi: 10.1108/07363769810226000]
34. Perret J, Vander Poorten E. Touching virtual reality: a review of haptic gloves. In: Proceedings of the 16th International Conference on New Actuators. 2018 Presented at: ACTUATOR 2018; June 25-27, 2018; Bremen, Germany.
35. Rice M, Cheong YL, Ng J, Chua PH, Theng YL. Co-creating games through intergenerational design workshops. In: Proceedings of the Designing Interactive Systems Conference. 2012 Presented at: DIS '12; June 11-15, 2012; Newcastle Upon Tyne, UK URL: https://doi.org/10.1145/2317956.2318012 [doi: 10.1145/2317956.2318012]
36. Li J, Li C, Chia BX, Chen X, Pham TP, Theng YL. Exergaming as a community program for older adults: the effects of social interaction and competitive information. J Aging Phys Act 2021 Jun 01;29(3):466-474 [FREE Full text] [doi: 10.1123/japa.2020-0188] [Medline: 33120364]
37. Chu A, Biancarelli D, Drainoni ML, Liu JH, Schneider JI, Sullivan R, et al. Usability of learning moment: features of an e-learning tool that maximize adoption by students. West J Emerg Med 2019 Dec 09;21(1):78-84 [FREE Full text] [doi: 10.5811/westjem.2019.6.42657] [Medline: 31913823]
38. Proulx CE, Higgins J, Gagnon DH. Occupational therapists' evaluation of the perceived usability and utility of wearable soft robotic exoskeleton gloves for hand function rehabilitation following a stroke. Disabil Rehabil Assist Technol 2023 Aug;18(6):953-962 [FREE Full text] [doi: 10.1080/17483107.2021.1938710] [Medline: 34190657]
39. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
40. Dierckx de Casterlé B, Gastmans C, Bryon E, Denier Y. QUAGOL: a guide for qualitative data analysis. Int J Nurs Stud 2012 Mar;49(3):360-371 [FREE Full text] [doi: 10.1016/j.ijnurstu.2011.09.012] [Medline: 21996649]
41. Sheng B, Zhao J, Zhang Y, Xie S, Tao J. Commercial device-based hand rehabilitation systems for stroke patients: state of the art and future prospects. Heliyon 2023 Mar;9(3):e13588 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e13588] [Medline: 36873497]
42. Kim HW, Kankanhalli A. Investigating user resistance to information systems implementation: a status quo bias perspective. MIS Q 2009;33(3):567-582 [FREE Full text] [doi: 10.2307/20650309]
43. Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertainty 1988 Mar;1(1):7-59 [FREE Full text] [doi: 10.1007/bf00055564]

Abbreviations

HERO: Hand Extension Robot Orthosis
RQ: research question
SUS: System Usability Scale

Edited by G Eysenbach; submitted 26.03.23; peer-reviewed by K Bower, T Patel; comments to author 19.07.23; revised version received 21.11.23; accepted 17.03.24; published 24.04.24.

Please cite as:
Lee EWJ, Tan WW, Pham BTP, Kawaja A, Theng YL
Addressing Data Absenteeism and Technology Chauvinism in the Use of Gamified Wearable Gloves Among Older Adults: Moderated Usability Study
JMIR Serious Games 2024;12:e47600
URL: https://games.jmir.org/2024/1/e47600
doi:10.2196/47600
PMID:38656778

Original Paper

Views of Specialist Clinicians and People With Multiple Sclerosis on Upper Limb Impairment and the Potential Role of Virtual Reality in the Rehabilitation of the Upper Limb in Multiple Sclerosis: Focus Group Study

Amy Webster[1], MSc; Matthieu Poyade[2], PhD; Elaine Coulter[1], PhD; Lisa Forrest[1], MSc; Lorna Paul[1], PhD

[1]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, United Kingdom

[2]School of Simulation and Visualisation, Glasgow School of Art, Glasgow, United Kingdom

Corresponding Author:
Amy Webster, MSc
School of Health and Life Sciences
Glasgow Caledonian University
Cowcaddens Road
Glasgow, G4 0BA
United Kingdom
Phone: 44 141 273 1345
Email: Amy.Webster@gcu.ac.uk

Abstract

Background: Finding enjoyable and effective long-term approaches to rehabilitation for improving the upper limb (UL) function of people with multiple sclerosis (MS) is challenging. Using virtual reality (VR) could be a solution to this challenge; however, there is a lack of reporting on the views of people with MS and clinicians on VR-based approaches and recommendations for games for rehabilitation.

Objective: This study aims to identify common UL problems and their related current therapeutic approaches for people with MS, and to explore the opinions of people with MS and specialist clinicians on VR and obtain suggestions for the development and design of VR games.

Methods: Separate focus groups were conducted with people with MS, recruited through the MS Society UK's research network, and clinicians, recruited through the MS Trust Therapists in MS network. A total of 10 people with MS (2 focus groups) and 8 clinicians (5 physiotherapists, 2 occupational therapists, and 1 MS nurse in 2 focus groups) were involved. The focus groups were recorded and transcriptions were analyzed using theme-based content analysis.

Results: People with MS commonly reported that their UL problems interfered with activities of daily living and resulted in the loss of meaningful hobbies such as writing. Many people with MS neglected UL exercise and found strategies for adapting to the UL impairments. Similarly, clinicians stated UL rehabilitation was neglected within their service and that it was challenging to find interesting treatment strategies. VR was suggested by both participant groups as a solution, as it was convenient for people with MS to access and it could provide a more engaging and disguised approach to exercise. There were shared concerns with cybersickness and disengagement with using VR approaches. Both groups agreed games should be meaningful and adaptable for users but suggested different VR activities, with clinicians suggesting games directly reflecting activities of daily living and people with MS suggesting more abstract activities.

Conclusions: VR was well received by both people with MS and clinicians for UL rehabilitation. Recommendations were made for the development of VR rehabilitation games which are personalized and customizable for the varying abilities of people with MS.

(JMIR Serious Games 2024;12:e51508) doi:10.2196/51508

KEYWORDS

virtual reality; multiple sclerosis; upper limb rehabilitation; coproduction; activities of daily living; exercise games; upper limb impairment

Introduction

Background

Multiple sclerosis (MS) is an inflammatory demyelination disorder of the central nervous system that is estimated to affect 2.8 million people worldwide [1]. Over a third of the people with MS have upper limb (UL) dysfunction, including weakness, tremors, and spasms in one or both ULs [2]. This can result in difficulties with activities of daily living (ADL), negatively impacting quality of life and the likelihood of remaining in employment [3,4]. Problems specifically with dexterity are related to higher health care costs [5] and a higher association with depression-like psychological measures compared to problems with lower limb function [6]. Rehabilitation and physical exercise improve motor function for people with MS [7,8]. The evidence regarding UL rehabilitation is lacking in comparison with the lower limb, despite the high frequency of UL impairments and their impact on ADL [9]. In addition, there are particular challenges in finding effective yet motivating rehabilitation strategies in MS due to the long-term, progressive nature of the disease and diversity of symptoms [10].

Virtual reality (VR) is increasing in popularity in rehabilitation research and is proposed as a possible approach to encourage long-term rehabilitation [11]. VR includes digital environments that often simulate real-world experiences with reported benefits of high motivation and engagement, with real-time feedback [12]. VR has shown promising results within MS populations, but this evidence is limited in comparison with stroke, especially regarding UL function [13]. Our systematic review, investigating the effect of VR in improving UL function in MS, found early, but limited, evidence suggesting VR has the potential to improve function in people with MS [14]. There was also a low number of dropouts in most studies within the review, supporting that VR could improve adherence compared with conventional rehabilitation; therefore, VR could be useful in conditions such as MS, where prolonged rehabilitation is required.

VR is often investigated alongside video games played within a VR setting, which can be commercially available or specifically tailored games designed with a target population in mind. Commercially available exercise games, targeted at a healthy population, can be unsuitable for disabled individuals and lead to discouragement and anxiety [15]. It is beneficial to involve a sample of target users in the creation and development of effective VR-based gamified approaches [16]. This process is known as coproduction [17]. To date, no study has systematically coproduced VR games specifically for UL rehabilitation in people with MS.

Objectives

The aims of this study were to determine the views of people with MS and specialist clinicians on UL dysfunction or function in MS, challenges faced by clinicians when delivering UL therapy, barriers and motivators for exercise in MS, opinions on VR, and suggestions for development and design of VR games. These findings will guide the future development of VR applications and interventions for UL rehabilitation for people with MS.

Methods

Ethical Considerations

Ethics approval for this study was provided by the School of Health and Life Sciences Ethics Committee at Glasgow Caledonian University (HLS/PSWAHS/20/002). Informed consent was obtained from the participants and clinicians.

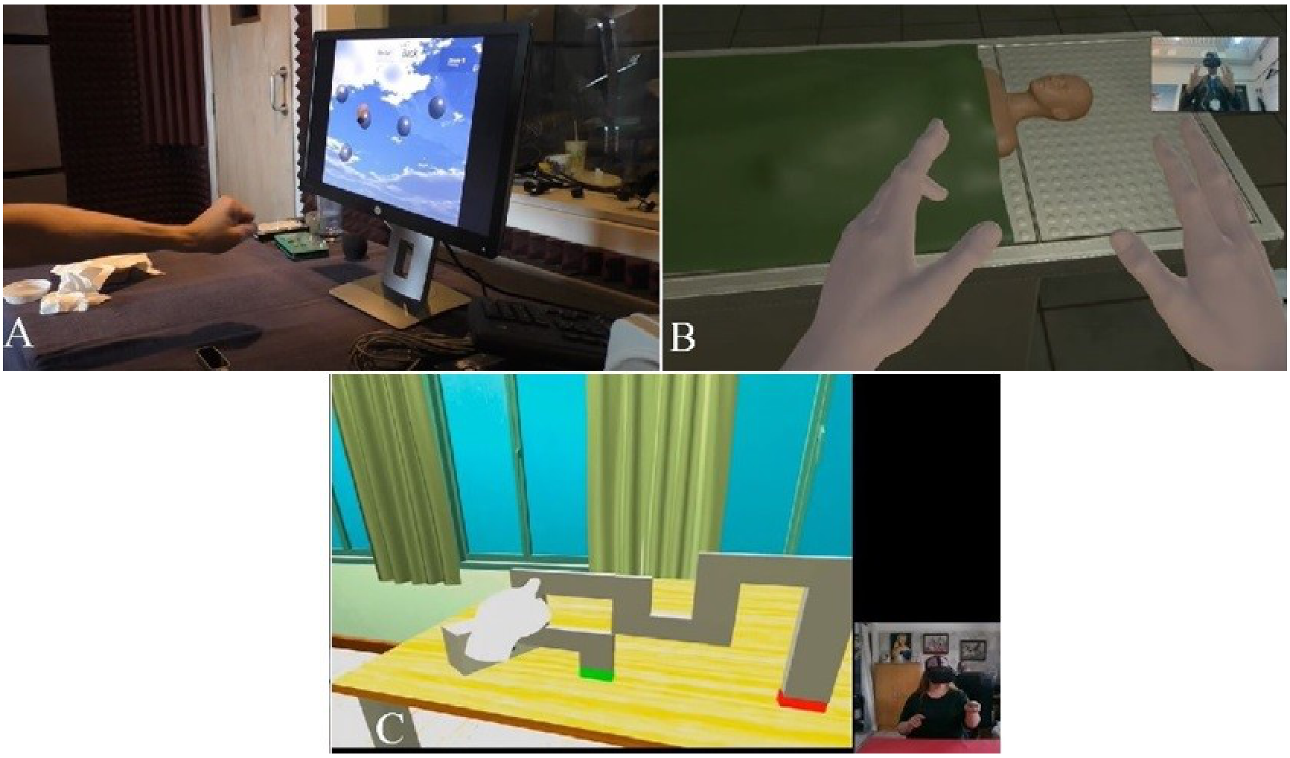
Recruitment

The study aimed to recruit up to 12 people with MS and 12 specialist MS clinicians to take part in online focus groups. The sample size was determined in line with the design of other similar studies and general recommendations for qualitative analysis [18,19]. To be included in the study, people with MS were required to be aged ≥18 years and have a diagnosis of MS (self-reported) with self-reported UL impairment. Clinicians were required to have experience (any duration) in delivering MS rehabilitation within the National Health Service (NHS) or the third sector. In addition, all participants were required to have access to and the ability to operate videoconference software. There were no specified exclusion criteria. Participants with MS were identified through the MS Society UK's research network, which advertised the study to its members. Those who were interested in participating contacted the research team directly and were emailed a participant information sheet. In terms of recruitment of clinicians, the MS Trust Therapists in MS network advertised the study to its members. Interested clinicians contacted the research team and were emailed a participant information sheet.

Coproduction Focus Groups

The focus groups for people with MS and clinicians were conducted separately with a maximum of 5 people per focus group. To comply with COVID-19 pandemic regulations at the time, focus groups were held online using Zoom (Zoom Video Communications) or Teams (Microsoft Corp) videoconference software; this also provided an opportunity for recruitment of participants from across the United Kingdom and Ireland. The focus groups were conducted in a semistructured style using a focus group schedule split broadly into three sections important for the development of VR interventions for UL problems in people with MS: (1) UL dysfunction and exercise or therapy; (2) opinions on VR; and (3) suggestions for development and design of any developed VR games (Multimedia Appendix 1). In addition, clinicians were asked what information and feedback they would want from a patient's VR therapy session. The questions included prompts that allowed more targeted responses from participants regarding their experiences and views [20]. Within the focus groups, participants were shown three videos demonstrating different commercially available head-mounted devices (HMDs) and hand-tracking devices: (1) a nonimmersive VR set up using a Leap Motion controller and computer monitor, which is a hand motion capture device that allows users to visualize their hand movements and interact with virtual environments; (2) immersive VR using the Oculus Rift HMD with a mounted Leap Motion device for hand tracking; and (3) immersive VR using the Oculus Quest, with in-built hand tracking (Figure 1 [21,22]; Multimedia Appendix 2). Videos were shown as participants were unable to try these devices since the focus groups were online due to the COVID-19 pandemic. These videos attempted to contextualize and demonstrate the different VR and motion capture devices in terms of users interacting with environments, possible hand movements, and previous games developed from prior research. After watching the videos, participants were encouraged to share their initial thoughts on each of the technologies. The focus groups involving people with MS and clinicians lasted approximately 90 minutes and 60 minutes, respectively. The focus groups were facilitated by a female researcher (AW) who had been involved in the recruitment of participants and an additional senior, female researcher (LP) attended.

Figure 1. Stills from videos shared with participants during focus groups, demonstrating different virtual reality technology. (A) Video 1 shows Leap Motion only; (B) video 2 shows Leap Motion and Oculus Rift; and (C) video 3 shows Oculus Quest.



Data Analysis

All focus groups were audio recorded and transcribed verbatim. Qualitative analysis of the data was performed based on theme-based content analysis (TBCA) as described in the study by Neale and Nichols [23]. This qualitative method groups responses into content-related themes to enable researchers to view the user preferences more easily and has been used to influence the development or evaluation of a VR environment [23-25]. TBCA is a flexible qualitative data analysis method that involves five key steps: (1) data collection, (2) data collation, (3) raw theme definition and classification, (4) higher order theme selection, and (5) presentation of classification matrix [23]. Owing to the large number of higher order themes, we added an additional step by grouping the higher order themes into main themes. The raw themes were assigned independently by 2 researchers in the transcripts of people with MS (AW and LF) and clinicians (AW and LP). After agreement on the raw themes, the responses were then independently grouped by 2 researchers (AW and LP) into higher order themes. Any discrepancies in assigning the themes were resolved through consultation with a third reviewer, if necessary. Once the higher order themes were determined, the main themes were determined by 2 researchers (AW and LP). The main themes with their associated raw and higher order themes are presented in tables. The raw and higher order themes were quantified manually within the matrix based on the number of responses necessary to display popularity or consensus [23], and example quotes for each higher order theme were included. Focus groups of people with MS and clinicians were analyzed separately to allow comparison of the findings between the 2 groups.

Results

Participant Demographics

A total of 10 people with MS were recruited to the study and took part in 1 of 2 focus groups, each of which had 5 participants. Most participants with MS were female (7/10, 70%), with a mean age of 56.4 (SD 16.5) years and a mean time since diagnosis of 14.4 (SD 12.3) years. Participants had varying MS types (Table 1).

A total of 8 clinicians were recruited (5 physiotherapists, 2 occupational therapists, and 1 MS specialist nurse). Among them, 6 participants worked in the NHS and 2 worked in other settings. There were 2 focus groups for clinicians with 4 participants in each group. All clinicians were female, with a mean age of 46.2 (SD 9.6) years, and the mean length of experience was 17.9 (SD 10. 2) years.

Table 1. Demographic details of people with multiple sclerosis.

Participant ID	Age (years; mean 56.4, SD 16.5)	Sex	Multiple sclerosis type	Time since diagnosis (years; mean 14.4, SD 12.3)
P1	60	Female	SPMS[a]	30
P2	38	Female	RRMS[b]	4
P3	68	Male	SPMS	35
P4	58	Female	SPMS	1
P5	42	Female	SPMS	11
P6	28	Female	RRMS	3
P7	56	Female	PPMS[c]	5
P8	70	Male	PPMS	16
P9	60	Male	SPMS	12
P10	84	Female	SPMS	27

[a]SPMS: secondary progressive multiple sclerosis.

[b]RRMS: relapsing-remitting multiple sclerosis.

[c]PPMS: primary progressive multiple sclerosis.

People With MS: TBCA

Overview

Following TBCA of the focus groups of people with MS, 20 higher order themes were determined based on the grouping of the assigned raw themes. These 20 higher order themes were grouped into four main themes: (1) Impact of MS on the UL; (2) Exercising with MS; (3) Views of people with MS on VR; and (4) Recommendations for development and user requirements (Table 2). A full version of this table, including more example quotes from participants, is available in Multimedia Appendix 3.

Table 2. Main, higher order, and raw themes from theme-based content analysis of people with multiple sclerosis focus groups.

Main theme and higher order themes (number of responses)	Raw themes (number of responses)
Impact of MS[a] on the UL[b]	
Interference with functional activities (35)	Dressing (8); eating (6); dropping items (5); writing (5); grooming (3); dependence on others for activities of daily living (3); carrying items (3); and traveling (2)
Symptoms and signs that impact activities (25)	Fatigue (10); numbness (6); sensory overload (4); weakness (3); tremors (2); proprioception (2); and coordination (1)
Strategies people with MS adopt to assist with ADL[c] (24)	Strategies for functional activities (8); adapting (7); making a difference (5); technology assistance (2); and mobility-assistance equipment (2)
Struggle with loss of meaningful activities and skills (14)	Loss of skills (6); impact of losing ability to write (4); and keeping meaningful activities (4)
UL actions people with MS find difficult (13)	Dexterity (6); range of motion (4); and grip (3)
Sharing and sympathy (13)	Sharing strategies (4); sharing advice on exercise (4); taking advice (3); and sympathizing (2)
Difficulty with progression and unpredictable nature of MS (10)	Variation in MS (6); unpredictable (2); and progression (2)
Exercising with MS	
Views and attitudes on exercise (49)	Maintenance (10); negative perceptions of exercise (8); keeping muscle strength (8); determined to exercise (7); benefits of exercise (6); multitask approach (4); legs focus (3); and in control (3)
Previous experience of UL rehabilitation or exercise (40)	Outcomes from UL exercise or rehabilitation (12); neglecting UL exercise or rehabilitation (10); UL equipment (6); UL physiotherapy (4); driven for UL exercise (3); UL exercise resources (3); and adherence (2)
Barriers to exercise (28)	Personal barriers (8); environmental barriers (8); COVID-19 barriers (7); and verbal disengagement (5)
Facilitators to exercise (28)	Verbal encouragement (10); health care professionals (8); MS center (4); gym facilitators (3); and pushing self for results (3)
Adverse effects of exercise (11)	Induce symptoms (4); tiring (3); recovery time after exercise (2); affecting socializing (1); and overdoing exercise (1)
Approaches to exercise used by people with MS (26)	Routine (7); exercise bikes (6); exercise aims (5); low impact or stretching exercise (4); and physiotherapy approaches (4)
Views on group vs individual exercise (26)	Competition in exercise (10); motivation of group exercise (5); downsides of group exercise (5); importance of socializing in exercise (2); camaraderie (2); enjoyment (1); and interest in group exercise (1)
Views of people with MS on VR[d]	
Positive views on VR (55)	Home use (9); outcome benefits (6); personal opinions on VR (5); fun (5); adaptable (5); positives of technology (5); wireless convenience (5); accessibility convenience (4); incentives (3); meaningful (3); online socializing (2); and immersion (1)
Negative views on VR (40)	Cybersickness (17); HMD[e] discomfort (6); technology discomfort (5); HMD dislike (3); disengagement (3); accessibility concerns (3); and unsuitability (3)
Views on trying or participating in VR rehabilitation (25)	Openness to VR (12); challenging (4); safety considerations (3); need results (2); technology considerations (2); and unsuitable for them (2)
Recommendations of people with MS for development and user requirements	
Considerations for development of VR games (84)	Mindful of target audience (9); tracking progress (8); discouragement of feedback (8); knowing UL outcomes (7); end result (6); score targets (6); challenging self (6); competition in games (5); education (5); time feedback (4); supervision (4); community involvement (3); multipurpose (3); continuous development (3); be fun (3); hardware (2); and learning patterns concern (2)
Suggestions for VR activities (36)	Suggested UL actions (9); game ideas (7); real-life vs abstract tasks (4); haptic activities (4); strength in games (4); writing and drawing (3); demonstrated games (3); additional objectives (2); and atmosphere (1)
Importance of choice (23)	Offer different movements (8); having a variety of games (6); personal preferences (6); and variety of different levels (3)

[a]MS: multiple sclerosis.

[b]UL: upper limb.

[c]ADL: activities of daily living.

[d]VR: virtual reality.

[e]HMD: head-mounted device.

Impact of MS on the UL

The most common higher order theme was "Interference with functional activities" with 35 responses (Table 2). Participants reported a wide range of activities they found difficult to perform due to their MS, the most frequent being ADL, including personal care, eating, and carrying heavy items. "Symptoms and signs that impact activities" had the second highest number of responses (n=25), where participants particularly noted the impact of fatigue on activity (n=10); however, sensory problems such as numbness and pins and needles were also highlighted. Other MS symptoms impacting UL function were, for example, weakness, tremors, and coordination problems. In "Strategies people with MS adopt to assist with ADL" (n=24), because of losing function, participants discussed the use of assistive equipment, for example, button fasteners, specialized cups, and voice control. Other strategies included using their less affected hand or pacing to manage fatigue. The remaining 4 higher order themes had fewer responses. In brief, dexterity, range of joint movement, and grip were the main "UL actions people with MS find difficult" (n=13). These were often compounded by the unpredictability and progressive nature of MS ("Difficulty with progression and unpredictable nature of MS," n=10). Participants reported the emotional impact of losing the ability to carry out personal and meaningful activities specifically because of loss of UL function ("Struggle with loss of meaningful activities and skills," n=14), with one participant stating the following:

> *I used to be a writer and it was very, very hard because I couldn't write anymore...I was really motivated [to relearn writing], felt really cut off from the world.* [P8; age 70 years; male participant with primary progressive multiple sclerosis]

The final higher order theme was "Sharing and sympathy" (n=13 responses), where participants empathized and shared experiences and suggestions of assistive equipment.

Exercising With MS

Most responses under this main theme related to "Views and attitudes on exercise" (n=49; Table 2). Participants were motivated to exercise with a "*use it or lose it*" attitude and a desire to, if not improve then at least maintain, their function and prevent further deterioration. Participants also described negative perceptions of exercise, such as finding it "*very boring*" and guilt from not participating in exercise. In "Previous experience with UL rehabilitation or exercise" (n=40), many participants (5/10, 50%) discussed not undertaking any UL exercise or rehabilitation, currently or previously. Many UL programs previously undertaken by some participants aimed to build strength, reduce pain, and improve hand function with varying outcomes. There were similar numbers of responses in terms of "Barriers to exercise" (n=28) and "Facilitators to exercise" (n=28). Personal barriers to exercise included comorbidities, MS symptoms (fatigue, pain, and bladder and bowel dysfunction), difficulty using exercise equipment, and expense. The COVID-19 pandemic had negatively impacted the participants' exercise due to services closing down. Environmental barriers to exercise included lack of local facilities and not having space to exercise at home. Verbal encouragement was described as both a barrier (could be off putting) and a facilitator (motivating) to exercise. Other facilitators were seeing improvements, feeling motivated, and the attitudes of health care professionals, personal trainers, and carers. Conversely, health care professionals with a lack of experience in MS overwork people with MS, leading to exhaustion ("Adverse effects of exercise," n=11). Participants undertook many different forms of exercise ("Approaches to exercise used by people with MS," n=26), including exercise bikes, Pilates and yoga, dog walking, and gym exercises. There were varying "Views on group versus individual exercise" (n=26). Some found competition within a group to be motivating while others did not, with one participant suggesting social support and camaraderie was more important than competition:

> *I'm not too fussed about being in competition with others, but if it was a more social thing that would maybe encourage me to perhaps join in a group that's doing something together.* [P4; age 58 years; female participant with secondary progressive multiple sclerosis (SPMS)]

Negative aspects of group exercise included the fear of letting others down.

Views on VR

The initial reaction to VR was positive ("Positive views on VR," n=55; Table 2). Participants stated it looked fun or enjoyable with the potential to improve or maintain muscle strength, dexterity, and spatial awareness, especially with repeating the actions and concurrently perhaps learning a new skill (for example, playing the piano):

> *I think [VR's] still very good because... it's...maintaining those motor skills that is so easily slip away when you're not using them.* [P9; age 60 years; male participant with SPMS]

There were positive comments in relation to the convenience and accessibility of VR facilitating exercise at home at a suitable time and eliminating travel to physiotherapy services and gyms. Participants highlighted that the wireless HMD was more convenient as it was portable and did not need a computer. The advantage of linking up with others online was raised. However, "Negative views on VR" (n=40) were related to concerns regarding cybersickness, linked to dizziness and balance problems:

> *With MS a lot of people suffer from nausea or motion sickness. That can be a concern for the headsets.* [P6; age 28 years; female participant with relapse and remitting multiple sclerosis]

Other negative responses related to the HMD discomfort regarded weight, usability concerns, wearing it with glasses, and being disconnected from the real world. Two participants indicated that interest in VR may reduce over time. Participants were also concerned about fatigue and the usefulness of VR for UL sensory dysfunction. Most participants (6/10, 60%) expressed they were open to trying VR ("Views on trying or participating in VR rehabilitation," n=25), but would like to understand the benefits, long-term outcomes, and any safety issues.

Recommendations for Development and User Requirements

With regard to "Considerations for development of VR games" (n=84), a variety of UL movements was desirable with clarity in terms of the aim and outcome in relation to the UL being important (Table 2). Competition within the VR games, interacting with others or challenging themselves, were frequently discussed as being motivating. Tracking improvements during VR gameplay was vital to some participants, including monitoring improvements in score, exercise time (rather than countdown which could be stressful), and progressive challenges. The games should offer the ability to challenge users, with one participant saying the following:

> *That challenge to try and be better the next time, whereas if you've got no idea...you've got nothing to fight against or to work against.* [P10; age 84 years; female participant with SPMS]

Conversely, other participants emphasized the potential demotivating effect of feedback given the progressive nature of MS, by warning that score feedback should not be "*disheartening*," and should therefore be made optional to the user. There was a strong feeling that the VR games should be "*fun*" with abstract gameplay potentially being more fun. Participants felt that demonstrations and supervision to assess progress were important. They also stated that the VR games had to account for the differences in the ability of people with MS and that older people may need more basic VR games. The idea of the VR games having an educational outcome or in learning a new skill was suggested to help with engagement. Participants suggested that reaching, punching, and other aerobic activities could be incorporated ("Suggestions for VR activities," n=36). Having haptic approaches was frequently proposed with gripping, squishing games, such as kneading bread. Participants proposed activities with a cognitive element, such as a puzzle or maze, and whole limb movements, such as Whack-a-Mole (Mattel), writing or drawing. Participants liked the VR piano which had been demonstrated. There was a variety of opinions in terms of abstract or real-life activity with most preferring abstract games but some ADL-type activity was also suggested. "Importance of choice" (n=23) related to having variety in games, UL movements, and levels of difficulty with abstract games or real-life gamified tasks, with 1 participant declaring the following:

> *I'd like to make sure I'm not doing a whole lot of exercises that are all doing the same things...Got to be mixing them up: one for coordination, one for dexterity.* [P1; age 60 years; female participant with SPMS]

Clinicians: TBCA

Overview

From the clinician focus groups, there were 15 higher order themes grouped into four main themes: (1) Current methods and challenges for delivering UL rehabilitation; (2) Clinicians' views on VR; (3) Recommendations for development and user requirements; and (4) Implementation of VR into practice (Table 3). A full version of this table, including more example quotes from participants, is available in Multimedia Appendix 4.

Table 3. Main, higher order, and raw themes from theme-based content analysis of clinician focus groups.

Main themes and higher order themes (number of responses)	Raw themes (number of responses)
Current methods and challenges for delivering UL[a] rehabilitation	
Challenges clinicians face when delivering exercise for people with MS[b] (52)	MS-specific challenges (13); patient adherence (11); service challenges (9); UL-related challenges (7); patient differences (6); challenges with current methods of delivery (4); COVID-19 impacts (2)
Recommended UL exercises for people with MS (29)	Actions (10); systematic approach (7); functional tasks (6); strength and range of movement (5); relapse care (1)
Experience with long-term, progressive condition (24)	Deterioration (11); acceptance in patients (8); difficulty with patient improvements (5)
Factors clinicians consider when prescribing exercise for the UL (22)	Meaningful and patient-focused (9); patient assessments (6); symptoms (4); repetition (3)
Current methods of UL exercise delivery for people with MS (15)	Technological approaches (4); programs (4); accessible equipment (3); clinician routines (2); patient lead (2)
Socializing in exercise (14)	social motivation (6); support (5); recommending social exercise (3)
Clinicians' views on VR[c]	
Positive views on VR (50)	Solutions to current challenges (10); personal opinions on VR (7); facilitating movements or tasks (6); VR-specific qualities (6); meaningful (5); engagement (5); visualization (4); novel (3); cognitive appeal (2); adaptability (2)
Negative views on VR (38)	Disengagement (10); cybersickness and safety (8); HMD[d] discomfort (7); accessibility concerns (5); feedback concerns (5); validity concerns (3)
Questioning benefits and the unknowns of VR (14)	Questioning purpose of VR (4); questioning benefits of VR (4); neural mechanisms (3); research (2); different VR systems (1)
Clinicians' recommendations for development and user requirements	
Considerations for developing VR games for people with MS (41)	Communication between clinician and patient (12); purposeful (7); social components (7); selecting tasks (4); slower tasks (3); competition (3); feedback for clinician (3); positive feedback (2); end point (2)
Suggestions for VR activities (18)	Activities of daily living (6); hobbies (6); objectives (6)
Importance of choice (15)	Preferences (6); having variety (5); setup (4)
Implementation of VR into practice	
Suggestions for incorporation of VR into practice (18)	Home use (7); VR in clinics (7); long-term treatment (4);
Challenges with implementation of VR into practice (24)	Funding (7); demanding on services (6); availability of equipment (5); risk (3); adjustment (2); uncertainty of practice (2)
Finding the target audience for VR (8)	Who would use VR (3); niche group (3); age (2)

[a]UL: upper limb.
[b]MS: multiple sclerosis.
[c]VR: virtual reality.
[d]HMD: head-mounted device.

Current Methods and Challenges for Delivering UL Rehabilitation

"Recommended UL exercises for people with MS" (n=29) included strength training and active movements related to functional activity, such as hand-to-mouth movements (Table 3). Treatment for the UL often involved equipment such as Therabands and Theraputty but also technology such as the Gloreha robotic system and functional electrical stimulators with different models of care for UL exercises described as part of community-based classes, within third sector organizations and online programs ("Current methods of UL exercise delivery for people with MS," n=15). Within "Factors clinicians consider when prescribing exercise for the UL" (n=22), most responses were regarding meaningful and goal-focused exercises. Clinicians also considered the patient's symptoms, for example, spasticity, pain, and the ability of patients. The importance of repetition of movement was reinforced. Most responses (n=52) were in relation to "Challenges clinicians face when delivering exercise for people with MS." Clinicians expressed that UL-focused exercise was neglected compared to the lower limb and the challenge of making UL exercise interesting:

> *A bit more difficult for upper limb things...it's much easier to maybe...go for a walk with somebody or you know, or cycle or whatever. Upper limb is maybe a wee bit more difficult.* [C6; physiotherapist]

Clinicians also mentioned the use of Theraputty described as "*juvenile*" and lists of exercises "*boring*."

Service-related challenges included limited time and capacity to see patients and large geographic areas to cover. Other challenges were keeping patients engaged with exercise in the long term, especially at home, and finding an activity that would be attractive to patients. Under "Experience with long-term, progressive condition" (n=24), clinicians raised being realistic about improvements with a progressive condition while also keeping patients motivated, minimizing deterioration or maintenance, rather than improving:

> *Trying to motivate people with progressive MS, you're trying to get them to continue to maintain where they are rather than improve.* [C5; occupational therapist]

Clinicians expressed the positive benefits of "Socializing in exercise" (n=14) for support and motivation.

Clinicians' Opinions on VR

Clinicians were very positive about VR (n=50), describing it as being interactive, fun, meaningful, and a novel potential approach to rehabilitation, which could help engagement (Table 3). They were positive about the escapism aspect and the potential to improve mental health:

> *What appeals about VR stuff is that it is focused and takes you into a different place...You're doing tai chi on a beautiful, Japanese garden rather than actually in your grumpy living room...I think even that in terms of the escapism aspect, maybe from a mental wellbeing.* [C1; physiotherapist]

Clinicians liked the visual feedback to help with, for example, coordination, but which could also reinforce movements and introduce a cognitive component. Clinicians commented that VR provided the opportunity to undertake activities not possible within the clinic and to exercise without the activity seeming like an exercise. Most of the "Negative views of VR" (n=38) were regarding patient safety using VR headsets, especially cybersickness, including dizziness and disorientation, specifically in patients with vestibular issues. Other general concerns with HMDs were usability with glasses, the weight of the HMD, and feeling claustrophobic. Clinicians suggested that VR activities should not be too simplistic to avoid patronizing patients and at an appropriate skill level. The longevity of engagement of patients after the initial novelty was questioned. Clinicians also questioned the use of VR for activities that can be done in the real world and similarly how VR activities might translate to real function. The importance of feedback on the quality of movement as well as the quantity was highlighted. Finally, accessibility and digital poverty were also raised. The final higher order theme was "Questioning benefits and the unknowns of VR" (n=14), where some clinicians felt there was insufficient evidence on the purpose and benefits of VR and its effect on neural mechanisms:

> *I think it's important to think about how is [VR] different to just doing [activities] in real life as well...What can you augment in your rehab through this virtual reality that you can't just do in real life anyway?* [C7; physiotherapist]

Recommendations for Development and User Requirements

Under "Considerations for developing VR games for people with MS" (n=41), clinicians discussed the importance of the VR games having purposeful activity, translation of tasks into real life, and having an end point (Table 3). The games should consider movements of individual joints of the UL with extension movements at the wrist and fingers being important as where people with MS lose the most function. Games should incorporate strength, coordination, proprioception, and range of motion exercise as well as exercises for the core. Feedback was important, with clinicians able to monitor the program. Clinicians were not interested in scores for the games but wished feedback on the quality of the movements and patient engagement. Clinicians stated that undertaking VR activities with others or in group settings with elements of competition was desirable. Clinicians provided "Suggestions for VR activities" (n=18), including ADL such as putting on makeup; writing or chopping vegetables; and hobbies including pottery, sewing, or piano playing. Clinicians raised the "Importance of choice" (n=15) in the VR setup, choice of games, and choice within games, for example, levels of difficulty, to appeal to as many people as possible:

> *I think, it is about having a variety of things that push as many buttons with patients that you can manage and cover as many options as you can.* [C2; physiotherapist]

Implementation of VR Into Practice

Under "Suggestions for incorporation of VR into practice" (n=18), clinicians felt long-term, regular use of VR was needed for positive outcomes (Table 3). Home use was felt to encourage frequent use, with clinicians monitoring progress remotely, thus saving in person contact time. There were a number of "Challenges with implementation of VR into practice" (n=24) with cost and funding (service and individual) being the most commonly reported, which included potential increased demand on services:

> *I know if I brought it to my bosses they would want a breakdown of cost of monthly rate, how are we going to utilise it, how often are we going to utilise it. What figures could we get from this particular item and what outcomes could we achieve.* [C4; MS specialist nurse]

Equipment-related challenges were ownership, availability, supply of equipment, and infection control. A full risk assessment would be required before implementation and guidance would be needed on intervention duration and frequency. Clinicians discussed for whom VR would be appropriate, in terms of age or other factors, and identified this as an area for future research ("Finding the target audience for VR," n=8).

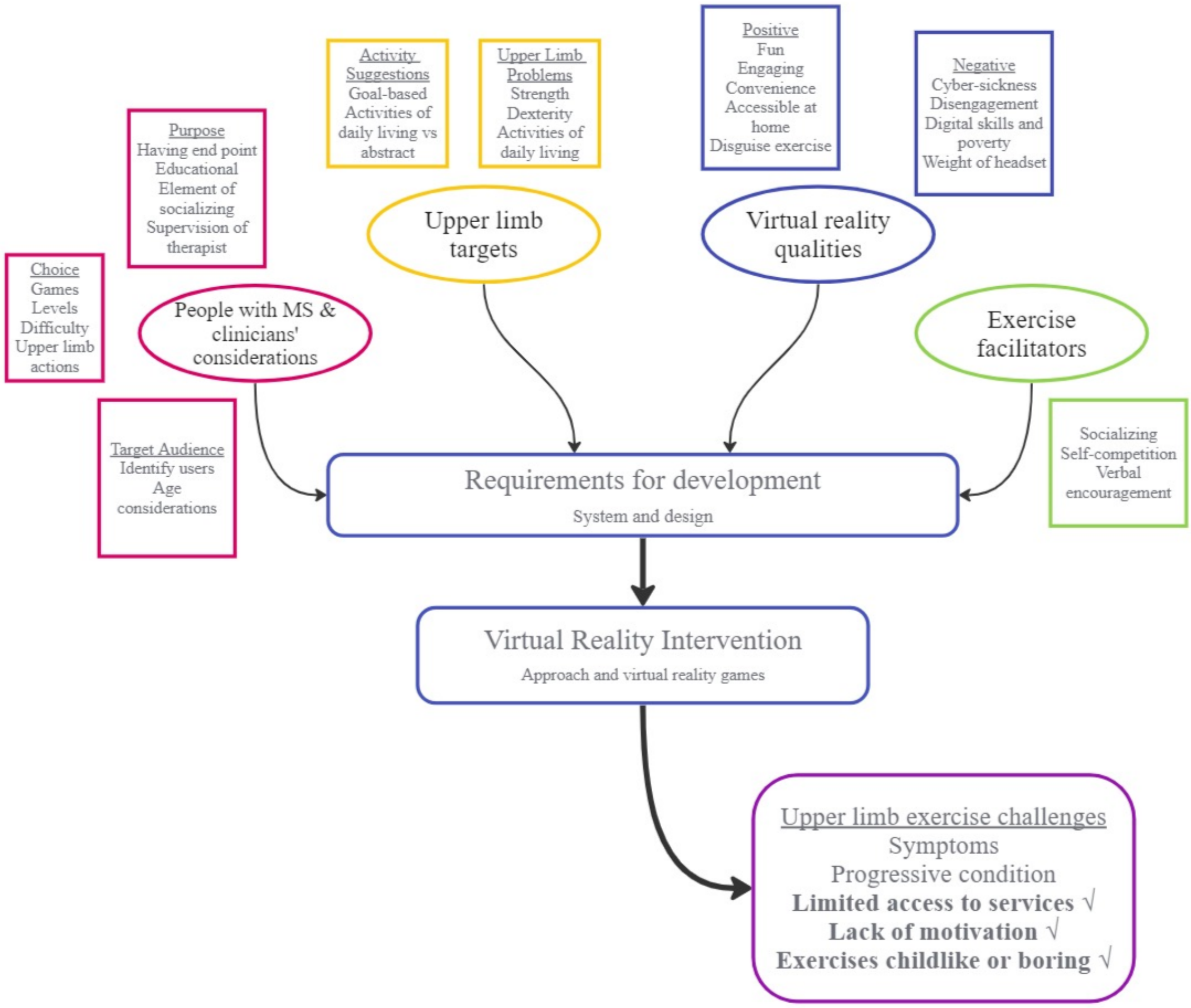
Discussion

Principal Findings

This study aimed to explore the views of people with MS and clinicians on UL impairment associated with MS and the potential role of VR as a rehabilitation approach to address this impairment. The discussion focuses on the combined findings from the 2 groups of participants: people with MS and clinicians (Figure 2). Figure 2 is a visual representation of the principal findings based on the higher number of responses assigned, which should inform the development of VR applications and interventions aiming to improve the UL function for people with MS and how VR could tackle challenges of existing UL exercise raised by clinicians and people with MS in this study.

The findings agree with those of previous studies that people with MS commonly have UL impairments that impact function, including problems with dexterity and ADL, which leads to loss of meaningful activities [26-28]. Despite UL difficulties, UL exercise was neglected due to MS symptoms, such as fatigue, lack of motivation, and dislike of exercise, as well as the challenges clinicians faced regarding time constraints and finding appropriate therapies that were not childlike or boring. Lack of focus on UL rehabilitation has been reported previously in MS [9] and in other long-term neurological conditions such as stroke [29]. The progressive and unpredictable nature of MS was raised by both groups, and consequently, clinicians raised the importance of setting realistic expectations with therapy, sometimes focusing on maintenance of function rather than improvement.

Figure 2. Flow diagram of the results from theme-based content analysis of people with multiple sclerosis (MS) and clinicians' coproduction focus groups and how this will guide the requirements for developing virtual reality–based games or interventions, which will aim to tackle certain achievable upper limb exercise challenges within MS.



Both groups (people with MS and clinicians) were optimistic about the use of VR and believed VR could be a solution to their exercise challenges. Positive comments, including avoiding traveling, being accessible, and being engaging or fun addressed the identified barriers for UL rehabilitation. This concurs with previous VR studies [30,31] and specifically in UL rehabilitation in MS, with a recent home-based, feasibility study using the Oculus Quest 2 VR headset in which participants described VR as fun, interesting, and innovative [32]. Participants in the study by Kamm et al [32] suggested adding difficulty levels and scoring to their exercises, competitive elements previously described to be motivating by people with MS using nonimmersive exercise games delivered through the Nintendo Wii [33]. In this study, both groups were especially positive regarding the immersive approach of the Oculus Quest. Participants thought the escapism properties and visualization of movements could potentially "disguise exercise," which may occur with the "fun" element of VR reducing the perception of exertion during exercise [34], therefore encouraging more UL therapy.

Negative views about VR were also expressed, mainly the potential for cybersickness. Cybersickness is thought to be caused by conflict of stimuli, leading to nausea, disorientation, and pain in the eyes and head [35]. Women are more susceptible to cybersickness [36], which is relevant in MS, with a higher number of women affected. Although cybersickness with VR has been reported previously in people with MS [37], there are development strategies for reducing cybersickness, such as designing VR activity with less overall movement within the virtual environment. Cybersickness is, however, thought to reduce over time with exposure to VR [38]. There were unnecessary concerns raised for those wearing glasses as the HMD can accommodate glasses, but there were valid concerns about the weight of the HMD for some users. Disengagement was another concern both groups expressed, with limited data on long-term adherence to VR in MS rehabilitation. Exercise is a behavioral intervention, and long-term adherence to exercise can be supported by evidence-based behavior change techniques [39]. These behavior change techniques, such as goal setting, rewards, and feedback, can be incorporated into VR games or activity to support long-term engagement in UL exercise. While VR can be more engaging than other methods of exercise [40], frequent performance, feedback on progress, and adjusting levels of difficulty can maximize VR engagement for those with long-term neurological conditions [41]. Finally, clinicians had specific concerns regarding digital poverty, the technical ability of people with MS, and insufficient technical services to support VR.

Considerations for VR game development align with user-centered design principles for VR in motor rehabilitation in survivors of stroke, such as being fun, tracking progress, having an element of competition, challenging oneself, and providing feedback [42], and are not specific to any clinical population. Participants raised that VR development should be mindful of the different end users (people with MS) who may differ in ability and preferences. Clinicians suggested VR would appeal to younger individuals with MS, whereas people with MS felt older people with MS might need more basic gameplay. While there is some, albeit limited, evidence for lower usability scores for older VR users compared with younger users, there can be higher user enjoyment [43], and there is moderate evidence for good usability of VR in older populations [44]; therefore, this concern may be overly cautious.

Consideration of the end user links to the importance of choice when designing VR interventions, with a variety of games to appeal to as many as possible. Participants felt the games should include different movements, levels of immersion, level of difficulty, or feedback on performance. Accommodating individual preferences is a key element for the design of VR games for rehabilitation, as it increases user engagement [45]. However, our previous systematic review found that a choice of games was rarely included in VR interventions in MS [14].

There were differing views in terms of the type of feedback people wished from VR. Some people with MS wanted to track scores and visualize results, which is supported by reward theories for users during both entertainment and serious games [46]. Conversely, concerns were raised about feedback potentially being discouraging or demotivating, especially given the variable nature of MS. As an example, countdown timers provide slight pressure to motivate players to increase engagement [47]; however, in this study, people with MS felt they could be stressful. Feedback on the duration of exercise completed was appealing to people with MS, as reported previously [19]. As well as the quantity of VR exercise, clinicians also wished feedback on the quality of movement when performing the games. Rehabilitation often involves highly repetitive movements to stimulate neuroplasticity; however, stroke specialist therapists have also previously reported concerns that quality of movement in VR rehabilitation for UL maybe sacrificed for a good gaming outcome [18], although this has not been explored in people with MS. Both groups were interested in the reported outcomes of using VR approaches which, if positive, would increase engagement.

Clinicians and people with MS felt VR activity had to be related to the patient's personalized and meaningful goals, which is known to increase motivation in physiotherapy settings [48]; however, this is often neglected in VR regimes [14]. Goals need to be adjusted over time in a progressive condition, such as MS, and to avoid disengagement as raised earlier. Participants with MS frequently stated that their goals were related to not only improvement but also the maintenance of ability and the prevention of further deterioration. In terms of suggestions for VR activities, the groups differed with clinicians suggesting ADL or hobby simulations and people with MS being more ambivalent, stressing activity to be fun with a variety of real-life and abstract VR games. Previous studies of VR have often involved ADL activities such as cooking or other kitchen activity [49,50]. Although VR can provide a safe environment to practice ADL for people with mobility issues [51] people with MS in this study were less interested in ADL, especially kitchen simulations. Both groups suggested an "end result," such as creating a drawing, or learning a new skill would be positive and facilitate a feeling of accomplishment. There were also suggestions from people with MS to incorporate haptic activities, such as grabbing and gripping. However, the user is not able to receive tactile feedback when interacting with a virtual environment, and handheld controllers may need to be considered for some VR activities [52]. Another solution could be to incorporate pseudo haptics, the use of different stimuli such as visual or auditory stimuli, to mimic a variety of haptic properties in a virtual environment [53]. This is an emerging field that could be explored in VR for people with MS. Similarly, as many of the participants suggested finger-related exercises, it is important that VR systems use good hand-tracking motion capture devices to allow visualization of the movement of fingers and wrists within a VR setting.

Many people with MS were supportive of VR for home use, as being more convenient and accessible. However, there was recognition that users needed demonstration of the technology and a level of clinician supervision. Assessing quality of movements and monitoring of patient progress are reported challenges for VR home use [54]. A recent study with a small number of participants found VR to be feasible for home-based UL rehabilitation in people with MS, after 3 supervised sessions [32], but larger studies of home-based VR for UL rehabilitation are required. There was agreement in both groups that an element of social interaction could be considered in the development of VR games. Generally, there is a lack of evidence on the effect of socialization within UL therapy, but it may improve adherence and motivation [55] and provide better outcomes [56]. Specifically in relation to VR, there is some evidence that social aspects increase motivation through competition [57], but participants in our study were more interested in self-competition rather than competing against others. This is similar to a study of a walking app for MS where users were less interested in sharing their goals or achievements with others [58].

Strengths and Limitations

Recruiting participants through online sources may result in a biased sample, as those comfortable with technology and access to online services are more likely to take part. Being online allowed the involvement of people with MS with varying abilities and clinicians who worked in the NHS and the third sector across the United Kingdom. However, the online nature meant it was not possible for participants to physically test the VR equipment and explore their reactions. While it can also be challenging to engage all participants in online focus groups, this was resolved by asking questions using participants' names or by getting participants to use the raise hand function within the videoconference software and encourage discussion between participants.

The TBCA methodology groups responses into themes to quantify them but does not allow consideration of the interaction between participants. Participants had a number of specific questions, such as the long-term outcomes of using immersive VR, the optimal target users for VR (level of disability), and the extent of translation of VR activity into "real-life" function. However, there is currently a lack of literature to provide responses to these questions, which highlights areas for future research.

Conclusions

This is the first study exploring the views of people with MS and clinicians in terms of VR for UL rehabilitation for people with MS and has highlighted the current challenges in UL rehabilitation even though UL impairment is common and impacts meaningful activity. Overall, people with MS often found dexterity-related activities difficult, which impacted multiple ADL and challenges faced in therapy related to motivation, lack of resources, and difficulty finding interesting UL exercises. There was positive support for VR for UL exercise. Overall, to improve engagement and satisfaction for the user, this study suggests any VR games developed for people with MS should (1) be fun and engaging; (2) have clear aims related to the individual user's goals; (3) offer personalization, such as a variety of games (abstract and ADL based), different movements, levels of difficulty, and methods of feedback; (4) monitor quality as well as quantity of movement during gameplay; (5) incorporate design features to reduce the potential for cybersickness; (6) consider if the games can incorporate education or skill development; (7) incorporate aspects of social interaction; and (8) consider including haptic properties. The findings support the need for the creation of bespoke serious games rather than using commercially available exercise games, which can discourage users with motor dysfunction [15,59]. Overall, future development of VR games for UL rehabilitation should focus on a personalized and customizable approach to encourage long-term engagement to improve meaningful outcomes for people with MS.

Acknowledgments

The authors would like to thank all the participants who took part in this study, as well as the MS Society, UK, for funding this study (grant number 115).

Authors' Contributions

AW, MP, EC, and LP designed and conceptualized this study. AW, LF, and LP performed the analysis. AW and LP wrote and prepared the manuscript, with support from MP, EC, and LF.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Focus group interview questions for both participant groups.

[DOCX File , 26 KB - games_v12i1e51508_app1.docx]

Multimedia Appendix 2

Compilation of 3 videos shown to participants during focus groups showing 3 different virtual reality systems.

[MP4 File (MP4 Video), 20062 KB - games_v12i1e51508_app2.mp4]

Multimedia Appendix 3

Main, higher order, and raw themes from theme-based content analysis of people with multiple sclerosis focus groups, with example quotes.
[DOCX File , 26 KB - games_v12i1e51508_app3.docx]

Multimedia Appendix 4
Main, higher order, and raw themes from theme-based content analysis of clinician focus groups, with example quotes.
[DOCX File , 25 KB - games_v12i1e51508_app4.docx]

References

1. Walton C, King R, Rechtman L, Kaye W, Leray E, Marrie RA, et al. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. Mult Scler 2020 Dec 11;26(14):1816-1821 [FREE Full text] [doi: 10.1177/1352458520970841] [Medline: 33174475]
2. Marrie RA, Cutter GR, Tyry T, Cofield SS, Fox R, Salter A. Upper limb impairment is associated with use of assistive devices and unemployment in multiple sclerosis. Mult Scler Relat Disord 2017 May;13:87-92. [doi: 10.1016/j.msard.2017.02.013] [Medline: 28427709]
3. Simmons RD, Tribe KL, McDonald EA. Living with multiple sclerosis: longitudinal changes in employment and the importance of symptom management. J Neurol 2010 Jun;257(6):926-936. [doi: 10.1007/s00415-009-5441-7] [Medline: 20084515]
4. Goverover Y, Genova HM, DeLuca J, Chiaravalloti ND. Impact of multiple sclerosis on daily life. In: Chiaravalloti ND, Goverover Y, editors. Changes in the Brain: Impact on Daily Life. New York, NY: Springer; 2017:145-165.
5. Koch MW, Murray TJ, Fisk J, Greenfield J, Bhan V, Jacobs P, et al. Hand dexterity and direct disease related cost in multiple sclerosis. J Neurol Sci 2014 Jul 15;341(1-2):51-54. [doi: 10.1016/j.jns.2014.03.047] [Medline: 24768310]
6. Alonso RN, Eizaguirre MB, Cohen L, Quarracino C, Silva B, Pita MC, et al. Upper limb dexterity in patients with multiple sclerosis: an important and underrated morbidity. Int J MS Care 2021;23(2):79-84 [FREE Full text] [doi: 10.7224/1537-2073.2019-083] [Medline: 33880084]
7. Dalgas U, Ingemann-Hansen T, Stenager E. Physical exercise and MS recommendations. Int MS J 2009 May;16(1):5-11. [Medline: 19413920]
8. Etoom M, Khraiwesh Y, Lena F, Hawamdeh M, Hawamdeh Z, Centonze D, et al. Effectiveness of physiotherapy interventions on spasticity in people with multiple sclerosis: a systematic review and meta-analysis. Am J Phys Med Rehabil 2018 Dec;97(11):793-807. [doi: 10.1097/PHM.0000000000000970] [Medline: 29794531]
9. Lamers I, Maris A, Severijns D, Dielkens W, Geurts S, Van Wijmeersch B, et al. Upper limb rehabilitation in people with multiple sclerosis: a systematic review. Neurorehabil Neural Repair 2016 Oct 10;30(8):773-793. [doi: 10.1177/1545968315624785] [Medline: 26747125]
10. Burks JS, Bigley GK, Hill HH. Rehabilitation challenges in multiple sclerosis. Ann Indian Acad Neurol 2009 Oct;12(4):296-306 [FREE Full text] [doi: 10.4103/0972-2327.58273] [Medline: 20182578]
11. Howard MC. A meta-analysis and systematic literature review of virtual reality rehabilitation programs. Comput Human Behav 2017 May;70:317-327. [doi: 10.1016/j.chb.2017.01.013]
12. Cano Porras D, Siemonsma P, Inzelberg R, Zeilig G, Plotnik M. Advantages of virtual reality in the rehabilitation of balance and gait: systematic review. Neurology 2018 May 29;90(22):1017-1025. [doi: 10.1212/WNL.0000000000005603] [Medline: 29720544]
13. Maggio MG, Russo M, Cuzzola MF, Destro M, La Rosa G, Molonia F, et al. Virtual reality in multiple sclerosis rehabilitation: a review on cognitive and motor outcomes. J Clin Neurosci 2019 Jul;65:106-111. [doi: 10.1016/j.jocn.2019.03.017] [Medline: 30898488]
14. Webster A, Poyade M, Rooney S, Paul L. Upper limb rehabilitation interventions using virtual reality for people with multiple sclerosis: a systematic review. Mult Scler Relat Disord 2021 Jan;47:102610. [doi: 10.1016/j.msard.2020.102610] [Medline: 33212400]
15. Plow M, Finlayson M. A qualitative study exploring the usability of Nintendo Wii Fit among persons with multiple sclerosis. Occup Ther Int 2014 Mar;21(1):21-32 [FREE Full text] [doi: 10.1002/oti.1345] [Medline: 23613154]
16. Kristensson P, Matthing J, Johansson N. Key strategies for the successful involvement of customers in the co-creation of new technology-based services. Int J Serv Ind Manag 2008;19(4):474-491. [doi: 10.1108/09564230810891914]
17. Windasari NA, Visita L. User engagement mechanisms of online co-design service: does user innovativeness matter? Asian Acad Manag J 2019 Jun 27;24(1):59-82. [doi: 10.21315/aamj2019.24.1.3]
18. Schmid L, Glässel A, Schuster-Amft C. Therapists' perspective on virtual reality training in patients after stroke: a qualitative study reporting focus group results from three hospitals. Stroke Res Treat 2016;2016:6210508 [FREE Full text] [doi: 10.1155/2016/6210508] [Medline: 28058130]
19. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Dec 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]

20. Jiménez TR, Orozco M. Prompts, not questions: four techniques for crafting better interview protocols. Qual Sociol 2021 Jun 05;44(4):507-528. [doi: 10.1007/s11133-021-09483-2]
21. Webster A, Poyade M, Rea P, Paul L. The co-design of hand rehabilitation exercises for multiple sclerosis using hand tracking system. In: Rea PM, editor. Biomedical Visualisation. Cham, Switzerland: Springer; 2019:83-96.
22. Hollywood RA, Poyade M, Paul L, Webster A. Proof of concept for the use of immersive virtual reality in upper limb rehabilitation of multiple sclerosis patients. In: Rea PM, editor. Biomedical Visualisation. Volume 11. Cham, Switzerland: Springer; 2022:73-93.
23. Neale H, Nichols S. Theme-based content analysis: a flexible method for virtual environment evaluation. Int J Hum Comput Stud 2001 Aug;55(2):167-189. [doi: 10.1006/ijhc.2001.0475]
24. Junus IS, Santoso HB, Isal RY, Utomo AY. Usability evaluation of the student centered e-learning environment. Int Rev Res Open Dis Learn 2015 Nov 02;16(4):62-82. [doi: 10.19173/irrodl.v16i4.2175]
25. Latham B, Poyade M, Finlay C, Edmond A, McVey M. New tools in education: development and learning effectiveness of a computer application for use in a university biology curriculum. In: Rea PM, editor. Biomedical Visualisation. Volume 2. Cham, Switzerland: Springer; 2019:29-46.
26. Bertoni R, Lamers I, Chen CC, Feys P, Cattaneo D. Unilateral and bilateral upper limb dysfunction at body functions, activity and participation levels in people with multiple sclerosis. Mult Scler 2015 Oct 06;21(12):1566-1574. [doi: 10.1177/1352458514567553] [Medline: 25662346]
27. Cattaneo D, Lamers I, Bertoni R, Feys P, Jonsdottir J. Participation restriction in people with multiple sclerosis: prevalence and correlations with cognitive, walking, balance, and upper limb impairments. Arch Phys Med Rehabil 2017 Jul;98(7):1308-1315 [FREE Full text] [doi: 10.1016/j.apmr.2017.02.015] [Medline: 28336344]
28. Newsome SD, von Geldern G, Shou H, Baynes M, Marasigan RE, Calabresi PA, et al. Longitudinal assessment of hand function in individuals with multiple sclerosis. Mult Scler Relat Disord 2019 Jul;32:107-113 [FREE Full text] [doi: 10.1016/j.msard.2019.04.035] [Medline: 31085489]
29. Meadmore KL, Hallewell E, Freeman C, Hughes A. Factors affecting rehabilitation and use of upper limb after stroke: views from healthcare professionals and stroke survivors. Top Stroke Rehabil 2019 Mar 13;26(2):94-100. [doi: 10.1080/10749357.2018.1544845] [Medline: 30422096]
30. Dias P, Silva R, Amorim P, Lains J, Roque E, Pereira IS, et al. Using virtual reality to increase motivation in poststroke rehabilitation. IEEE Comput Grap Appl 2019 Jan 1;39(1):64-70. [doi: 10.1109/mcg.2018.2875630]
31. Pagliari C, Di Tella S, Jonsdottir J, Mendozzi L, Rovaris M, De Icco R, et al. Effects of home-based virtual reality telerehabilitation system in people with multiple sclerosis: a randomized controlled trial. J Telemed Telecare 2021 Dec 01;30(2):344-355. [doi: 10.1177/1357633x211054839]
32. Kamm CP, Blättler R, Kueng R, Vanbellingen T. Feasibility and usability of a new home-based immersive virtual reality headset-based dexterity training in multiple sclerosis. Mult Scler Relat Disord 2023 Mar;71:104525 [FREE Full text] [doi: 10.1016/j.msard.2023.104525] [Medline: 36738693]
33. Forsberg A, Nilsagård Y, Boström K. Perceptions of using videogames in rehabilitation: a dual perspective of people with multiple sclerosis and physiotherapists. Disabil Rehabil 2015;37(4):338-344 [FREE Full text] [doi: 10.3109/09638288.2014.918196] [Medline: 24833535]
34. Palacios-Ceña D, Ortiz-Gutierrez R, Buesa-Estellez A, Galán-Del-Río F, Cachon Perez JM, Martínez-Piedrola R, et al. Multiple sclerosis patients' experiences in relation to the impact of the kinect virtual home-exercise programme: a qualitative study. Eur J Phys Rehabil Med 2016 Jul;52(3):347-355 [FREE Full text] [Medline: 26883340]
35. LaViola Jr JJ. A discussion of cybersickness in virtual environments. SIGCHI Bull 2000 Jan 01;32(1):47-56. [doi: 10.1145/333329.333344]
36. Stanney K, Fidopiastis C, Foster L. Virtual reality is sexist: but it does not have to be. Front Robot AI 2020 Jan 31;7:4 [FREE Full text] [doi: 10.3389/frobt.2020.00004] [Medline: 33501173]
37. Kalron A, Frid L, Fonkatz I, Menascu S, Dolev M, Magalashvili D, et al. The design, development, and testing of a virtual reality device for upper limb training in people with multiple sclerosis: single-center feasibility study. JMIR Serious Games 2022 Oct 12;10(3):e36288 [FREE Full text] [doi: 10.2196/36288] [Medline: 36094809]
38. Kemeny A, Chardonnet JR, Colombet F. Getting Rid of Cybersickness: In Virtual Reality, Augmented Reality, and Simulators. Cham, Switzerland: Springer; 2020.
39. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 May 23;6:42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
40. Mouatt B, Smith AE, Mellow ML, Parfitt G, Smith RT, Stanton TR. The use of virtual reality to influence motivation, affect, enjoyment, and engagement during exercise: a scoping review. Front Virtual Real 2020 Dec 23;1. [doi: 10.3389/frvir.2020.564664]
41. Zimmerli L, Jacky M, Lünenburger L, Riener R, Bolliger M. Increasing patient engagement during virtual reality-based motor rehabilitation. Arch Phys Med Rehabil 2013 Oct;94(9):1737-1746. [doi: 10.1016/j.apmr.2013.01.029] [Medline: 23500181]
42. Charles D, Holmes D, Charles T. Virtual reality design for stroke rehabilitation. In: Rea PM, editor. Biomedical Visualisation. Volume 6. Cham, Switzerland: Springer; 2020:53-87.

43. Lorenz M, Brade J, Klimant P, Heyde C, Hammer N. Age and gender effects on presence, user experience and usability in virtual environments-first insights. PLoS One 2023 Mar 27;18(3):e0283565 [FREE Full text] [doi: 10.1371/journal.pone.0283565] [Medline: 36972245]
44. Tuena C, Pedroli E, Trimarchi PD, Gallucci A, Chiappini M, Goulene K, et al. Usability issues of clinical and research applications of virtual reality in older people: a systematic review. Front Hum Neurosci 2020 Apr 8;14:93 [FREE Full text] [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
45. Lohse K, Shirzad N, Verster A, Hodges N, Van der Loos HF. Video games and rehabilitation: using design principles to enhance engagement in physical therapy. J Neurol Phys Ther 2013 Dec;37(4):166-175. [doi: 10.1097/NPT.0000000000000017] [Medline: 24232363]
46. Richter G, Raban DR, Rafaeli S. Studying gamification: the effect of rewards and incentives on motivation. In: Reiners T, Wood LC, editors. Gamification in Education and Business. Cham, Switzerland: Springer; 2015:21-46.
47. Yildirim IG. Time pressure as video game design element and basic need satisfaction. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI EA '16; May 7-12, 2016; San Jose, CA p. 7-12 URL: https://dl.acm.org/doi/10.1145/2851581.2892298 [doi: 10.1145/2851581.2892298]
48. Dekker J, de Groot V, Ter Steeg AM, Vloothuis J, Holla J, Collette E, et al. Setting meaningful goals in rehabilitation: rationale and practical tool. Clin Rehabil 2020 Jan 18;34(1):3-12. [doi: 10.1177/0269215519876299] [Medline: 31530186]
49. Høeg E, Becermen B, Bruun-Pedersen JR, Serafin S. Co-creating virtual reality applications for motor rehabilitation with physiotherapists. In: Proceedings of the 8th EAI International Conference, ArtsIT 2019, and 4th EAI International Conference, DLI 2019 on Interactivity, Game Creation, Design, Learning, and Innovation. 2019 Presented at: ArtsIT DLI '19; November 6-8, 2019; Aalborg, Denmark p. 379-389 URL: https://link.springer.com/chapter/10.1007/978-3-030-53294-9_26 [doi: 10.1007/978-3-030-53294-9_26]
50. Pau M, Cocco E, Arippa F, Casu G, Porta M, Menascu S, et al. An immersive virtual kitchen training system for people with multiple sclerosis: a development and validation study. J Clin Med 2023 May 30;12(9):3222 [FREE Full text] [doi: 10.3390/jcm12093222] [Medline: 37176662]
51. Lee JH, Ku J, Cho W, Hahn WY, Kim IY, Lee S, et al. A virtual reality system for the assessment and rehabilitation of the activities of daily living. Cyberpsychol Behav 2003 Aug;6(4):383-388. [doi: 10.1089/109493103322278763] [Medline: 14511450]
52. Knierim P, Kosch T, Schwind V, Funk M, Kiss F, Schneegass S, et al. Tactile drones - providing immersive tactile feedback in virtual reality through quadcopters. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2017 Presented at: CHI EA '17; May 6-11, 2017; Denver, CO p. 6-11 URL: https://dl.acm.org/doi/10.1145/3027063.3050426 [doi: 10.1145/3027063.3050426]
53. Collins K, Kapralos B. Pseudo-haptics: leveraging cross-modal perception in virtual environments. Senses Soc 2019 Oct 14;14(3):313-329. [doi: 10.1080/17458927.2019.1619318]
54. Threapleton K, Drummond A, Standen P. Virtual rehabilitation: what are the practical barriers for home-based research? Digit Health 2016 Apr 29;2:2055207616641302 [FREE Full text] [doi: 10.1177/2055207616641302] [Medline: 29942551]
55. Clarke R, Coote S. Perceptions of participants in a group, community, exercise programme for people with multiple sclerosis. Rehabil Res Pract 2015;2015:123494-123497 [FREE Full text] [doi: 10.1155/2015/123494] [Medline: 26491567]
56. Arntzen EC, Straume BK, Odeh F, Feys P, Zanaboni P, Normann B. Group-based individualized comprehensive core stability intervention improves balance in persons with multiple sclerosis: a randomized controlled trial. Phys Ther 2019 Aug 01;99(8):1027-1038 [FREE Full text] [doi: 10.1093/ptj/pzz017] [Medline: 30722036]
57. Anderson-Hanley C, Snyder AL, Nimon JP, Arciero PJ. Social facilitation in virtual reality-enhanced exercise: competitiveness moderates exercise effort of older adults. Clin Interv Aging 2011;6:275-280 [FREE Full text] [doi: 10.2147/CIA.S25337] [Medline: 22087067]
58. Geurts E, Van Geel F, Feys P, Geurts E. WalkWithMe: personalized goal setting and coaching for walking in people with multiple sclerosis. In: Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization. 2019 Presented at: UMAP '19; June 9-12, 2019; Larnaca, Cyprus p. 51-60 URL: https://dl.acm.org/doi/10.1145/3320435.3320459 [doi: 10.1145/3320435.3320459]
59. Cuesta-Gómez A, Sánchez-Herrera-Baeza P, Oña-Simbaña ED, Martínez-Medina A, Ortiz-Comino C, Balaguer-Bernaldo-de-Quirós C, et al. Effects of virtual reality associated with serious games for upper limb rehabilitation inpatients with multiple sclerosis: randomized controlled trial. J Neuroeng Rehabil 2020 Jul 13;17(1):90 [FREE Full text] [doi: 10.1186/s12984-020-00718-x] [Medline: 32660604]

Abbreviations

ADL: activities of daily living
HMD: head-mounted device
MS: multiple sclerosis
NHS: National Health Service
SPMS: secondary progressive multiple sclerosis

TBCA: theme-based content analysis
UL: upper limb
VR: virtual reality

Edited by C Prahm; submitted 02.08.23; peer-reviewed by J Wu, C Srikesavan, L Bulle; comments to author 21.12.23; revised version received 14.02.24; accepted 14.03.24; published 26.04.24.

Please cite as:
Webster A, Poyade M, Coulter E, Forrest L, Paul L
Views of Specialist Clinicians and People With Multiple Sclerosis on Upper Limb Impairment and the Potential Role of Virtual Reality in the Rehabilitation of the Upper Limb in Multiple Sclerosis: Focus Group Study
JMIR Serious Games 2024;12:e51508
URL: https://games.jmir.org/2024/1/e51508
doi:10.2196/51508
PMID:38669680

Original Paper

A Virtual Reality Serious Game for the Rehabilitation of Hand and Finger Function: Iterative Development and Suitability Study

Michael Bressler[1], MSc; Joachim Merk[1], PhD; Tanja Gohlke[1]; Fares Kayali[2], Prof Dr; Adrien Daigeler[1], Prof Dr Med; Jonas Kolbenschlag[1], Prof Dr Med; Cosima Prahm[1], PhD

[1]BG Klinik Tuebingen, Clinic for Hand, Plastic, Reconstructive and Burn Surgery, University of Tuebingen, Tuebingen, Germany

[2]Institute for Teacher Education, University of Vienna, Vienna, Austria

Corresponding Author:
Michael Bressler, MSc
BG Klinik Tuebingen
Clinic for Hand, Plastic, Reconstructive and Burn Surgery
University of Tuebingen
Schnarrenbergstrasse 95
Tuebingen, 72076
Germany
Phone: 49 650 4432134
Email: michael.bressler.hprv@gmail.com

Abstract

Background: Restoring hand and finger function after a traumatic hand injury necessitates a regimen of consistent and conscientious exercise. However, motivation frequently wanes due to unchallenging repetitive tasks or discomfort, causing exercises to be performed carelessly or avoided completely. Introducing gamification to these repetitive tasks can make them more appealing to patients, ultimately increasing their motivation to exercise consistently.

Objective: This study aims to iteratively develop a serious virtual reality game for hand and finger rehabilitation within an appealing and engaging digital environment, encouraging patient motivation for at least 2 weeks of continuous therapy.

Methods: The development process comprised 3 distinct stages, each of which was subject to evaluation. Initially, a prototype was created to encompass the game's core functionalities, which was assessed by 18 healthy participants and 7 patients with impaired hand function. Subsequently, version 1 of the game was developed and evaluated with 20 patients who were divided into an investigation group and a control group. On the basis of these findings, version 2 was developed and evaluated with 20 patients who were divided into an investigation group and a control group. Motivation was assessed using the Intrinsic Motivation Inventory (IMI), while the application's quality was rated using the Mobile Application Rating Scale and the System Usability Scale. User feedback was gathered using semistructured interviews.

Results: The prototype evaluation confirmed the acceptance and feasibility of the game design. Version 1 significantly increased motivation in 2 IMI subscales, *effort* ($P<.001$) and *usefulness* ($P=.02$). In version 2, a significant increase in daily performed exercises was achieved ($P=.008$) compared to version 1, with significantly higher motivation in the IMI subscale *effort* ($P=.02$). High Mobile Application Rating Scale scores were obtained for both versions 1 and 2, with version 2 scoring 86.9 on the System Usability Scale, indicating excellent acceptability. User feedback provided by the semistructured interviews was instrumental in the iterative development regarding improvements and the expansion of the playable content.

Conclusions: This study presented a virtual reality serious game designed for hand and finger rehabilitation. The game was well received and provided an environment that effectively motivated the users. The iterative development process incorporated user feedback, confirming the game's ease of use and feasibility even for patients with severely limited hand function.

(JMIR Serious Games 2024;12:e54193) doi:10.2196/54193

KEYWORDS

video games; virtual reality; exercise therapy; physical therapy; hand rehabilitation; finger rehabilitation

Introduction

Background

Our hands are essential tools for managing daily life and are thus at high risk of injury. Therefore, comprehensive and successful rehabilitation to quickly restore hand function is essential for patients' quality of life and ability to work. A key factor in successful rehabilitation is to maintain patients' motivation to conscientiously participate in the process [1]. However, this is complicated by the fact that as part of their therapy, patients may experience pain reactions during or after exercise and must endure them. In addition, long-term repetitive activities are often monotonous and tend to be performed more and more carelessly without the supervision of an occupational therapist or physiotherapist [2]. However, to achieve the best possible outcome of the therapeutic process, it is necessary for patients to perform their exercises regularly, usually even daily [3,4].

In the last decade, serious games for health have become more popular and have shown a positive effect on the rehabilitation process [5,6]. The application of serious games for health covers a wide range of domains, such as training for behavioral change [7], cognitive exercises [8], the treatment of perceptual disorders [9], or physiotherapeutic exercising (eg, for pain [10] or multiple sclerosis [11,12]). The concept of gamification is an attempt to enrich a context, for example therapeutic exercising, with elements and principles used in game design [5,13]. The main target here is to positively influence the player's attitude, enjoyment, and perceived usefulness toward the game [14]. Gamification can also contribute to improving personal health behavior [15].

The use of video games allows for the creation of exciting adventures for patients who are experiencing certain limitations due to age, illness, or disabilities and can significantly improve their mood [16]. The highly immersive experience that can be generated by the application of virtual reality (VR) technology promises to increase the positive effects on the rehabilitation process even further. The term *virtual reality* in the context of rehabilitation is often used to describe any type of computer-based system, regardless of the level of immersion. Strictly speaking, however, VR refers to a system in which the viewer is surrounded by a computer-generated 3D environment and can move around in this artificial world in real time, view it from different angles, and interact with it [17]. The cost of such immersive systems dropped dramatically after 2013, for example, a 90° field-of-view head-mounted display (HMD) was US $35,000 in 2013 and US $600 in 2016, thus enabling affordable VR hand therapy [18].

Serious Games for Health Regarding Hand Rehabilitation

Input systems for the real-time capture of the patient's hand and finger movements presented in the literature range from haptic devices, such as joysticks [19], robots that allow the fingers to be moved in a targeted manner [20], and data gloves [21,22], to wearable inertial tracking devices [23] and optical tracking systems with either externally placed cameras, such as the Nintendo Wii [24], or low-cost, camera-based tracking systems (eg, the Leap Motion controller that can be used stationary in front of a screen [25,26] or mounted onto a VR HMD). An alternative built-in-one setup is provided by the Meta (formerly Oculus) Quest 2 HMD [12,27]. Such markerless optical tracking generally enables a very simple setup and is also especially beneficial for patients with severely injured skin, burns, or allodynia [18,28].

Many VR and non-VR applications designed for arm and hand rehabilitation can be found in the literature, for example, for the purpose of grasping exercises after-stroke rehabilitation [21] or dexterity training for multiple sclerosis [12]. For practicing the hand and fingers in particular, several examples are given for rehabilitative tasks and activities to be performed in VR, such as playing a virtual piano, catching butterflies, picking flower petals, or solving puzzles [29,30]. Furthermore, assessment-like tasks can be found in VR, such as stacking cylinders into a pegboard or stacking cubes [31,32]. Most of these examples lack a concept that motivates the patient to stick with the game for a longer period but rather rely on the effects of technological novelty and already existing intrinsic motivation. Some tasks provide concepts, such as point systems, eventually combined with playtime and connected to a leaderboard. These rather competitive game elements influence mainly extrinsic motivation and have little effect on intrinsic motivation [33].

Creating Motivational Serious Games for Health

Various types of players exist, and they differ in the degree to which they can be motivated by intrinsic or extrinsic motivation [34]. Although both are vital for engagement, games often focus on extrinsic motivators, such as rewards, achievements, or points, which can be harmful to intrinsic motivation [35]. Intrinsic motivation, by contrast, can be supported by self-initiation and choice [36]. Concepts from self-determination theory, such as competence, relatedness, and autonomy, can help create designs that provide sustaining engagement [37]. Game-based approaches related to self-determination theory can also be drawn from behavior change technology [38,39]. Other key factors in intrinsic motivation are informational feedback and clear game goals, which serve as proof of effectiveness for the patient and can also contribute to extending health beneficial behavior beyond the context of the game [40,41]. Work on gamification discusses the roles of intrinsic and extrinsic motivation, as both are important and not sufficiently studied empirically [33,42].

Objectives

The primary objective of this study was to iteratively design and evaluate a serious game for the rehabilitation of hand and finger function in a patient-centered approach. In contrast to hand rehabilitation games presented in the literature, StableHandVR (BG Klinik Tübingen) had a stronger focus on different motivational factors to promote sustained user engagement for a variety of player types. Similarly, the game was designed to be feasible even for patients with severely limited hand function. The secondary objective was to compare the motivational effects of the rehabilitation game across the design iterations and a control group.

Methods

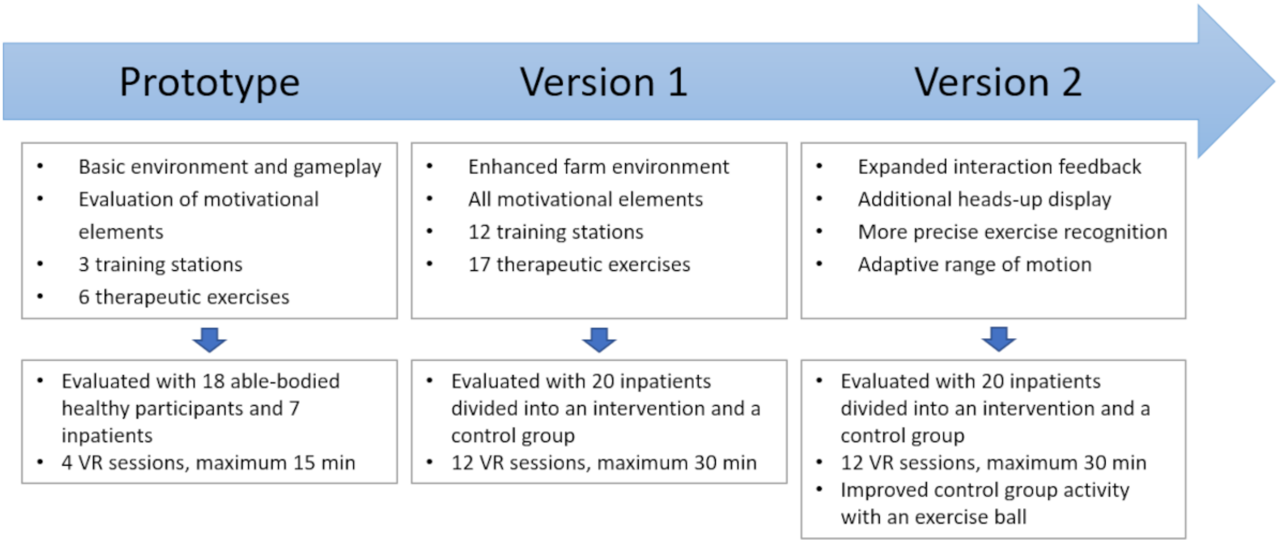
Ethical Considerations

Participant recruitment for the study was conducted in compliance with the Declaration of Helsinki and followed the ethical guidelines by the University of Tübingen, Germany. This study was also approved by the ethics committee of the University Clinic of Tübingen (470/2019B02). Before the initiation of the study, informed consent was obtained from all participants. All data used for this study was anonymized. No compensation was provided to the participants.

Study Design

This study presents the iterative design and development process of the serious game StableHandVR in 3 steps. First, a prototype was created, which provided the core game mechanics and was tested for usability and feasibility. In total, 4 game elements to maintain motivation were designed and evaluated; additional user feedback was collected. On the basis of this preliminary investigation, version 1 of the game was developed expanding the playable content to 3 weeks of training. An intervention group played the game, while a control group watched 360° videos in VR for 12 days during inpatient rehabilitation to evaluate motivational effects. Subsequently, based on the repeatedly gathered feedback and user observation, version 2 was developed and evaluated with the control group's activity being expanded to the use of a training ball to exercise the injured hand while watching the VR videos. Figure 1 presents an overview of the 3 development stages and their evaluation.

Figure 1. StableHandVR underwent 3 successive iterations of development and evaluation. VR: virtual reality.



Apparatus and Setup

All versions of the game were developed in Unity (version 2021, Unity Technologies) for the Meta Quest 2 HMD, running as stand-alone application. The game relied primarily on the use of the inbuilt optical finger and hand tracking feature of the Meta Quest 2; no controllers were required. A physiotherapist could optionally supervise by streaming the visual contents of the HMD onto an Android tablet. The study was conducted in treatment rooms at the hospital with an exercise area of approximately 2×3 meters.

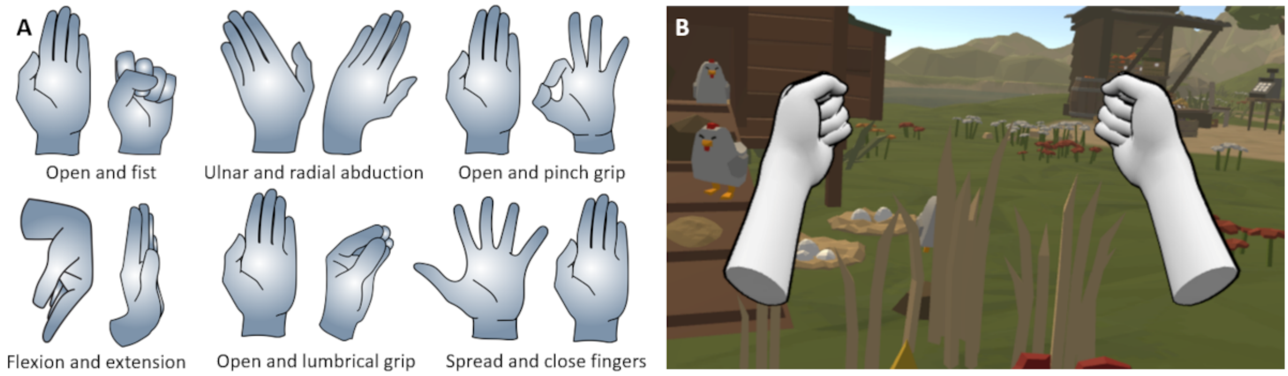
Basic Game Design

StableHandVR aimed to transfer traditional physiotherapy hand and finger exercises into an immersive and motivating virtual world. The inspiration for placing StableHandVR in a natural environment was derived from a study conducted in the 1980s on patients undergoing postoperative recovery [43]. An early feasibility study coined the setting of the game to be a farm environment that provided several training stations to perform exercises [27]. Each station of StableHandVR included a specific task (eg, feeding and milking the cows, preparing a meal, or repairing a tractor), and its completion was divided into 6 exercises, each to be repeated 10 times. Therefore, 60 exercise repetitions had to be performed at each station, and the station's environment would adapt with each repetition, according to the task at hand.

Therapeutic Exercises

The exercises integrated into the serious game were selected by a peer group of physiotherapists and are based on conventional hand mobility therapy [4]. They involved hand and finger movements, wrist movements, and forearm rotations. The prototype included a set of 6 basic exercises, such as closing the hand into a fist or gripping for holding a book (Figure 2).

Figure 2. (A) Overview of the 6 exercises that were used in the prototype. (B) Preview hands as seen in version 2 demonstrated each exercise to the user.



In version 1, the number of exercises was expanded to a total of 17 different movements (a complete list is presented in Multimedia Appendix 1). Throughout the game, the difficulty of the exercises progressively increased. This was achieved by incorporating compound movements, such as simultaneously closing the hand into a fist while pronating the wrist. Moreover, each movement provided the option to be performed with both hands moving in synchrony or in opposite directions, thus adding a further level of complexity and skill requirement to the gameplay.

Performing the exercises neither involved direct interaction with the environment, such as plucking flower petals with the fingertips [44], nor a direct transfer of the patient's hand movements to control the environment [45]. Instead, the environment would adapt automatically at each successful repetition of the exercise, according to the respective task of the station. This design decision was made to ensure good optical tracking of the hands by always being positioned to be clearly visible to the cameras. Furthermore, the original exercise should not be falsified or complicated by being combined with a virtual interaction. Moreover, this design allowed a dynamic composition of exercises for each station on each training day, adjustable for every patient.

In the prototype, preview hands were introduced that would appear in front of the player to demonstrate and clarify the requested hand movements at the beginning of each exercise. The preview hands disappeared after 2 complete repetitions. In version 2, the 3D hand model was expanded by a forearm to provide a more comprehensive visualization of the exercises containing a rotation of the wrist. In addition, instead of dark gray hands, the color was changed to a lighter gray, and the outlines of the hands were highlighted to enhance visibility (Figure 2).

Exercise Tracking and Dynamic Range Adjustment

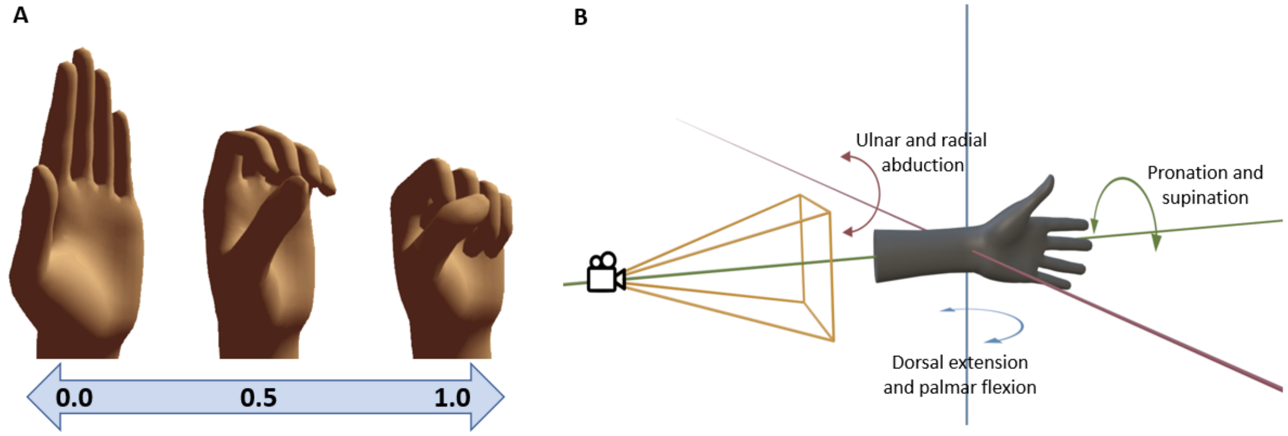
For the prototype, a dedicated component was developed to define and track exercise movements. This component used the hand position model supplied by the Meta Quest software development toolkit. In detail, it made use of the flexion angles of the finger joints within the provided hierarchical bone model to store and reproduce hand positions. By using the HMD, it was then possible to record various hand positions, such as an open hand or a closed fist, and subsequently use these stored positions to define exercise movements by specifying a respective start and end position as well as optional middle positions.

Due to the absence of a forearm in the tracking model of the Quest software development toolkit, the direct extraction of wrist rotation angles was not available. To compensate for this limitation, a reference coordinate system was used in replacement of a forearm bone to determine the rotation of the wrist. This coordinate system had its origin at the player's wrist and was spanned by the vertical axis of the VR environment and a forward axis based on the player's view direction and the forward direction of the hands, leveled within the VR environment by setting its vertical component to 0 (Figure 3). To accurately measure wrist rotation, it was necessary for the player to keep their arms bent forward during the exercise.

In addition to the interface for defining hand and finger exercises, the component was also able to track their execution. Therefore, the players' hand positions were compared to the specified exercise and expressed as a floating-point number within the interval (ie, 0,1). In this representation, 0 denoted the start position, and 1 indicated the end position. If the hand position deviated from the movement, the position was represented as –1.

For each exercise that was newly introduced, the game initially measured the range of motion (ROM) achieved by the player as represented within (0-1). This range was subsequently used to set a minimum target for the player to be exceeded when exercising. In version 2, an adaptive approach was implemented, where the target ROM was recalculated daily, based on the average ROM achieved in the preceding days. The tracking accuracy was also enhanced in version 2 by adding an additional middle position (Figure 3) to all exercise definitions, thus providing a more precise mapping of the player's ROM and adjustment to their skill.

Figure 3. (A) For each exercise movement, a start and end position were defined to determine the position of the hands within the interval (0,1) during exercising. In the second version, additionally a middle position was defined for more accurate measurement. (B) As Meta Quest 2 does not provide wrist tracking, its rotation had to be determined using a reference coordinate system. This coordinate system had its origin at the player's wrist and was spanned by the vertical axis (blue) and the leveled view direction (green). The third axis (red) was defined as perpendicular to the plane spanned by the first 2. This required the forearms to be held approximately along the view direction while exercising.



Interaction Design

Outside of the exercises, the player also used their hands to interact with the game (Figure 4). In the prototype, a teleportation system was created that enabled the player to switch between stations through predefined teleport points by pointing at them. In version 1, a waypoint network was established to allow the player to explore the farm environment also beyond the stations. Starting from version 1, the player was accompanied by a dog character that would provide guidance when touched. The dog's advice was displayed as text within a speech bubble. Nonplayable characters (NPCs) would respond to the player in a similar manner. In version 2, several improvements were implemented regarding the interaction. The teleport system was enhanced by introducing a navigation arrow with the purpose of guiding the player toward the next task. The touch interactions with the dog as well as with NPCs were improved by providing audiovisual feedback. All stations were fully supplemented with audio feedback during exercising; furthermore, an exercise counter was added that would display the number of remaining repetitions during an exercise.

Figure 4. Examples of in-game interaction: (A) petting the dog, (B) teleporting, (C) interaction with a nonplayable character, and (D) performing exercises at the Fireplace station.

Motivational Game Elements

Overview

In total, 4 game elements were developed to sustain patient motivation over the course of a 3-week treatment. These different elements were designed to cater to both intrinsic and extrinsic motivation, ensuring a wide range of motivational factors. The design prioritized preventing patient frustration resulting from limited hand movement and potential therapy-related discomfort, such as pain. At the same time, the game aimed to provide a challenge to less restricted or more competitive players, ensuring they remain engaged without getting bored. Furthermore, there should be no incentive to perform the exercises sloppily, for example, a time challenge. These motivational elements underwent first evaluation in the prototype and were further refined in versions 1 and 2.

Storytelling

The farm was populated with NPCs, which would provide the player with daily tasks, for example, to gather carrots from the vegetable field (Figure 5). In the prototype, initially only 1 NPC was implemented. However, starting from version 1, the farm was populated with 7 NPCs who assigned the player 2 daily tasks, each involving exercises at specific stations. Once the player completed these daily tasks, they gained access to exercise at all the other unlocked stations.

Figure 5. (A) A nonplayable character providing a daily task for the player. (B) The assessment station as seen in version 2 gave an overview of the player's progress. (C) Traffic Light Hands indicate that the movements were performed well. Yellow or green color indicated that the player had reached or exceeded their personal limit. As in version 2, this limit was adjusted daily according to their previous performance. (D) From version 1, the farm contained 12 stations that were unlocked over the course of the game.

Unlocking Rewards

Over the course of the game, the player was rewarded with additional exercise stations. In the prototype, the player unlocked a third station over the first 3 days to become playable on the fourth day. Starting from version 1, the player was rewarded with stations on each day after fulfilling their daily tasks, thus subsequently revealing all 12 stations over the course of the game (Figure 5).

Traffic Light Hands

Different hand colors were used to provide immediate feedback on the execution of movements. This was intended to guide the player toward performing terminal and correct movements while encouraging them to push their personal limit. A red hand color indicated that the patient reached their initially measured ROM, and the game recognized this as a successful repetition. As the patient continued to exceed this ROM, the hand color transitioned from red to orange and finally to green. In version 2, this system was adjusted daily based on the patient's ROM (Figure 5).

Scoring

In the prototype, the player received score points while exercising, based on the aforementioned achieved ROM. A highscore board placed in the middle of the farm presented the scores for each exercise. In version 1, the scoring system was omitted, and in version 2, an assessment station was introduced. This station provided players with visual feedback on their ROM progress throughout the training period for each exercise (Figure 5).

Evaluation

Overview

For all evaluations of StableHandVR, in-house patients from the hospital were recruited. These patients were undergoing inpatient rehabilitation because of limited hand function to such an extent that it restricted their professional and everyday activities. An overview of the type of injuries is given in Table 1. The patients took part in the VR sessions as part of their daily therapy schedule. Prerequisites for participation were basic mobility of the hand (no paralysis or total stiffness); limited hand function; no severe pain at rest (≥9 on a scale of 0-10); and the hand had to be free of stabilizing structures, such as splints, casts, or Kirschner wires. Inclusion criterion for all participants was a minimum age of 18 years.

Table 1. An overview of the injuries of the inpatients that were recruited for the evaluation of the prototype (n=7), version 1 (n=20), and version 2 (n=20).

Type of injury	Prototype, n (%)	Version 1, n (%)	Version 2, n (%)
Fractures in the wrist and hand area	3 (43)	11 (55)	6 (30)
Crush injuries or soft tissue injuries in the hand area	1 (14)	2 (10)	3 (15)
Tendon injuries in the area of the hand	0 (0)	2 (10)	1 (5)
Dislocations or ligament injuries in the area of the hand	1 (14)	1 (5)	1 (5)
Combination of the above points	2 (28)	4 (20)	9 (45)

Evaluation of the Prototype

The evaluation of the prototype involved 1 group of 25 participants consisting of 18 (72%) able-bodied individuals and 7 (28%) inpatients from the hospital (women: n=14, 56%; men: n=11, 25%. The age of the participants ranged from 18 to 56 years, with a mean age of 30.68 (SD 13.3) years. Among the 25 participants, 9 (36%) had prior experience with VR.

Each participant underwent 4 VR sessions conducted over 4 consecutive weekdays, with each session limited to 15 minutes. During each session, participants completed 1 exercise station and had the option to voluntarily complete a second one. In addition, only 1 of the 4 possible motivational game elements (Storytelling, Unlocking Rewards, Traffic Light Hands, and Scoring) was active during each of the sessions. In total, 3 motivational elements were evaluated in a randomized order over the first 3 sessions. In the fourth session, the player was consistently rewarded with access to a third exercise station, thus representing the fourth element, Unlocking Rewards.

After each session, participants rated their experience using 3 scales from the Intrinsic Motivation Inventory (IMI) questionnaire [46], specifically *interest and enjoyment*,*effort* and *pressure* on a 7-point Likert scale. In addition, participants were interviewed to gather their feedback on the game, suggestions for improvements, and ideas for additional content. The open-ended answers were evaluated based on the grounded theory [47] and thematic analysis [48].

Evaluation of Version 1

In total, 20 inpatients (women: n=6, 30%; men: n=14, 70%) from the hospital were equally assigned to either intervention or control group. The age ranged from 24 to 70 years, with a mean age of 48.8 (SD 12.3) years. Of the 20 inpatients, 8 (40%) had prior experience with VR.

During their 3-week inpatient rehabilitation program, both groups completed 12 VR sessions on consecutive weekdays in addition to their regular rehabilitation therapy. Each VR session was limited to 30 minutes and was supervised by a physiotherapist. In the intervention group, patients played the VR game and completed 2 mandatory tasks in each session. Additional training stations that were already unlocked could be voluntarily explored. New exercises were introduced every fourth day. In the control group, patients used the VR headset to watch a 360° video during each session, with durations ranging from 10 to 15 minutes. Following the final session, both groups were surveyed using the *interest and enjoyment*, *effort*, *usefulness*, and *pressure* subscales of the IMI questionnaire. The intervention group also evaluated the VR game using the Mobile Application Rating Scale (MARS) questionnaire [49], with the scales *engagement*, *functionality*, *aesthetics* and *impact on knowledge and attitudes* rated on a 5-point Likert scale, and was interviewed to gather feedback on the game. For further analysis of the user behavior, the game automatically recorded all user interactions in a time log.

Evaluation of Version 2

In total, 20 inpatients from the hospital were equally assigned to either the intervention or control group. One patient dropped out after the second session in the intervention group and another after the third session in the control group. The intervention group dropout was due to a dislike for the game, while the control group dropout was due to a transfer to another hospital. Both dropouts were replaced by 2 additional patients. All following analyses refer to the 18 remaining original patients and the 2 replacements.

The age of the final 20 patients (women: n=10, 50%; men: n=10, 50%) ranged from 22 to 61 years, with a mean age of 38.1 (SD 12.9) years; Of the 20 inpatients, 5 (25%) had prior experience with VR. The evaluation of version 2 followed a similar approach as version 1, with 2 modifications to the test protocol. First, the activity of the control group was extended by incorporating a crumple ball exercise for patients to engage their injured hand while watching the VR content. This addition aimed to provide an unspecific exercise for the injured hand. Second, patients in the intervention group had to rate the game using the System Usability Scale (SUS) [50] on a 5-point Likert scale after the last session. This measure was introduced to gather validated feedback on the usability of the game.

Data Analysis

All statistical analyses were performed with MATLAB for Windows (version R2021a; MathWorks) with a significance level of α=.05. To assess demographic effects, patients from the evaluations of versions 1 and 2 were consolidated and categorized into 5 age groups (<30, 30-39, 40-49, 50-59, and >60 years), into a group of women or men, and a VR-experienced versus no VR-experience group. All data gathered from the automatic tracking were tested for normal distribution using the Kolmogorov-Smirnov test. Due to the small sample sizes, nonparametric tests were used for all analyses. For multiple pair-wise comparisons, the Kruskal-Wallis test was used with the *P* level adjusted by Bonferroni correction. For single pair-wise comparisons, the Wilcoxon rank sum test was used. Demographic data, temporal data, and variables gathered from questionnaires were represented as mean (SD); the number of stations per day and the duration of exercise repetitions, gathered from the automatically tracked in-game data were represented as mean (SE).

Results

Prototype Results

All 25 participants completed the 4 VR sessions, and none of them reported vertigo or discomfort at any point. Over the course of the 4 sessions, all participants learned to operate the game without assistance. The mean VR playtime for each session decreased from 14.5 (SD 0.65) minutes for the first session to 10.6 (SD 0.82) minutes for the last session. All participants were able to perform the 6 exercises in a way that the game could recognize them. The mean duration needed to perform 1 repetition of an exercise decreased from the first to the fourth session from 7.1 (SD 0.35) to 6 (SD 0.38) seconds for the patients and from 5.6 (SD 0.26) to 4.7 (SD 0.24) seconds for the able-bodied participants.

The evaluation of the game elements with the IMI questionnaire resulted in high scores for all 4 elements on the 2 subscales *interest and enjoyment* and *effort* and low scores on the *pressure* subscale (Multimedia Appendix 2). There were neither substantial differences between the elements nor over the course of the 4 days. When asked about their most favored motivational element in the interviews, the Traffic Light Hands and Unlocking Rewards both were mentioned most frequently (8 mentions), followed by Storytelling (6 mentions) and Scoring (3 mentions), which lacked significance for many participants, as they always achieved the full ROM and therefore the maximum number of points.

The overall feedback gathered from the interviews was highly positive; the farm setting displayed in the game was widely regarded as pleasant and appealing (10 mentions). The direct feedback of the Traffic Light Hands (6 mentions), the interactive aspects of the station environments that would adapt during the execution of the exercises (5 mentions), the animals on the farm including the dog companion (4 mentions), and the general idea of gamifying a rather boring rehabilitation activity (4 mentions) were also positively mentioned. The least favored experiences were the recognition of the exercises (4 mentions), and the low number of exercise stations (4 mentions). The most mentioned suggestions for improvements were more variety in general (6 mentions), more exercise stations (4 mentions), enhanced storytelling (2 mentions), more variety in exercise movements (2 mentions), and a larger farm area to explore (2 mentions). When asked about suggestions for additional content, many suggestions were made for additional training stations related to farm work, mostly regarding animals (7 mentions) but also regarding the farm infrastructure (4 mentions), such as the farmhouse or the tractor, and activities regarding the lake (3 mentions).

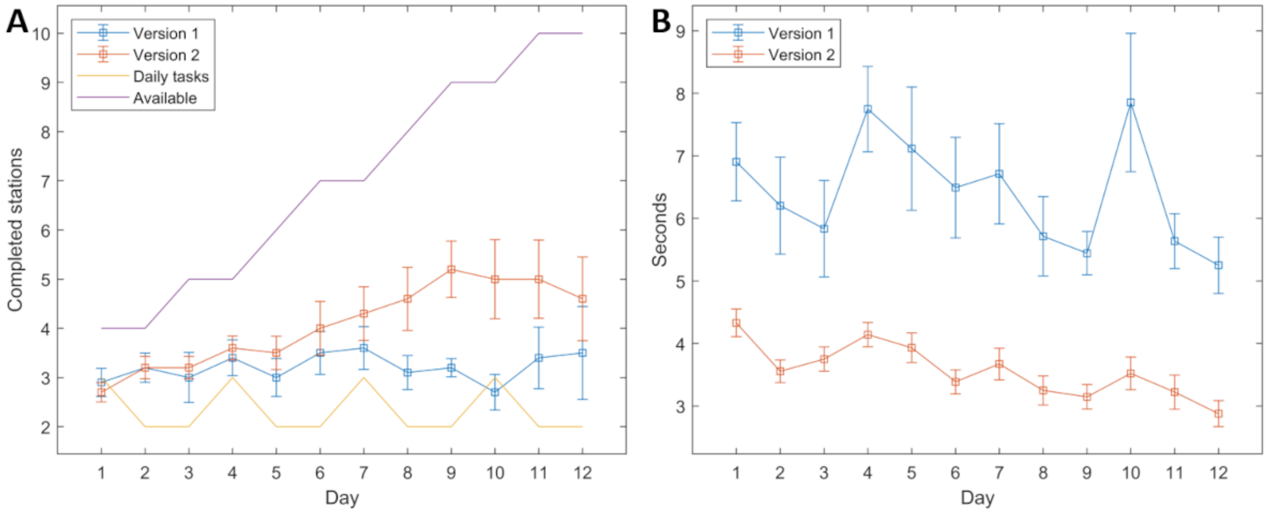
Version 1 Results

All 20 patients in both groups completed the 12 VR sessions; no patient reported any experience of discomfort or motion sickness. The mean VR playtime of the intervention group was 25.3 (SD 5.9) minutes on the first day and 26.4 (SD 10.6) minutes on the last day, with a mean playtime of 22.8 (SD 7.1) minutes over all days. The mean number of stations played per day was higher than the mandatory amount specified by the daily tasks on all days except for the 1st and the 10th day (Figure 6). Due to their limited hand function, some patients had problems performing the exercises such that they were recognized by the game, especially on the 10th day when the last set of exercises was introduced. The mean duration of the execution of 1 exercise repetition increased strongly on the days when new exercises were introduced (Figure 6) but decreased over the course of the 12 days from 6.9 (SE 0.62) seconds to 5.3 (SE 0.45) seconds.

The MARS rating resulted in high scores for all subscales: *engagement* (mean 4.18, SD 0.42), *functionality* (mean 4.17, SD 0.47), *aesthetics* (mean 4.2, SD 0.48), and *impact on knowledge and attitudes* (mean 4.7, SD 0.35; Multimedia Appendix 3). All IMI subscale scores were slightly higher in the intervention group than in the control group (Multimedia Appendix 4); significant differences were found for the subscales *effort* (*P*<.001) and *usefulness* (*P*=.02).

The user observation indicated that most patients could navigate the game independently without assistance after the initial 2 days, except when new exercises were introduced. The visual and textual instructions provided by the preview hands and the dog companion were occasionally unclear, requiring clarification from the assisting physical therapist. Some patients had difficulties with orienting themselves on the farm, especially on the first days. At the training stations, it was not always clear for some patients where the action was taking place to signify task completion. Patients mentioned difficulties regarding exercise counting, and they expressed the need for a visible counter and more auditory feedback at the stations to signal when a repetition was completed.

Figure 6. (A) The mean number and SE of completed stations in version 1 and version 2 as well as the minimum amount demanded by the daily tasks and the maximum possible number of stations for each day. (B) The mean duration and SE that were required to execute 1 repetition of 1 exercise in version 1 and version 2. New exercises were introduced on days 1, 4, 7, and 10.



Both groups reported perceiving the VR sessions as a vacation from their inpatient stay. They were able to momentarily forget about their injuries, felt being transported to another place, and experienced a sense of tranquility. The feedback from the intervention group regarding their game experience again was highly positive. The Traffic Light Hands were frequently mentioned for motivating patients to extend their limits toward terminal movements (6 mentions) and providing feedback on correct exercises (4 mentions); the immersive scenario, which made patients lose track of time during practice (6 mentions), and the virtual representation of their hand, which made them forget their injury (4 mentions), were also mentioned frequently. Some participants expressed their wish for an option to observe their progress, for example, as a score system that was not present in version 1.

Version 2 Results

A total of 20 patients completed the 12 VR sessions, of whom 19 (95%) did not report any discomfort or motion sickness and 1 (5%) answered with "don't know." The mean VR playtime of the intervention group was 20.3 (SD 5.9) minutes on the first day and 16.7 (SD 8.5) minutes on the last day, with an overall mean of 20.9 (SD 8.9) minutes on all days. The mean number of stations played was higher than in version 1 after the second day (Figure 6), and the overall mean of stations played was significantly higher (P=.008). All patients were able to perform the exercises in a way that was recognized by the game. The mean duration for performing 1 exercise repetition decreased from 4.6 (SE 0.24) seconds on the first day to 2.6 (SE 0.23) seconds on the last day. On the days with new movements being introduced, the increases were not as pronounced as in version 1 (Figure 6). The overall mean duration for performing 1 exercise repetition was significantly lower than that in version 1 (P<.001).

Compared to version 1, the MARS rating resulted in slightly higher scores for the subscales *engagement* (mean 4.24, SD 0.71) and *functionality* (mean 4.28, SD 0.42) and in slightly lower scores for the subscales *aesthetics* (mean 4.0, SD 0.63) and *impact on knowledge and attitudes* (mean 4.14, SD 0.73; Multimedia Appendix 3). The SUS rating resulted in a mean score of 86.9 (SD 3.3), which ranges in the fourth quartile and represents excellent acceptability. The scores of the IMI subscales were higher for the intervention group (Multimedia Appendix 4) than for the control group, with significant differences for the subscale *effort* (P=.02).

Similar to version 1, the user observation indicated that patients were able to operate the game without assistance after the first 2 days but required some assistance in learning new exercise movements. Orientation on the farm and at the stations was comprehensible for all patients, and performing the exercises caused the patients less difficulty than in version 1. Both groups reported the relaxing effect of the VR experience, and the most mentioned categories regarding the overall feedback from the interviews were similar to version 1: the Traffic Light Hands (5 mentions), the immersive game experience (5 mentions), and the virtual representation of the hands (3 mentions). Furthermore, the execution of the exercises was mostly described as working well (5 mentions). The newly designed Scoring element, providing a progress overview of the ROM at the assessment station, experienced the same issues as in the prototype, namely, always showing the possible maximum of points for the most exercises. Therefore, it was described as not very meaningful and was predominantly not used. Suggestions were made to provide other scenarios, such as a dungeon- or sci-fi–themed environment. Patients also raised the wish for a more personalized experience, and suggestions were made, for example, custom paint for the farmhouse or customizable virtual hands. Finally, recommendations for activities regarding the lake were mentioned again, such as fishing or riding a boat.

Demographic Effects

Regarding age, the 40- to 49-year age group showed the lowest number of completed stations among all age groups as well as the lowest IMI scores for *interest and enjoyment*, *effort*, and *usefulness*, while high scores for *pressure* were reported. The 30- to 39-year age group completed the most stations, and the <30-year age group reported the highest IMI scores for *interest and enjoyment*, *effort*, and *usefulness* and the lowest *pressure* score. Men played more stations than women patients and accordingly had higher IMI scores for *interest and enjoyment*, *effort*, and *usefulness*, with a lower *pressure* score. Patients with no prior VR experience completed more stations per day than patients with prior VR experience. However, the latter rated higher IMI scores for *interest*, *effort*, and *usefulness* but also a higher score for the *pressure* scale. The complete list of values is presented in Table 2.

Table 2. Number of stations completed per day and Intrinsic Motivation Inventory scores (1-7 scale)[a].

Demographic group	Stations per day, mean (SE)	Interest and Enjoyment, mean (SD)	Effort, mean (SD)	Pressure, mean (SD)	Usefulness, mean (SD)
Gender					
Men (n=11)	3.89 (0.6)	6.4 (0.6)	6.1 (1.0)	2.0 (0.6)	6.5 (0.6)
Women (n=9)	3.34 (0.5)	6.3 (1.1)	6.0 (1.0)	1.9 (0.7)	6.3 (1.2)
VR[b] Experience					
Prior VR experience (n=4)	3.04 (0.5)	6.4 (0.9)	6.1 (0.6)	2.2 (0.7)	6.5 (0.5)
No prior VR experience (n=16)	3.79 (0.6)	6.3 (0.8)	6.0 (1.0)	1.9 (0.7)	6.4 (0.9)
Age group (y)					
<30 (n=2)	3.58 (0.3)	6.6 (0.6)	7.0 (0)	1.5 (0.7)	6.9 (0.2)
30-39 (n=4)	4.98 (1.5)	6.0 (0.6)	5.8 (0.7)	1.7 (0.4)	6.3 (0.6)
40-49 (n=3)	2.97 (0.9)	5.9 (2.0)	5.7 (1.5)	2.1 (0.7)	5.5 (2.1)
50-59 (n=8)	3.08 (0.3)	6.6 (0.5)	6.0 (1.0)	2.1 (0.8)	6.8 (0.2)
>60 (n=3)	4.05 (1.1)	6.3 (0.3)	6.3 (0.7)	2.0 (0.3)	6.1 (0.4)

[a]The data were consolidated from version 1 (n=10) and version 2 (n=10) and categorized by demographic groups.

[b]VR: virtual reality.

Discussion

Principal Findings

In this study, we iteratively developed a serious health game for hand and finger rehabilitation with the deliberate goal of contributing to the long-term engagement of patients. The game was developed for the Meta Quest 2 because this HMD with built-in finger tracking allowed for a very simple and fast setup. A second reason for this decision was the finger tracking of the device, which produced better results than the Leap Motion or UltraLeap controllers when performing certain finger positions that were considered important, for example, the thumb touching ≥1 long fingers.

The iterative development and evaluation steps ensured a patient-centered design process by user observation and feedback through semistructured interviews. In general, participants reported a high level of immersion that allowed them to temporarily escape from their inpatient setting into another world. This was reported by both the intervention group and control groups and impressively confirms the potential of VR as described in the literature [18,51]. The setting of the farm environment and the display of nature were overall received very positively, a design decision that was also inspired by literature [43]. However, in the original intervention group of version 2, one patient quit after the second session because they did not like the game in general. In terms of engagement, we observed the entire spectrum from patients who tended to be underchallenged to patients well within their physical and cognitive limitations. We attribute these differences not only to varying levels of hand function but also to contrasting user preferences (eg, purpose vs mastery) [34], possibly coined by the practice of playing video games.

We expected a decrease in motivation related to the higher age of the participants and thus less familiarity with video games [52]. The familiarity was reflected by the IMI subscale *pressure*, which was clearly lowered for the <30- and 30- to 39-year age groups. However, while the 40- to 49-year and 50- to 59-year age groups completed the fewest stations per day, the >60-year age group reached the second highest value, only surpassed by the 30- to 39-year age group and before the <30-year age group. Differences in the number of stations completed between genders are smaller than between VR and no VR experience. The other IMI scales *interest*, *effort*, and *usability* show only minor deviations for all demographic groups. From these findings, it could be concluded that while age, gender, or the effect of technological novelty due to using VR for the first time might affect the overall motivation, these demographic effects did not harm the intrinsic motivation of the patients.

Compared to their control groups, higher IMI scores were measured for both intervention groups, and significant differences were found for the subscales *effort* in version 1 and version 2 and *usefulness* in version 1. It should be mentioned that, surprisingly the control groups were also very motivated to watch the videos, which again demonstrates the potential of using VR in the context of rehabilitation. However, we assume that in an unsupervised setting, the motivation in the control groups would have decreased more compared to the intervention groups [2] because the game, as used in our study, still represented a form of supervision. The assurance of performing the therapeutic exercises correctly may have been a key factor for the significantly higher IMI subscale *usefulness* of the intervention group, as scored in version 1 [41]. According to the MARS questionnaire, awareness for conscious health behavior could be raised for the intervention groups, which is linked to perceived attitude, enjoyment, and especially usefulness [14,15].

To engage a wide range of player types, 4 motivational game elements offering different motivational factors were developed. The element Storytelling should provide tasks and a context, Unlocking Rewards should offer new incentives for discovery, the Traffic Light Hands would provide direct feedback, and the Score would provide long-term feedback [34]. Among these 4, the first 3 elements can be considered rather successful, as they scored high on the IMI questionnaire and the subscales *effort* and partially also *usability* were significantly higher than in the control groups. In particular, the Traffic Light Hands received positive feedback in the interviews. This direct biofeedback, which is a critical factor for intrinsic motivation [41,53], not only encouraged patients to perform terminal movements but also gave them confidence to exercise correctly.

In contrast, the Scoring element already proved to be unsuccessful in the prototype stage because it lacked meaningfulness in terms of showing the patient's ROM improvement. This was mainly caused by the fact that many players were able to perform most exercises with full range from the beginning. Due to the ambiguity, the element was discontinued in version 1 to avoid negative feedback and discouragement. Due to the frequent demand to visually represent the progression of ROM, the element was reintroduced in version 2 with a different design, which did not work properly for the same reasons. Therefore, the resulting lack of competitive motivational elements might be the main reason why the game was not challenging enough for some patients [34]. Furthermore, the lack of visualization of the therapy progress deprived patients of evidence of efficacy, which is also a key motivating factor for health games [41].

The comparison of versions 1 and version 2 shows a significantly higher number of stations played per day, from which it can be concluded that the overall motivation in version 2 was higher. Besides the improved interaction feedback, also reflected by a high SUS score, the most obvious improvement was the tracking accuracy of the therapeutic exercises, which can clearly be seen by the mean time that was required for a single exercise repetition being reduced by almost half. This circumstance may be mainly responsible for the fact that more stations were played. However, this interpretation must also consider that the mean age of the patients in version 2 was 10 years lower than in version 1. In both versions, a similar amount of time was spent in the game, on average approximately 22 minutes. As there were no significant differences in the IMI scores between version 1 and version 2, we conclude that the improvements did not affect the intrinsic motivation, which was already high in version 1.

Limitations

Patients attended the VR sessions after their daily rehabilitation routine and may have been exhausted, which could potentially affect their motivation negatively. By contrast, we can assume that by being part of their daily routine, the supervised participation in VR sessions, although voluntary, was accompanied by stronger motivation than if it had been unsupervised [2]. The large variation in the degree of injury and impairment between individual patients may have influenced the outcome of measured and observed motivation. A larger sample size would better compensate for this effect.

Due to the short evaluation time of only 12 days, it was only partially possible to measure the long-term course of motivation.

Outlook and Future Challenges

While we acknowledge the challenges in making the game equally appealing and challenging for all patients, our approach of incorporating a mix of extrinsic and intrinsic motivational elements was generally successful. By individually adjusting the composition of the movements, the challenge level of the game could be easily increased for underachieving players. It has also become apparent that there is a great need for a working score system or at least a therapy progress indicator. To build up compliance and adherence among patients, this proof of effectiveness is a key factor and needs to be improved. Due to the missing forearm tracking and inaccuracies in detecting individual phalanges [54], especially within an unusual finger or hand position, the correct execution of the patient's movements and the assessment of their ROM was limited. We expect this feature to improve over the next few years, enabling us to overcome these limitations. In the meantime, a point system could be implemented that is simply based on the number of exercises already completed.

In the context of the general shortage of therapists, it would be beneficial to use StableHandVR in an unsupervised setting, for example, at home for several weeks, following the inpatient stay. The evaluation of StableHandVR in an unsupervised setting for an extended period would also be an exciting next step from the perspective of this study. In such a setting, we see the biggest challenge for our game, both in terms of user engagement and in substituting the physical therapist, whose assistance is currently still required for correctly learning new exercises.

Finally, StableHandVR could easily be transferred into other domains where hand and finger exercises are required, such as in stroke rehabilitation [21] or multiple sclerosis [12]. As StableHandVR allows for the simple creation of additional therapeutic exercise movements, it could be extended within a short time for other motor exercises. Furthermore, we think that the promising use of different motivational elements as in StableHandVR would also be beneficial for a variety of other applications that use gamification in a therapeutic context to achieve user engagement.

Conclusions

This study showcased a VR game designed for hand and finger rehabilitation exercises. The iterative development process allowed user feedback to be incorporated into further development. The game was well received, offering an engaging environment and various elements that effectively motivated the users. Despite impaired hand function, the tracking of therapeutic movements proved to be reliable in operating the game. The high SUS score confirms the ease of use of the game, even for patients with physical limitations. With ongoing technical advancements in optical finger tracking, we anticipate even greater accuracy in the future, paving the way for automated medical assessments and telerehabilitation scenarios. This creates the potential for StableHandVR to become an unsupervised yet engaging VR health game for postrehabilitation home use.

Acknowledgments

Funding for this project was provided by the German Social Accident Insurance (Deutsche Gesetzliche Unfallversicherung [DGUV]; FR237). The authors thank student Dominik Lammer for his valuable assistance in designing and implementing the prototype and version 1.

Authors' Contributions

MB contributed to the game design, implemented all versions, contributed to the evaluation, analyzed the data, and wrote the manuscript. JM contributed to the game design, contributed to the evaluation, and revised the manuscript. TG contributed to the game design, performed the experiments, collected the data, and revised the manuscript. FK contributed to the game design, contributed to the interpretation of the results, and revised the manuscript. AD and JK contributed to the critical revision of the intellectual content and approved the final version. CP contributed to the methodology and the game design, revised the manuscript, and oversaw overall direction and planning. All authors provided critical feedback and helped shape the research, analysis, and manuscript. All authors contributed to the article and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the therapeutic exercises.
[PDF File (Adobe PDF File), 708 KB - games_v12i1e54193_app1.pdf]

Multimedia Appendix 2
Intrinsic Motivation Inventory (IMI) scores per game element in the prototype. During the prototype evaluation, participants rated their experience for each game element using the IMI questionnaire. Error bars indicate the SD of the mean.
[PNG File , 59 KB - games_v12i1e54193_app2.png]

Multimedia Appendix 3
The mean scores and SD for version 1 and 2 rated by the intervention groups using the Mobile Application Rating Scale questionnaire.
[PNG File , 86 KB - games_v12i1e54193_app3.png]

Multimedia Appendix 4
Mean scores and SD of the evaluated Intrinsic Motivation Inventory scales of the intervention group compared to the control group for version 1 (A) and version 2 (B). Statistically significant at *P*<.05.
[PNG File , 54 KB - games_v12i1e54193_app4.png]

References

1. Maclean N, Pound P, Wolfe C, Rudd A. Qualitative analysis of stroke patients' motivation for rehabilitation. BMJ 2000 Oct 28;321(7268):1051-1054 [FREE Full text] [doi: 10.1136/bmj.321.7268.1051] [Medline: 11053175]
2. Minetama M, Kawakami M, Teraguchi M, Kagotani R, Mera Y, Sumiya T, et al. Supervised physical therapy vs. home exercise for patients with lumbar spinal stenosis: a randomized controlled trial. Spine J 2019 Aug;19(8):1310-1318. [doi: 10.1016/j.spinee.2019.04.009] [Medline: 30986577]
3. Foley N, McClure JA, Meyer M, Salter K, Bureau Y, Teasell R. Inpatient rehabilitation following stroke: amount of therapy received and associations with functional recovery. Disabil Rehabil 2012;34(25):2132-2138. [doi: 10.3109/09638288.2012.676145] [Medline: 22524794]
4. Duncan SF, Flowers CW. Therapy of the Hand and Upper Extremity. Cham, Switzerland: Springer; 2015.
5. Kato PM. Video games in health care: closing the gap. Rev Gen Psychol 2010 Jun 01;14(2):113-121. [doi: 10.1037/A0019441]

6. Primack BA, Carroll MV, McNamara M, Klem ML, King B, Rich M, et al. Role of video games in improving health-related outcomes: a systematic review. Am J Prev Med 2012 Jun;42(6):630-638 [FREE Full text] [doi: 10.1016/j.amepre.2012.02.023] [Medline: 22608382]
7. Terlouw G, Kuipers D, van't Veer J, Prins JT, Pierie JP. The development of an escape room-based serious game to trigger social interaction and communication between high-functioning children with autism and their peers: iterative design approach. JMIR Serious Games 2021 Mar 23;9(1):e19765 [FREE Full text] [doi: 10.2196/19765] [Medline: 33755023]
8. Chang CH, Yeh CH, Chang CC, Lin YC. Interactive somatosensory games in rehabilitation training for older adults with mild cognitive impairment: usability study. JMIR Serious Games 2022 Jul 14;10(3):e38465 [FREE Full text] [doi: 10.2196/38465] [Medline: 35834303]
9. Stammler B, Flammer K, Schuster T, Lambert M, Karnath H. Negami: an augmented reality app for the treatment of spatial neglect after stroke. JMIR Serious Games 2023 Feb 27;11:e40651 [FREE Full text] [doi: 10.2196/40651] [Medline: 36848215]
10. Beltran-Alacreu H, Navarro-Fernández G, Godia-Lledó D, Graell-Pasarón L, Ramos-González Á, Raya R, et al. A serious game for performing task-oriented cervical exercises among older adult patients with chronic neck pain: development, suitability, and crossover pilot study. JMIR Serious Games 2022 Feb 01;10(1):e31404 [FREE Full text] [doi: 10.2196/31404] [Medline: 35103608]
11. Kalron A, Frid L, Fonkatz I, Menascu S, Dolev M, Magalashvili D, et al. The design, development, and testing of a virtual reality device for upper limb training in people with multiple sclerosis: single-center feasibility study. JMIR Serious Games 2022 Sep 12;10(3):e36288 [FREE Full text] [doi: 10.2196/36288] [Medline: 36094809]
12. Kamm CP, Blättler R, Kueng R, Vanbellingen T. Feasibility and usability of a new home-based immersive virtual reality headset-based dexterity training in multiple sclerosis. Mult Scler Relat Disord 2023 Mar;71:104525 [FREE Full text] [doi: 10.1016/j.msard.2023.104525] [Medline: 36738693]
13. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
14. Baptista G, Oliveira T. Gamification and serious games: a literature meta-analysis and integrative model. Comput Human Behav 2019 Mar;92:306-315. [doi: 10.1016/j.chb.2018.11.030]
15. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
16. Russoniello CV, O'Brien K, Parks JM. The effectiveness of casual video games in improving mood and decreasing stress. J Cyber Ther Rehabil 2009;2(1):53-66 [FREE Full text]
17. Tieri G, Morone G, Paolucci S, Iosa M. Virtual reality in cognitive and motor rehabilitation: facts, fiction and fallacies. Expert Rev Med Devices 2018 Feb 10;15(2):107-117. [doi: 10.1080/17434440.2018.1425613] [Medline: 29313388]
18. Hoffman HG, Boe DA, Rombokas E, Khadra C, LeMay S, Meyer WJ, et al. Virtual reality hand therapy: a new tool for nonopioid analgesia for acute procedural pain, hand rehabilitation, and VR embodiment therapy for phantom limb pain. J Hand Ther 2020 Apr;33(2):254-262 [FREE Full text] [doi: 10.1016/j.jht.2020.04.001] [Medline: 32482376]
19. Ganjiwale D, Pathak R, Dwivedi A, Ganjiwale J, Parekh S. Occupational therapy rehabilitation of industrial setup hand injury cases for functional independence using modified joystick in interactive computer gaming in Anand, Gujarat. Natl J Physiol Pharm Pharmacol 2018:1. [doi: 10.5455/njppp.2019.9.0620211201800l]
20. Liu S, Meng D, Cheng L, Huang F. A virtual reality based training and assessment system for hand rehabilitation. In: Proceedings of the 9th International Conference on Intelligent Control and Information Processing. 2018 Presented at: ICICIP '18; November 9-11, 2018; Wanzhou, Chaina p. 33-38 URL: https://ieeexplore.ieee.org/document/8606705 [doi: 10.1109/ICICIP.2018.8606705]
21. da Silva Cameirão M, Bermúdez I Badia S, Duarte E, Verschure PF. Virtual reality based rehabilitation speeds up functional recovery of the upper extremities after stroke: a randomized controlled pilot study in the acute phase of stroke using the rehabilitation gaming system. Restor Neurol Neurosci 2011;29(5):287-298. [doi: 10.3233/RNN-2011-0599] [Medline: 21697589]
22. Gabyzon ME, Engel-Yeger B, Tresser S, Springer S. Using a virtual reality game to assess goal-directed hand movements in children: a pilot feasibility study. Technol Health Care 2016;24(1):11-19. [doi: 10.3233/THC-151041] [Medline: 26409528]
23. Yang X, Yeh SC, Niu J, Gong Y, Yang G. Hand rehabilitation using virtual reality electromyography signals. In: Proceedings of the 5th International Conference on Enterprise Systems. 2017 Presented at: ES '17; September 22-24, 2017; Beijing, China p. 125-131 URL: https://ieeexplore.ieee.org/document/8119378 [doi: 10.1109/es.2017.27]
24. Standen P, Threapleton K, Richardson A, Connell L, Brown D, Battersby S, et al. A low cost virtual reality system for home based rehabilitation of the arm following stroke: a randomised controlled feasibility trial. Clin Rehabil 2017 Mar 10;31(3):340-350 [FREE Full text] [doi: 10.1177/0269215516640320] [Medline: 27029939]
25. Alimanova M, Borambayeva S, Kozhamzharova D, Kurmangaiyeva N, Ospanova D, Tyulepberdinova G, et al. Gamification of hand rehabilitation process using virtual reality tools: using leap motion for hand rehabilitation. In: Proceedings of the 1st IEEE International Conference on Robotic Computing. 2017 Presented at: IRC '17; April 10-12, 2017; Taichung, Taiwan p. 336-339 URL: https://ieeexplore.ieee.org/document/7926560 [doi: 10.1109/irc.2017.76]

26. Charles D, Pedlow K, McDonough S, Shek K, Charles T. Close range depth sensing cameras for virtual reality based hand rehabilitation. J Assist Technol 2014 Sep 09;8(3):138-149. [doi: 10.1108/JAT-02-2014-0007]
27. Pereira MF, Prahm C, Kolbenschlag J, Oliveira E, Rodrigues NF. A virtual reality serious game for hand rehabilitation therapy. In: Proceedings of the 2020 IEEE 8th International Conference on Serious Games and Applications for Healt. 2020 Presented at: SeGAH '20; August 12-14, 2020; Vancouver, BC p. 1-7 URL: https://ieeexplore.ieee.org/document/9201789 [doi: 10.1109/segah49190.2020.9201789]
28. Wu YT, Chen KH, Ban SL, Tung KY, Chen LR. Evaluation of leap motion control for hand rehabilitation in burn patients: an experience in the dust explosion disaster in Formosa Fun Coast. Burns 2019 Feb;45(1):157-164. [doi: 10.1016/j.burns.2018.08.001] [Medline: 30322737]
29. Shin JH, Kim MY, Lee JY, Jeon YJ, Kim S, Lee S, et al. Effects of virtual reality-based rehabilitation on distal upper extremity function and health-related quality of life: a single-blinded, randomized controlled trial. J Neuroeng Rehabil 2016 Feb 24;13(1):17 [FREE Full text] [doi: 10.1186/s12984-016-0125-x] [Medline: 26911438]
30. Then JW, Shivdas S, Tunku Ahmad Yahaya TS, Ab Razak NI, Choo PT. Gamification in rehabilitation of metacarpal fracture using cost-effective end-user device: a randomized controlled trial. J Hand Ther 2020 Apr;33(2):235-242. [doi: 10.1016/j.jht.2020.03.029] [Medline: 32430167]
31. Postolache O, Lourenço F, Dias Pereira JM, Girão P. Serious game for physical rehabilitation: measuring the effectiveness of virtual and real training environments. In: Proceedings of the 2017 IEEE International Instrumentation and Measurement Technology Conference. 2017 Presented at: I2MTC '17; May 22-25, 2017; Turin, Italy p. 1-6 URL: https://ieeexplore.ieee.org/abstract/document/7969978 [doi: 10.1109/i2mtc.2017.7969978]
32. Tang HK, Feng ZQ, Xu T, Yang XH. VR system for active hand rehabilitation training. In: Proceedings of the 4th International Conference on Information, Cybernetics and Computational Social System. 2017 Presented at: ICCSS '17; July 24-26, 2017; Dalian, China p. 316-320 URL: https://ieeexplore.ieee.org/document/8091432 [doi: 10.1109/iccss.2017.8091432]
33. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput Human Behav 2017 Jun;71:525-534. [doi: 10.1016/j.chb.2015.08.048]
34. Tondello GF, Wehbe RR, Diamond L, Busch M, Marczewski A, Nacke LE. The gamification user types hexad scale. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play. 2016 Presented at: CHI PLAY '16; October 16-19, 2016; Austin, TX p. 229-243 URL: https://dl.acm.org/doi/10.1145/2967934.2968082 [doi: 10.1145/2967934.2968082]
35. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol Bull 1999 Nov;125(6):627-692. [doi: 10.1037/0033-2909.125.6.627] [Medline: 10589297]
36. Biddiss E, Irwin J. Active video games to promote physical activity in children and youth: a systematic review. Arch Pediatr Adolesc Med 2010 Jul 05;164(7):664-672. [doi: 10.1001/archpediatrics.2010.104] [Medline: 20603468]
37. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
38. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Dec;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
39. Cheek C, Fleming T, Lucassen MF, Bridgman H, Stasiak K, Shepherd M, et al. Integrating health behavior theory and design elements in serious games. JMIR Ment Health 2015 Apr 21;2(2):e11 [FREE Full text] [doi: 10.2196/mental.4133] [Medline: 26543916]
40. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
41. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
42. Seaborn K, Fels DI. Gamification in theory and action: a survey. Int J Hum Comput Interact 2015 Feb;74:14-31. [doi: 10.1016/j.ijhcs.2014.09.006]
43. Ulrich RS. View through a window may influence recovery from surgery. Science 1984 Apr 27;224(4647):420-421. [doi: 10.1126/science.6143402] [Medline: 6143402]
44. Wang Z, Wang P, Xing L, Mei L, Zhao J, Zhang T. Leap Motion-based virtual reality training for improving motor functional recovery of upper limbs and neural reorganization in subacute stroke patients. Neural Regen Res 2017 Nov;12(11):1823-1831 [FREE Full text] [doi: 10.4103/1673-5374.219043] [Medline: 29239328]
45. Jiang Y, Li Z, He M, Lindlbauer D, Yan Y. HandAvatar: embodying non-humanoid virtual avatars through hands. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 1-17 URL: https://dl.acm.org/doi/10.1145/3544548.3581027 [doi: 10.1145/3544548.3581027]

46. Reynolds L. Measuring intrinsic motivations. In: Reynolds RA, Woods R, Baker JD, editors. Handbook of Research on Electronic Surveys and Measurements. Hershey, PA: IGI Global; 2007:170-173.
47. Adams A, Lunt P, Cairns P. A qualitative approach to HCI research. In: Cairns P, Cox AL, editors. Research Methods for Human-Computer Interaction. Cambridge, MA: Cambridge University Press; 2008:138-157.
48. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
49. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the mobile application rating scale (MARS). PLoS One 2020 Nov 2;15(11):e0241480 [FREE Full text] [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
50. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:207-212.
51. Pereira MF, Prahm C, Kolbenschlag J, Oliveira E, Rodrigues NF. Application of AR and VR in hand rehabilitation: a systematic review. J Biomed Inform 2020 Nov;111:103584 [FREE Full text] [doi: 10.1016/j.jbi.2020.103584] [Medline: 33011296]
52. Koivisto J, Hamari J. Demographic differences in perceived benefits from gamification. Comput Human Behav 2014 Jun;35:179-188. [doi: 10.1016/j.chb.2014.03.007]
53. Kayali F, Luckner N, Purgathofer P, Spiel K, Fitzpatrick G. Design considerations towards long-term engagement in games for health. In: Proceedings of the 13th International Conference on the Foundations of Digital Games. 2018 Presented at: FDG '18; August 7-10, 2018; Malmö, Sweden p. 1-8 URL: https://dl.acm.org/doi/10.1145/3235765.3235789 [doi: 10.1145/3235765.3235789]
54. Abdlkarim D, Di Luca M, Aves P, Maaroufi M, Yeo S, Miall RC, et al. A methodological framework to assess the accuracy of virtual reality hand-tracking systems: a case study with the Meta Quest 2. Behav Res Methods 2024 Feb 13;56(2):1052-1063 [FREE Full text] [doi: 10.3758/s13428-022-02051-8] [Medline: 36781700]

Abbreviations

HMD: head-mounted display
IMI: Intrinsic Motivation Inventory
MARS: Mobile Application Rating Scale
NPC: nonplayable character
ROM: range of motion
SUS: System Usability Scale
VR: virtual reality

Edited by R Baranyi; submitted 01.11.23; peer-reviewed by T Vanbellingen, R Patterson; comments to author 22.04.24; revised version received 20.05.24; accepted 03.06.24; published 27.08.24.

Please cite as:
Bressler M, Merk J, Gohlke T, Kayali F, Daigeler A, Kolbenschlag J, Prahm C
A Virtual Reality Serious Game for the Rehabilitation of Hand and Finger Function: Iterative Development and Suitability Study
JMIR Serious Games 2024;12:e54193
URL: https://games.jmir.org/2024/1/e54193
doi:10.2196/54193
PMID:39190432

Original Paper

Measuring the Reliability of a Gamified Stroop Task: Quantitative Experiment

Katelyn Wiley[1], MEDes; Phaedra Berger[1], BSc, BA⪼ Maximilian Achim Friehs[2,3,4], PhD; Regan Lee Mandryk[5], PhD

[1]Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada

[2]Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

[3]School of Psychology, University College Dublin, Dublin, Ireland

[4]Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[5]Department of Computer Science, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Katelyn Wiley, MEDes
Department of Computer Science
University of Saskatchewan
110 Science Place
University of Saskatchewan
Saskatoon, SK, S7N 5C9
Canada
Phone: 1 5878884567
Email: katelyn.wiley@usask.ca

Abstract

Background: Few gamified cognitive tasks are subjected to rigorous examination of psychometric properties, despite their use in experimental and clinical settings. Even small manipulations to cognitive tasks require extensive research to understand their effects.

Objective: This study aims to investigate how game elements can affect the reliability of scores on a Stroop task. We specifically investigated performance consistency within and across sessions.

Methods: We created 2 versions of the Stroop task, with and without game elements, and then tested each task with participants at 2 time points. The gamified task used points and feedback as game elements. In this paper, we report on the reliability of the gamified Stroop task in terms of internal consistency and test-retest reliability, compared with the control task. We used a permutation approach to evaluate internal consistency. For test-retest reliability, we calculated the Pearson correlation and intraclass correlation coefficients between each time point. We also descriptively compared the reliability of scores on a trial-by-trial basis, considering the different trial types.

Results: At the first time point, the Stroop effect was reduced in the game condition, indicating an increase in performance. Participants in the game condition had faster reaction times (P=.005) and lower error rates (P=.04) than those in the basic task condition. Furthermore, the game condition led to higher measures of internal consistency at both time points for reaction times and error rates, which indicates a more consistent response pattern. For reaction time in the basic task condition, at time 1, $r_{\text{Spearman-Brown}}$=0.78, 95% CI 0.64-0.89. At time 2, $r_{\text{Spearman-Brown}}$=0.64, 95% CI 0.40-0.81. For reaction time, in the game condition, at time 1, $r_{\text{Spearman-Brown}}$=0.83, 95% CI 0.71-0.91. At time 2, $r_{\text{Spearman-Brown}}$=0.76, 95% CI 0.60-0.88. Similarly, for error rates in the basic task condition, at time 1, $r_{\text{Spearman-Brown}}$=0.76, 95% CI 0.62-0.87. At time 2, $r_{\text{Spearman-Brown}}$=0.74, 95% CI 0.58-0.86. For error rates in the game condition, at time 1, $r_{\text{Spearman-Brown}}$=0.76, 95% CI 0.62-0.87. At time 2, $r_{\text{Spearman-Brown}}$=0.74, 95% CI 0.58-0.86. Test-retest reliability analysis revealed a distinctive performance pattern depending on the trial type, which may be reflective of motivational differences between task versions. In short, especially in the incongruent trials where cognitive conflict occurs, performance in the game condition reaches peak consistency after 100 trials, whereas performance consistency drops after 50 trials for the basic version and only catches up to the game after 250 trials.

Conclusions: Even subtle gamification can impact task performance albeit not only in terms of a direct difference in performance between conditions. People playing the game reach peak performance sooner, and their performance is more consistent within and across sessions. We advocate for a closer examination of the impact of game elements on performance.

(JMIR Serious Games 2024;12:e50315) doi:10.2196/50315

KEYWORDS
cognitive assessment; gamification; serious games; Stroop task; reliability

Introduction

Background

In 1886, James Cattell observed that it takes people longer to name the colors and pictures of objects than it does for them to read the corresponding word [1]. This experiment, along with others, paved the way for the development of what Cattell would call *mental tests* and what we now call *cognitive tasks*. On the basis of these and other results, JR Stroop developed a test of cognitive ability in which study participants read the color but not the meaning of a color word aloud [2]. The results revealed an interference effect if the word color and word meaning did not match. Typical cognitive tasks require people to respond to such visual or auditory cues, and data about their responses, often reaction time and accuracy, are collected. These data can then be used to study human cognition, create population norms, and inform medical decisions, such as dementia diagnoses [3].

Cognitive tasks are most useful when collecting high-quality, high-quantity data. However, this is a challenging process. Traditionally, capturing large data sets has been time consuming and expensive, requiring highly trained professionals to administer and score tasks with individual participants. With technological advancements, tasks can now be administered via computers, deployed remotely, and automatically scored [4,5]. This automation makes it easier to collect large quantities of data but raises new concerns about data quality. Many factors influence cognitive test performance beyond cognitive capacity, such as motivation, stereotype threat, and fatigue [6,7]. Cognitive tasks are often repetitive and boring, leading to high attrition rates [8] and suboptimal effort from participants [9,10].

In attempts to improve the quality of data collected by such tasks, researchers have increasingly turned to gamification, with the hope that tasks can be made more engaging through the addition of game elements, such as points and graphics.

Cognitive Task Gamification

Overview

Deterding et al [11] defined gamification as “the use of game design elements in nongame contexts.” In the context of cognitive tasks, this process typically involves layering game elements over an already existing task. For example, the Go No-Go task has commonly been gamified by adding points [12], narrative elements [13], and fun graphics [14] to the basic task.

Enjoyment and Motivation

Typically, tasks are gamified with the intent of increasing participant enjoyment and motivation. Nicholson [15] noted that gamification can target both extrinsic and intrinsic motivations depending on the game elements used. Reward-based elements, such as points, achievements, and badges, target extrinsic motivation, whereas elements such as play, exposition, and choice target intrinsic motivation. By targeting motivation, researchers aim to combat attrition and encourage repeated, prolonged play [16-18].

However, there is little examination of whether participants experience increased enjoyment when tasks are gamified. In a systematic review of gamified attention tasks, only 25 of the 74 studies reported results from an evaluation of gameplay [16]. When enjoyment is measured, the research shows mixed results. Some studies have found that gamification increases motivation; for example, participants in a stop signal task study experienced higher enjoyment and more flow-like experiences in the gamified condition (as opposed to the basic task) [19].

Other studies have found that certain game elements, especially thematic or narrative elements, can have a negative effect on self-reported enjoyment of cognitive tasks [8,20,21], possibly due to the “chocolate-covered broccoli” effect [22]. Tasks can only be gamified and retain the important elements of a task. When participants expect a fun game and must still complete a repetitive cognitive task, they may experience even lower enjoyment than if they expected a boring task [20]. Game elements can also be used to introduce other emotions. For example, Levy et al [23] found that some older Jewish participants were uncomfortable with their cooking-themed game as they required making recipes containing pork products.

Do these mixed findings imply that researchers should move away from gamifying tasks? Not necessarily, participants might not *enjoy* assessment games more than a control task, but the data they produced may still be of higher quality.

Performance

Groening and Binnewies [24] note that enjoyment is only one way to operationalize motivation, one closely linked to intrinsic motivation. They found that adding achievement-based game elements to a series of simple tasks did not improve self-reported motivation but did improve persistence—when participants could earn achievements, they engaged with a Stroop task for longer before voluntarily switching tasks, compared with when no achievements were available. Similarly, Mekler et al [25] found that when they gamified an image annotation task, participants generated significantly more annotations, despite no reported differences in intrinsic motivation or competence need satisfaction when compared with the basic task.

Adding game elements to a task may improve performance (without affecting enjoyment) in various ways. For example, Jung et al [26] compared the performance of participants who were given a numeric goal (ie, generating 22 ideas) with those who were asked to “do their best.” Participants who were given a specific goal generated higher quantity and higher quality responses. When completing cognitive tasks, participants are often instructed to respond “as quickly and accurately as possible.” This nebulous goal can be clarified and reinforced through game elements that provide immediate feedback such as scoring points for fast reactions or losing points for incorrect responses.

When designing gamified tasks for research and assessment purposes, it may be beneficial to focus on influencing performance rather than on enjoyment. Levy et al [23] noted that changes in emotions can influence cognitive abilities, which may interfere with the collection of valid and reliable data when using games as scientific tools. When Vanden Abeele et al [27] compared 2 games designed to measure psychoacoustic thresholds in preschoolers, they found that the more fully developed and motivating game was able to detect lower thresholds. As another example, Delisle and Braun [28] found that changing a task to resemble a fast-paced videogame normalized the performance of participants with attention-deficit/hyperactivity disorder (ADHD), meaning that participants with and without ADHD performed similarly on a gamified task (but differently on a standard task). In some cases, such an effect may be desired, but it depends on why the task is used and gamified.

Psychometric Properties of Gamified Tasks

Tasks may also be gamified with the goal of improving the psychometric properties of a task, such as validity (how well a task measures what it claims to measure) and reliability (how consistent the measurement obtained by the task is) [29]. There are also different types of evidence for reliability that must be considered when gamifying cognitive tasks. Internal consistency refers to the stability of the task data within an assessment; for example, the similarity of a participant's reaction time at the beginning of a task to their reaction time at the end of the task. Test-retest reliability refers to the stability of the task data over time; for example, how similar a participant's score on a task is at one time point compared with their score on the task a month later.

Typical cognitive tasks are boring, repetitive, and long partly because of the issue of reliability. From one trial to the next, people will perform quite differently, so multiple trials are needed to decrease measurement noise [30]. Adding game elements to a task may change the reliability of its measurement. Participants may be sufficiently engaged that their performance is more stable over time; for example, perhaps only 20 trials are needed for a reliable measure, instead of 200. Friehs et al [19] found that response variability in a gamified stop signal task was lower than that in the nongame version. Shorter tasks would require fewer resources to administer and would reduce the burden on participants, which would be particularly beneficial for clinical and pediatric populations.

Game elements also offer the ability to guide participants' performance. Most cognitive tasks use measures of reaction time and accuracy, which leads to classic speed-accuracy trade-offs—the faster a participant responds, the less accurate they will be, and vice versa. Individual participants also favor speed or accuracy differently than one another [30]. These behaviors can be manipulated through instructions (eg, asking participants to respond as quickly as possible). Game elements can also indirectly encourage participants to emphasize speed or accuracy, for example, by awarding points or feedback for faster or more accurate responses, generating more consistency across participants [30,31].

This Study

Overview

Few gamified cognitive tasks are subjected to rigorous examination of psychometric properties [16], despite their use in experimental and clinical settings. Parsons et al [32] noted that psychology lacks a standard practice of reporting the reliability of cognitive task measurements. This problem is exacerbated when tasks are adapted, such as gamification. Even small manipulations of cognitive tasks require extensive research to understand their effects [33].

In this study, we sought to research how game elements can affect the reliability of scores on a cognitive task, specifically the Stroop task. As a typical cognitive task that demonstrates robust experimental effects in the general population [34], the Stroop task is well suited for this research.

The Stroop Task

Building on the 1886 work by Cattell [1] with cognitive tasks, in 1935, Stroop [2] conducted an experiment in which he asked participants to either name the colors of colored rectangles or name the colors of mismatched words (eg, the word "blue" printed in red ink). Participants responded much more slowly when naming incongruent colored words, a paradigm we now call the Stroop effect [2].

Since Stroop's first experiment and subsequent development of the experimental protocol [35-37], the Stroop task has become one of the most widely used tasks in both cognitive and clinical psychology [34,38]. Recently, the Stroop task has been gamified for experimental and clinical applications. For example, Groening and Binnewies [39] used the Stroop task to investigate the effects of game elements on participants' motivation and performance. They found that when points and story elements were added to the task, participants were more persistent (they engaged with the task for longer before switching to a new task) and reported higher motivation. Gomez-Tello et al [40] used gamified tasks as part of a battery of tests for neuropsychological screening of children and found evidence of the Stroop effect in a gamified version of the task. However, previous studies have not considered the reliability of the Stroop effect in a gamified task, either in terms of internal consistency or test-retest reliability. Thus, we have little guidance when gamified tasks can or should not be used in assessments.

We created 2 versions of the Stroop task, with and without game elements, and tested each task with participants at 2 time points. In this paper, we report on the reliability of the gamified Stroop task in terms of internal consistency and test-retest reliability, compared with the control task. We also compared the reliability of these scores on a trial-by-trial basis. Our objective was to demonstrate how game elements can affect the reliability of scores on a Stroop task.

Methods

Ethical Considerations

This research project was approved on ethical grounds by the University of Saskatchewan Research Ethics Board (BEH 17-418). The participants were given GBP £6 (USD $8.3 at time of study) compensation at each time point.

Tasks

The control task was designed using the basic computerized Stroop task described by Macleod [34] and Hedge et al [41] as models. Participants were shown words in the middle of their screen in various colors (red, blue, green, or yellow). The word could be the same as the font color (congruent condition), a noncolor word (lot, ship, cross, or advice; neutral condition), or a nonmatching color word (eg, the word "blue" shown in green; incongruent condition). After each word, participants were asked to press a key corresponding to the font color (z-key for red, x-key for blue, n-key for green, and m-key for yellow). The participants first completed a training exercise to learn each keymap. The task consisted of 240 trials in each condition (congruent, neutral, and incongruent) for a total of 720 trials.

The gamified version was designed to increase reliability by manipulating the speed-accuracy trade-off [30] and improving engagement through game elements. On the basis of prior research, which demonstrated increased enjoyment from points and decreased enjoyment from themes added to a gamified task [20], we focused on adding points-based game elements to the Stroop task. Points-based elements also target extrinsic motivation (rather than intrinsic motivation), which may be more effective in influencing participant performance [24]. We followed the feedback category of the Gameful Design Heuristics from Tondello et al [42], which states that the system should offer users clear and immediate feedback, actionable feedback, and graspable progress.

Using feedback also allowed us to manipulate the speed-accuracy trade-off by preferentially awarding points for faster (but still correct) answers. In the game version of our task, participants saw their response time for each trial and whether they answered correctly. A record of the fastest response time was also displayed at the corner of the screen. They lost 5 points for any incorrect answer, gained 5 points for any correct answer, and were rewarded with a bonus of 25 points for responses that broke their previous "fastest time" record. A progress bar at the bottom of the screen tracked the points (Figure 1).

Figure 1. Game version of the task after a correct response was entered.

Participants

Participants were recruited through Prolific, a web-based platform for recruiting research participants. Web-based platforms are commonly used in human-computer interaction research to conduct studies [43] and have been shown to yield reliable data when precautionary methods for data gathering and analysis are used [44,45]. Each participant completed either the control task or the gamified task at 2 time points, 3 weeks apart (time 1 and time 2). The participants signed a consent form, were given instructions and training for the task, and then completed the task. After completion, they answered questionnaires collecting demographic information, including information about their experience with the task (Intrinsic Motivation Inventory [46]), their general gaming behavior, and self-reported attentional control (Attentional Control Scale [47]).

The study design was between-subjects, with half the participants completing the control version of the task and the other half completing the points version. The participants were randomly assigned to a condition. The study took approximately 40 minutes to complete.

Our analyses were based on the methods of Parsons et al [32] and Hedge et al [41]. Both studies used the same data sets, which had data from 47 (study 1) and 56 (study 2) participants for the Stroop task. In these studies, this sample size was sufficient to observe effects with medium effect sizes. Thus, based on these prior studies, we aimed to obtain approximately 50 participants for each condition [48].

We only analyzed data from participants who had completed both sessions. We also set quality thresholds and removed participants who did not meet them at either time point. Finally, we also removed outlying data points, such as individual trials that were much slower than the average for each participant, to reduce noise in the data, as the study was web-based, and we could not otherwise account for participant distraction from the tasks.

Statistical Analysis

Reaction Time and Error Rate Data

We conducted 2-way ANOVAs with task type (basic or game) and trial condition (congruent, neutral, or incongruent) for reaction time and error rate data. We used 1-way ANOVAs to compare the effect of task type on the skewness and kurtosis of the distribution of reaction time data for each participant. In addition, we conducted 3-way repeated measures ANOVAs (task type × trial type × time) for reaction time cost and error rate cost data. We also created groups representing low and high attentional control based on the median of 51.0 of our participants and then conducted 3-way repeated measures ANOVAs (task type × attention × time) for reaction time cost and error rate cost data.

Internal Consistency and Test-Retest Reliability

For measuring and reporting reliability, our analysis followed the recommendations from Parsons et al [32]. To evaluate internal consistency, we used a permutation approach, which involves repeatedly randomly splitting the data, calculating the reliability estimate, and then averaging all estimates. This approach provides a more stable estimate, independent of how trial stimuli and conditions are presented [32]. To evaluate test-retest reliability, we calculated the Pearson correlation between each time point. We also used intraclass correlation coefficients (ICCs) to indicate the degree of consistency and agreement between each time point. On the basis of Parson recommendations, we used ICCs labelled ICC(3,1) and ICC(2,1), as described by Shrout and Fleiss [49]. Finally, we plotted the test-retest reliability as the number of trials increased. To achieve this, we followed the method used by Hedge et al [41].

Results

Participants

For the first round of data collection (time 1), we received 135 responses, followed by 78 responses for time 2.

All participants met the criteria for questionnaire speed of completion (participants needed to spend an average of 1.5 seconds per item) and variance (participants needed to show some variance across items). In total, 13 participants were excluded because they too frequently provided an incorrect response on the Stroop task (total incorrect responses>1 SD above the mean number of incorrect responses) and because they responded to trials too slowly (mean reaction time>3 SD above the group mean reaction time). Before calculating the group mean reaction time, we also removed any individual trials that were slower than the average for each participant (reaction time>3 SD above the individual mean reaction time), as well as any remaining outlier trials that were slower than 2000 milliseconds. At time 1, we removed 1667 trials (out of 50,400). At time 2, we removed 1976 trials (out of 49,680). Notably, both at time 1 and time 2, significantly fewer trials needed to be removed from the game condition compared with the basic version; 38.6% of the removed trials were in the game condition at time 1, and 32.9% were in the game condition at time 2.

After exclusions, 65 participants remained (50 female, 13 male, 1 nonbinary, and 1 prefer not to disclose; mean age 23.91, SD 4.64 years), with 31 participants in the basic task condition and 34 participants in the game condition. Our sample had a high proportion of women because of the web-based platform we used [50]. The participants had a mean score of 51.8 (SD 7.54) on the Attentional Control Scale.

Intrinsic Motivation Inventory

At both time points, the basic task and game conditions showed no significant differences for any of the Intrinsic Motivation Inventory subscales (interest, competence, effort, and pressure).

Reaction Time and Error Rate Data

We averaged the reaction times and error rates across participants and then analyzed each measure by task type and trial condition at each time point. We also calculated reaction time and error rate costs (mean incongruent trials and mean congruent trials). Table 1 presents the descriptive statistics for each measure.

Histograms of reaction time for all participants are presented in Figure 2 by task type and time point. One-way ANOVAs revealed no significant effects of task type on the skewness and kurtosis of the distribution of reaction time data for each participant (Table 2).

The 2-way ANOVAs for reaction time and error rate demonstrated evidence of the Stroop effect at both time points (significant differences between incongruent trials and both congruent and neutral trials). Furthermore, congruence sequence effect analysis revealed the expected adaptive control effect but no effect of task condition, time, or an interaction between the 2 emerged. There were also significant differences between task conditions at time 1: participants in the game condition had faster reaction times and lower error rates than those in the basic task condition. There were no significant differences at time 2 (Tables 3 and 4).

Two-way repeated measures ANOVAs (task type × time) for reaction time cost and error rate cost data showed no significant interaction effects (Table 5). The 3-way repeated measures ANOVAs (task type × trial condition × time) for reaction time and error rate data showed no significant interaction effects (Table 5). On the basis of grouping our participants into low and high attentional control categories, we found a significant 3-way interaction between time, task type, and attention category for the error rate (Table 5). Participants who scored low in attentional control and were in the basic task condition had a lower error rate cost at time 1 than at time 2. In the game condition, participants who scored low on attentional control had a higher error rate cost at time 1 than at time 2. The error rate cost for participants who scored high on attentional control showed an opposite pattern. There were no significant simple 2-way interactions between task type and attention category at either time point.

Table 1. Descriptive statistics for reaction time and error rates, at times 1 and 2 for each task type.

	Time 1, mean (SD)	Time 2, mean (SD)
Basic task		
Congruent reaction time (milliseconds)	678 (103)	659 (104)
Neutral reaction time (milliseconds)	671 (94.0)	656 (94.7)
Incongruent reaction time (milliseconds)	796 (124)	758 (118)
Reaction time cost (milliseconds)	118 (50.9)	98.8 (39.8)
Congruent correct (%)	96.0 (2.86)	96.1 (2.52)
Neutral correct (%)	96.7 (2.33)	96.8 (2.43)
Incongruent correct (%)	93.1 (5.46)	93.6 (4.36)
Error rate cost (%)	2.86 (4.53)	2.55 (3.23)
Game task		
Congruent reaction time (milliseconds)	638 (94.5)	645 (95.3)
Neutral reaction time (milliseconds)	628 (84.1)	631 (79.1)
Incongruent reaction time (milliseconds)	753 (112)	730 (103)
Reaction time cost (milliseconds)	115 (48.8)	85.3 (42.3)
Congruent correct (%)	94.6 (3.70)	95.5 (2.50)
Neutral correct (%)	96.0 (2.53)	96.0 (2.79)
Incongruent correct (%)	92.1 (3.90)	93.0 (4.80)
Error rate cost (%)	2.52 (4.71)	2.53 (4.18)

Figure 2. Histograms of reaction time by time point and task type for each type of trial condition.

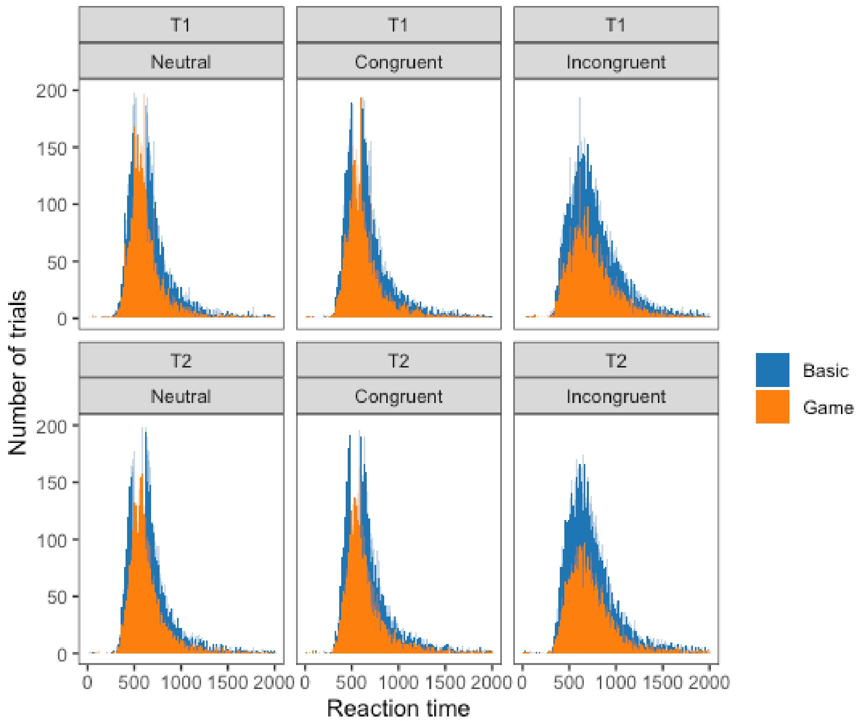


Table 2. ANOVA summary table for reaction time distribution.

	Mean squares	*F* test (*df*)	*P* values	Effect size
Time 1				
Skewness	0.316	1.863 (1)	.18	0.029
Kurtosis	0.003	0.001 (1)	.98	0.000
Time 2				
Skewness	0.317	1.852 (1)	.18	0.029
Kurtosis	1.159	0.358 (1)	.55	0.006

Table 3. ANOVA summary table for reaction time.

	Mean squares	*F* test (*df*)	*P* value	Effect size
Time 1				
Task type	85,185.015	8.107 (1)	.005	0.041
Condition	317,396.780	30.205 (3)	<.001	0.242
Task type × condition	49.167	0.005 (2)	.10	0.000
Time 2				
Task type	23,700.032	2.402 (1)	.12	0.013
Condition	201,201.515	20.394 (2)	<.001	0.178
Task type × condition	788.555	0.080 (2)	.92	0.001

Table 4. ANOVA summary table for error rate.

	Mean squares	*F* test (*df*)	*P* value	Effect size
Time 1				
Task type	0.005	4.012 (1)	.05	0.021
Condition	0.024	18.301 (2)	<.001	0.162
Task type × condition	0.000	0.148 (2)	.86	0.002
Time 2				
Task type	0.002	1.945 (1)	.17	0.010
Condition	0.018	15.402 (2)	<.001	0.140
Task type × condition	0.010	0.022 (2)	.98	0.000

Table 5. Repeated measures ANOVA summary table for reaction time and error rate.

	Mean squares	*F* test (*df*)	*P* value	Effect size
Reaction time cost				
Task type × time	880.934	1.105 (1)	.30	.017
Reaction time				
Trial type × task type × time[a]	317.106	0.616 (2)	.49	0.010
Attention × task type × time	1325.711	1.665 (1)	.20	0.012
Error rate cost				
Task type × time	<0.001	0.106 (1)	.75	0.002
Error rate				
Trial type × task type × time	0.000	0.615 (2)	.54	0.010
Attention × task type × time	39.218	5.493 (1)	.02	0.083

[a]Owing to the interaction violates the assumption of sphericity (*P*<.001), *P* values are derived using the Greenhouse-Geisser statistic.

Internal Consistency

Overview

We estimated the internal consistency of the basic task by using a permutation-based split-half approach [32] with 5000 random splits. Internal consistency ranged between 0 and 1, with higher numbers representing more consistency across an individual's complete set of trials.

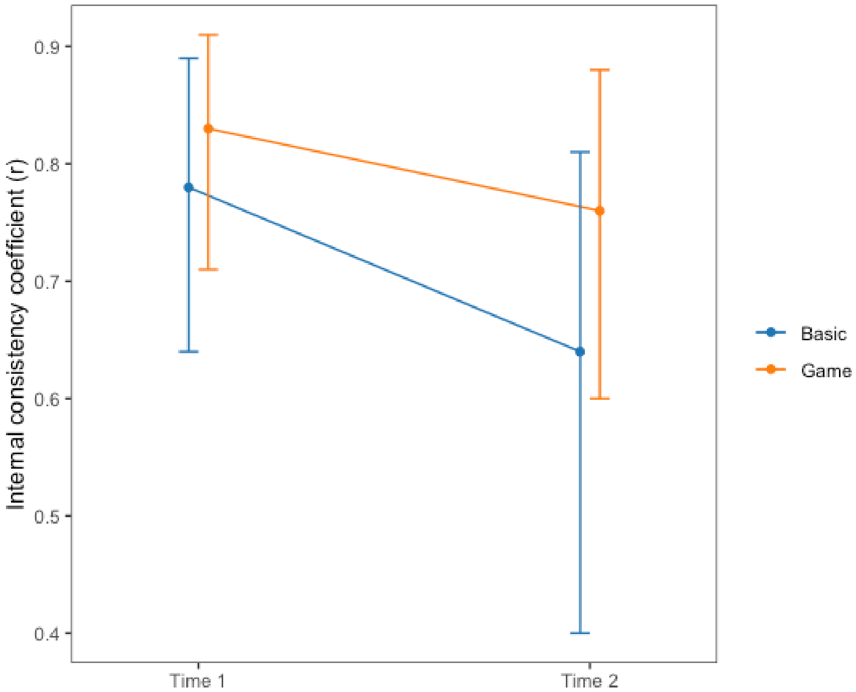
Reaction Time

When using the reaction time cost, the (Spearman-Brown corrected) split-half internal consistency for the basic task at time 1 was $r_{\text{Spearman-Brown}}$=0.78, 95% CI 0.64-0.89. At time 2, $r_{\text{Spearman-Brown}}$=0.64, 95% CI 0.40-0.81.

For the game condition at time 1, the split-half internal consistency was $r_{\text{Spearman-Brown}}$=0.83, 95% CI 0.71-0.91. At time 2, $r_{\text{Spearman-Brown}}$=0.76, 95% CI 0.60-0.88.

The internal consistency values were higher at both time 1 and time 2 for the game condition (Figure 3); however, converting the correlations to Fisher *z* scores indicated no significant differences between groups at each time point.

Figure 3. Internal consistency of reaction time cost for each time point and task type.



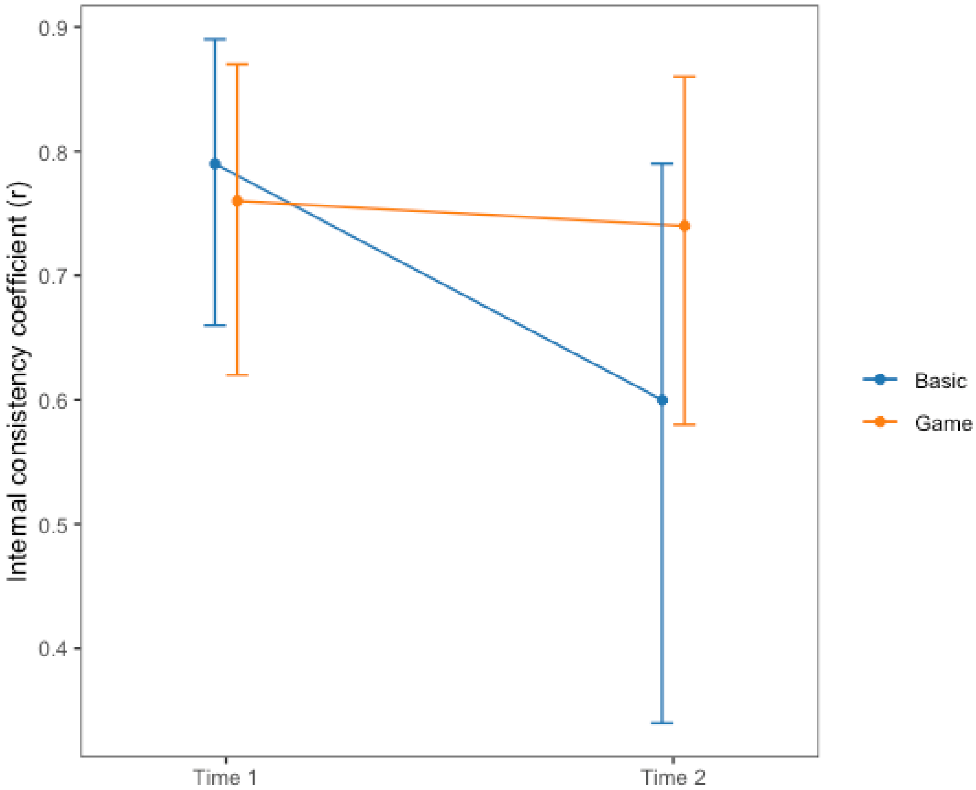
Error Rate

When using error rate cost, the (Spearman-Brown corrected) split-half internal consistency for the basic task at time 1 was $r_{\text{Spearman-Brown}}$=0.79, 95% CI 0.66-0.89. At time 2, $r_{\text{Spearman-Brown}}$=0.6, 95% CI 0.34-0.79.

For the game condition at time 1, the split-half internal consistency was $r_{\text{Spearman-Brown}}$=0.76, 95% CI 0.62-0.87. At time 2, $r_{\text{Spearman-Brown}}$=0.74, 95% CI 0.58,0.86.

The internal consistency values were higher at time 2 for the game condition at time 2 (Figure 4); however, similar to the reaction time data, converting the correlations to Fisher *z* scores indicated no significant differences between groups at each time point.

Figure 4. Internal consistency of error rate cost for each time point and task type.



Test-Retest Reliability

Reaction Time

Using reaction time cost data, for the basic task, the Pearson correlation between each time point indicated a test-retest reliability of 0.68, 95% CI 0.43-0.84. This correlation was significant (t_{29}=5.04; P<.001). For the game condition, we found a test-retest reliability of 0.58, 95% CI 0.31-0.77. This correlation was also significant (t_{32}=4.07; P<.001).

We also estimated the test-retest reliability between time 1 and time 2 with ICCs using the *psych* package in R (R Foundation for Statistical Computing) [51]. ICCs were used to measure the reliability of a measure between 2 time points. The ICC value can range from 0 to 1, with higher values indicating higher reliability. We report the results of 2-way mixed-effects models for absolute agreement, ICC(2,1), and consistency, ICC(3,1).

Using reaction time cost data, for the basic task, the estimated agreement was 0.61, 95% CI 0.36-0.78, and the estimated consistency was 0.66, 95% CI 0.46-0.80. For the game condition, the estimated agreement was 0.48, 95% CI 0.16-0.69, and the estimated consistency was 0.58, 95% CI 0.35-0.74.

Typically, cognitive tasks require many trials to reduce measurement noise. We plotted how ICC(3,1) changes as the number of trials increases, to see if a more stable estimate could be determined with fewer trials when using game elements. Figure 5 shows how the reliability of the Stroop effect (reaction time cost) changes with an increasing number of trials.

To investigate why the game condition shows lower test-retest reliability, we also plotted how the reliability of reaction time changes over time for each trial type (neutral, congruent, and incongruent trials; Figure 6). Comparing the plots suggests that the game condition reaches a higher level of consistency sooner for incongruent trials, compared with both neutral and congruent conditions. The basic task showed similar patterns of consistency across all trial types.

Figure 5. Test-retest reliability of reaction time cost as the number of trials increases for each task type.

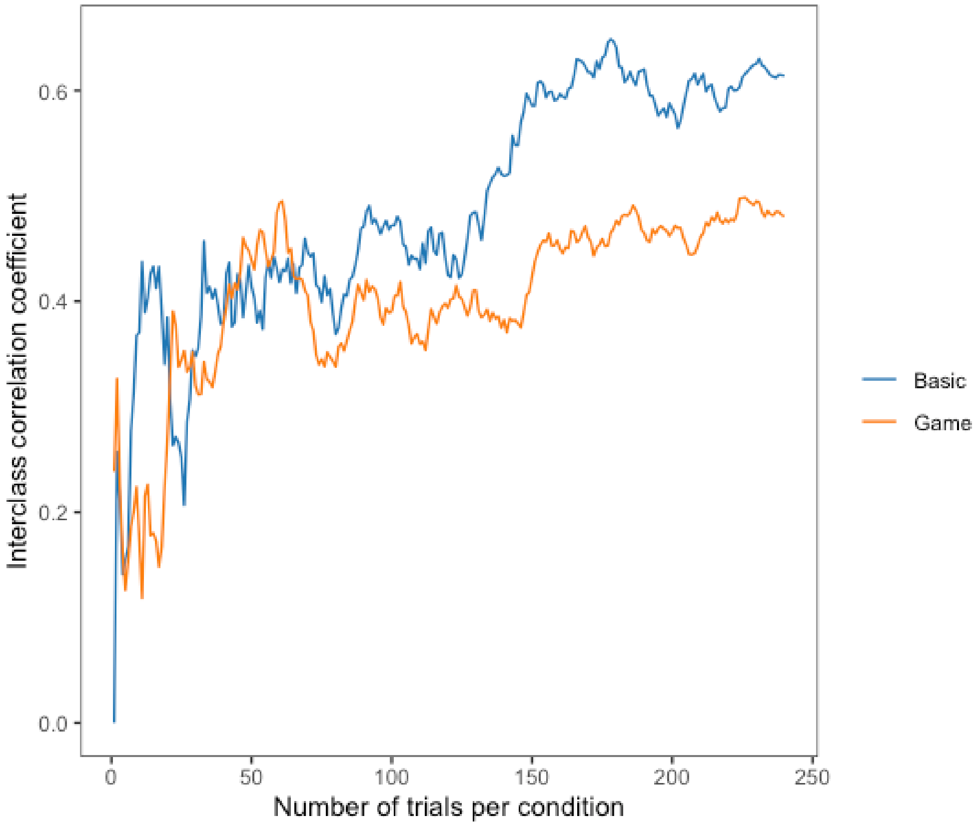
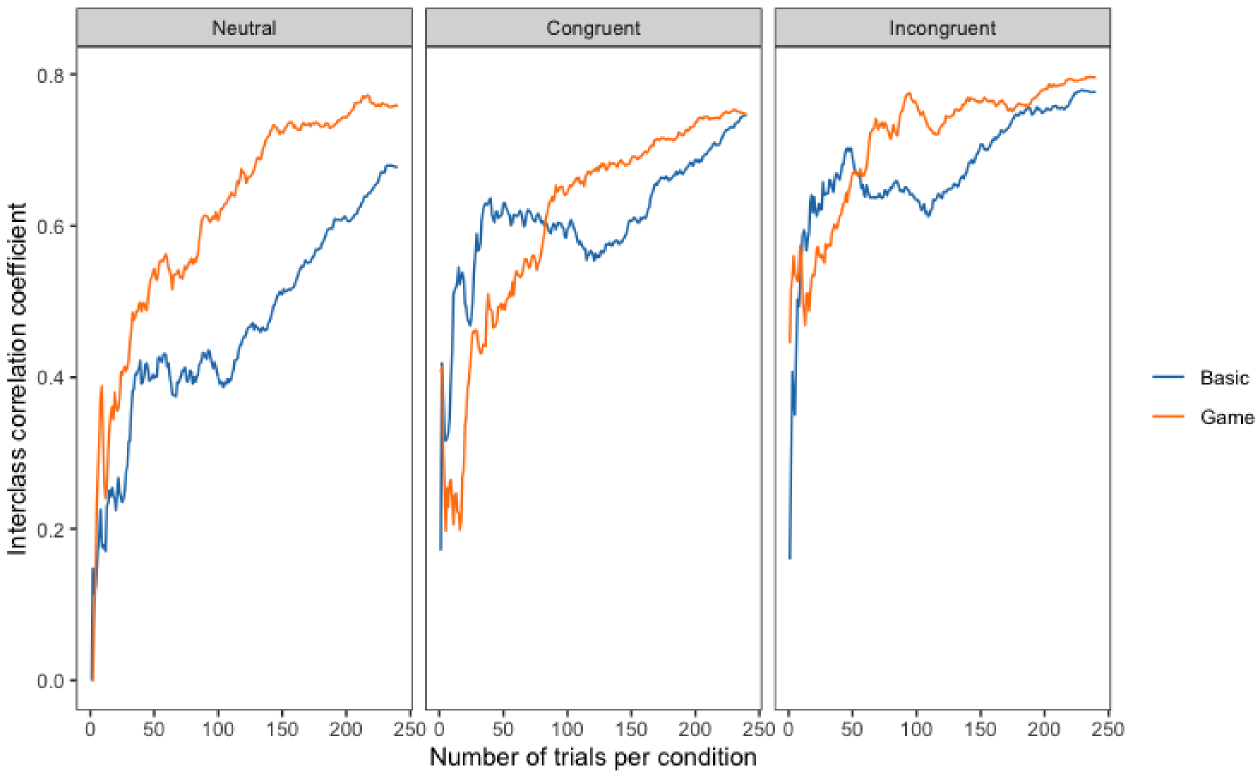


Figure 6. Test-retest reliability of reaction time as the number of trials increases for each trial type and task type.



Error Rate

Using error rate cost data, for the basic task, the Pearson correlation between each time point indicated a test-retest reliability of 0.55, 95% CI 0.24-0.76. This correlation was significant (t_{29}=3.56; *P*=.001). For the game condition, we found a test-retest reliability of 0.62, 95% CI 0.35-0.79. This correlation was also significant (t_{32}=4.45; *P*<.001).

Using error rate cost data, for the basic task, ICC(2,1) (estimated agreement) was 0.53, 95% CI 0.28-0.71, and ICC(3,1) (estimated consistency) was 0.53, 95% CI 0.28-0.71. For the game condition, ICC(2,1) was 0.62, 95% CI 0.42-0.77, and ICC(3,1) was 0.62, 95% CI 0.41-0.77.

We plotted how ICC(3,1) changes as the number of trials increases, to determine whether a more stable estimate could be determined with fewer trials when using game elements. Figure 7 shows how the reliability of the Stroop effect using the error rate cost changes with an increasing number of trials.

Similar to the reaction time, we plotted how the reliability of the number of errors changes over time for each trial type (neutral, congruent, and incongruent trials; Figure 8). The basic task showed similar patterns of consistency across all the trial types, whereas in the game condition, only the neutral and congruent conditions were similar—the reliability of the incongruent trials continued to increase over time.

Figure 7. Test-retest reliability of error rate cost as the number of trials increases, for each task type.

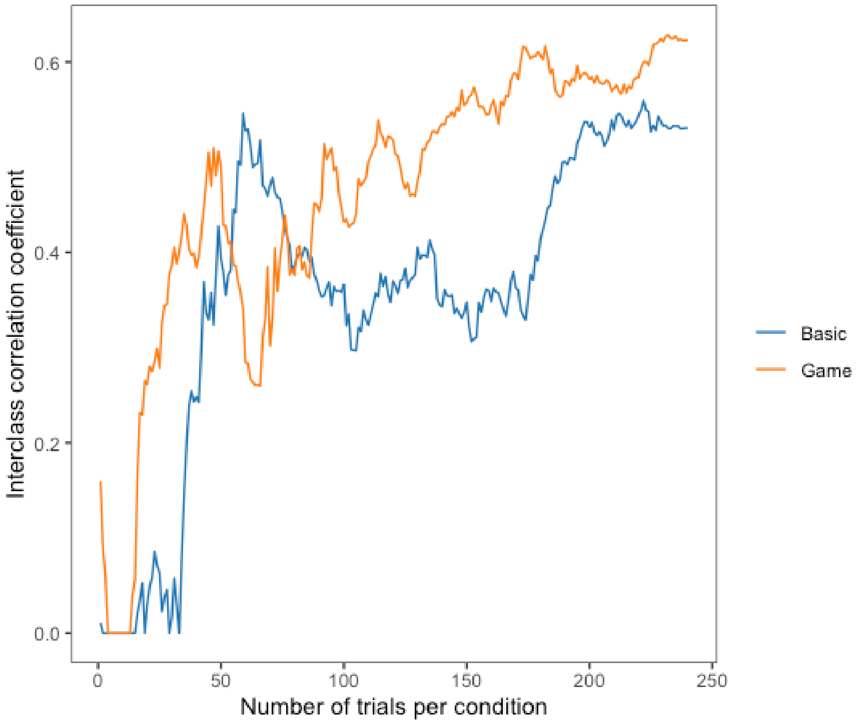
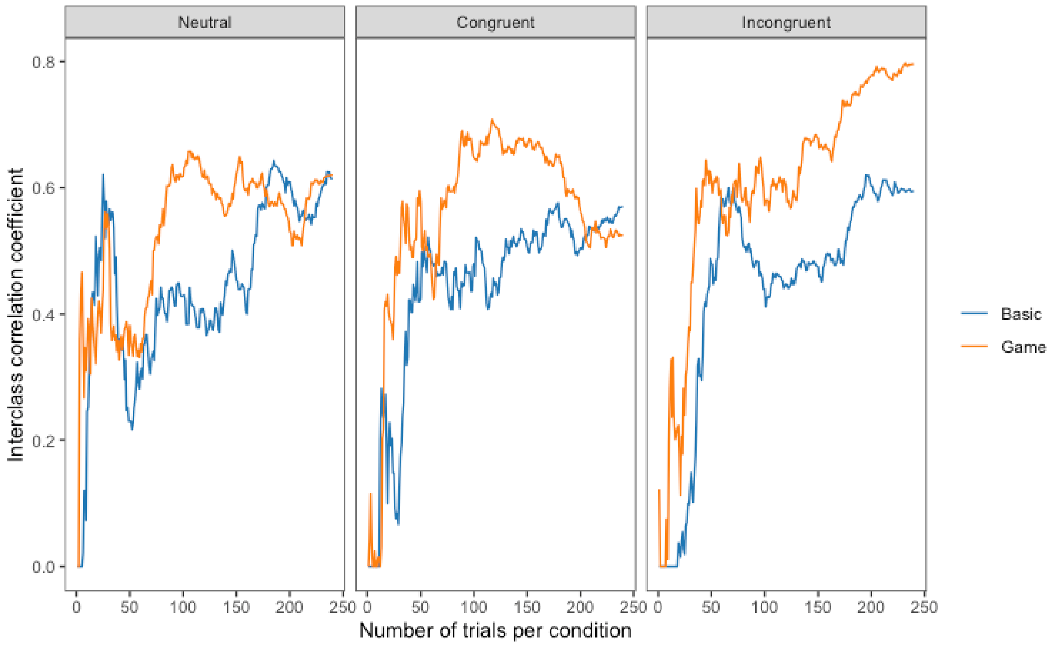


Figure 8. Test-retest reliability of error rate as the number of trials increases for each trial type and task type.



Discussion

Summary and Explanation of Findings

Performance

Both versions of the task demonstrated the Stroop effect, meaning that the effect is robust to the addition of certain game elements. Gamification can affect the validity of cognitive tasks; for example, adding graphics (especially those that change the stimuli participants respond to) can worsen performance compared with a control task [8,12,21]. In this study, in the game condition, reaction times and a progress bar were perpetually displayed on the screen. Graphics indicating gained or lost points also appeared between stimuli. These elements did not interfere with the validity of the Stroop task.

There were no significant differences in performance-based measures between the basic task and game conditions, with one exception: Participants in the game condition had significantly faster reaction times and lower error rates than those in the basic task condition but only at time 1. There may be several reasons for these results.

Points that function as extrinsic motivators have been shown to improve performance in cognitive tasks [25]; however, this effect may be short lived. Nicholson [15] noted that reward-based game elements can drive immediate spikes in engagement but only as long as continuous rewards are provided. In our game condition, participants were continually awarded points for accurate responses; however, for reaction time, they were only awarded bonus points for responses that broke their previous "fastest time" record. There is a physical limitation on how quickly participants can react to stimuli—once that threshold is met, it will be near impossible to improve further, and the motivating influence of the bonus points may be diminished.

In the game condition, participants may also learn faster and reach their "peak performance" sooner. Participants were quickly incentivized to put forth their best effort. This effect may be particularly pronounced when the cognitive demands of the task are higher. When we plotted the reliability of reaction time and error rate as the number of trials increased, the incongruent trials showed an improved pattern of consistency only in the game condition. Specifically, after approximately 50 to 100 trials, the reaction time remained consistent in the game, whereas there was a significant variation in the basic version, with a noticeable drop after 50 trials. A similar pattern was observed for the error rates. For the basic task, the plots of all 3 trial types showed similar patterns across both performance measures. This is especially noteworthy because incongruent trials are arguably the most important trials in the Stroop task, as they are the trials wherein cognitive conflict needs to be resolved. Improved performance in the incongruent trials also explains why the reliability of the Stroop effect (reaction time cost) appeared lower in the game condition—participants in that condition performed better and more consistently in the incongruent trials.

The differences between the basic task and game conditions may be emphasized by incongruent trials because they are more cognitively demanding than the congruent and neutral trials. Evidence suggests that game elements can differentially affect cognition depending on how participants experience the demands of the task. For example, gamification can normalize the performance of participants with ADHD [28].

Another indication of improved performance consistency comes in the form of a significantly smaller number of outlier trials that need to be removed from the game condition compared to the basic version. Approximately twice the number of far-out outlier trials were removed from the basic task. These trials were not considered valuable data and were essentially lost time for both the researcher and the participant. By reducing the number of trials that needed to be removed from performance, the time investment for participants was reduced. Furthermore, this means that the previous results are a conservative estimate of the game's reliability advantage because the most egregious outliers were already removed from the analysis.

Enjoyment

There were also no differences in the self-reported measures of motivation between the basic task and game conditions. These results align with those of other studies, which found that achievement-based game elements are only effective in promoting performance and not motivation [24,25].

Levy et al [23] note how carefully games must be designed to appropriately function as scientific tools and highlight the importance of using the research and data collection goals to inform the choice of game design. For this study, we specifically chose game elements that we thought would influence performance rather than enjoyment. Gamified tasks may be more successful if the game elements are just "good enough" to achieve the goals of the study without interfering with the validity of the task [23]. Because we wanted to improve participant performance irrespective of enjoyment, we did not add extraneous game elements, even if those elements would have made the game more fun.

Limitations and Future Work

One limitation of our study is the small sample size. The 2 task conditions were designed with subtle differences in the form of points and feedback. While this design was intentional, we also had a relatively small sample size, which may not have been powerful enough to reveal the small effects of our slight manipulation. We recruited 135 participants for time 1 with the intent of having at least 50 participants per condition. However, only 78 participants returned at time 2. It was difficult to incentivize participants to return to a web-based study. Future studies may find significant effects with a larger sample size.

Another limitation is that our sample was heavily skewed toward young adult female participants. We recruited participants through a web-based platform called Prolific. At the time of our study, a young woman made a video describing her hustle as a participant on the platform. Her video went viral on TikTok, resulting in an influx of new signups to Prolific, most of whom were, similar to the creator, female adults in their 20s [50]. However, given the fundamental nature of this research, this sampling bias is unlikely to have influenced the results.

The addition of points and feedback is one simple approach to gamification. Other game elements may produce different results. As discussed, we had theoretical and practical reasons for using points, but even within the category of points and achievement-based game elements, we could have made different design and mechanical choices. For example, adding a leaderboard system may have influenced participant behavior because of increased competition. Mekler et al [25] found that for an image annotation task, participants in the point condition significantly outperformed those in a control condition, where no game elements were used. However, participants in the points condition were, in turn, significantly outperformed by those in conditions where leaderboards and levels were used.

Future studies should investigate other game elements. Other cognitive tasks could also be investigated to determine how game elements affect reliability across task types that target different cognitive domains. Our same methods for investigating reliability could be applied to any gamified task.

Implications

In this study, we show that the Stroop effect is robust to the addition of simple points-based game elements. Adding points to a Stroop task does initially increase participant reaction time, but this gamification may be most effective in the short term. Our results also suggest that game elements may differently influence *parts* of a cognitive task, such as the more cognitively demanding incongruent trials.

We also provide an example of reporting psychometric data for a gamified task. Despite a long history of cognitive task gamification, the field lacks standard practices regarding how these tasks are made and measured [16]. Any advancement in how these tasks are designed and used requires a stronger base of knowledge on how individual game elements affect cognitive behavioral measures [25,32]. One of the most cited reasons for gamifying tasks is to address the limitations of standard neuropsychological testing [16]; however, these games will never be acceptable replacements for traditional tests if they are not subjected to the same rigorous standards of reliability and validity.

The results of this study suggest a potential advantage of using game-like tasks to assess cognitive functioning, especially for difficult-to-reach populations or individuals who cannot be subjected to prolonged testing. For example, gamified tasks have been shown to provide a more engaging environment that creates a more captivating setting that may aid in collecting data from populations with a lower attention span, such as children or groups of patients with concentration or attention deficits [52].

Our results suggest that the game condition may provide faster onboarding to true performance and improved consistency, as demonstrated descriptively through the lower proportion of outlier trials removed, the reaction time distributions, the split-half internal consistency values for reaction time and error rate, and reaction time cost by trial number charts. This faster onboarding is also supported by the significantly faster reaction times and lower error rates in the game condition at time 1. However, these trends do not result in significant performance differences between the basic task and game conditions in analyses of reaction time cost and also do not influence test-retest reliabilities, suggesting that the game elements we included neither significantly improved nor compromised performance in a gamified Stroop task.

Acknowledgments

The authors are grateful for the funding provided by the Natural Sciences and Engineering Research Council of Canada and the Saskatchewan-Waterloo Games User Research Training Program.

Data Availability

The data sets generated and analyzed in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

KW and RLM conceptualized the study and developed the methodology. KW and PB developed the tasks and administered the projects. KW curated the data. KW, MAF, and RLM performed analyses and interpreted the study findings. KW wrote the original draft of the manuscript, and all authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Cattell JM. The time it takes to see and name objects. Mind 1886;os-XI(41):63-65. [doi: 10.1093/mind/os-xi.41.63]
2. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935 Dec;18(6):643-662. [doi: 10.1037/h0054651]
3. Wenzel A, Rubin DC. Cognitive Methods and Their Application to Clinical Research. Washington, DC: American Psychological Association; 2005.
4. Luciana M. Practitioner review: computerized assessment of neuropsychological function in children: clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB). J Child Psychol Psychiatry 2003 Jul;44(5):649-663. [doi: 10.1111/1469-7610.00152] [Medline: 12831110]

5. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. Am J Alzheimers Dis Other Demen 2015 Feb;30(1):13-28 [FREE Full text] [doi: 10.1177/1533317514522852] [Medline: 24526761]
6. Schmader T, Johns M, Forbes C. An integrated process model of stereotype threat effects on performance. Psychol Rev 2008 Apr;115(2):336-356 [FREE Full text] [doi: 10.1037/0033-295X.115.2.336] [Medline: 18426293]
7. Boksem MA, Tops M. Mental fatigue: costs and benefits. Brain Res Rev 2008 Nov;59(1):125-139. [doi: 10.1016/j.brainresrev.2008.07.001] [Medline: 18652844]
8. Lumsden J, Skinner A, Coyle D, Lawrence N, Munafo M. Attrition from web-based cognitive testing: a repeated measures comparison of gamification techniques. J Med Internet Res 2017 Nov 22;19(11):e395 [FREE Full text] [doi: 10.2196/jmir.8473] [Medline: 29167090]
9. DeRight J, Jorgensen RS. I just want my research credit: frequency of suboptimal effort in a non-clinical healthy undergraduate sample. Clin Neuropsychol 2015;29(1):101-117. [doi: 10.1080/13854046.2014.989267] [Medline: 25494327]
10. Kirkwood MW, Kirk JW, Blaha RZ, Wilson P. Noncredible effort during pediatric neuropsychological exam: a case series and literature review. Child Neuropsychol 2010;16(6):604-618. [doi: 10.1080/09297049.2010.495059] [Medline: 20628928]
11. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
12. Lumsden J, Skinner A, Woods A, Lawrence NS, Munafò M. The effects of gamelike features and test location on cognitive test performance and participant enjoyment. PeerJ 2016 Jul 06;4:e2184 [FREE Full text] [doi: 10.7717/peerj.2184] [Medline: 27441120]
13. Chicchi Giglioli IA, de Juan Ripoll C, Parra E, Alcañiz Raya M. EXPANSE: a novel narrative serious game for the behavioral assessment of cognitive abilities. PLoS One 2018 Nov 9;13(11):e0206925 [FREE Full text] [doi: 10.1371/journal.pone.0206925] [Medline: 30412614]
14. Berg V, Rogers SL, McMahon M, Garrett M, Manley D. A novel approach to measure executive functions in students: an evaluation of two child-friendly apps. Front Psychol 2020 Jul 16;11:1702 [FREE Full text] [doi: 10.3389/fpsyg.2020.01702] [Medline: 32765379]
15. Nicholson S. A RECIPE for meaningful gamification. In: Reiners T, Wood L, editors. Gamification in Education and Business. Cham, Switzerland: Springer; 2015.
16. Wiley K, Robinson R, Mandryk RL. The making and evaluation of digital games used for the assessment of attention: systematic review. JMIR Serious Games 2021 Aug 09;9(3):e26449 [FREE Full text] [doi: 10.2196/26449] [Medline: 34383674]
17. Thirkettle M, Lewis J, Langdridge D, Pike G. A mobile app delivering a gamified battery of cognitive tests designed for repeated play (OU Brainwave): app design and cohort study. JMIR Serious Games 2018 Oct 30;6(4):e10519 [FREE Full text] [doi: 10.2196/10519] [Medline: 30377140]
18. Friehs MA, Dechant M, Schäfer S, Mandryk RL. More than skin deep: about the influence of self-relevant avatars on inhibitory control. Cogn Res Princ Implic 2022 Apr 08;7(1):31 [FREE Full text] [doi: 10.1186/s41235-022-00384-8] [Medline: 35394227]
19. Friehs MA, Dechant M, Vedress S, Frings C, Mandryk RL. Effective gamification of the stop-signal task: two controlled laboratory experiments. JMIR Serious Games 2020 Sep 08;8(3):e17810 [FREE Full text] [doi: 10.2196/17810] [Medline: 32897233]
20. Wiley K, Vedress S, Mandryk RL. How points and theme affect performance and experience in a gamified cognitive task. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3313831.3376697]
21. Birk MV, Mandryk RL, Bowey J, Buttlar B. The effects of adding premise and backstory to psychological tasks. In: Proceedings of the Conference on Human Factors in Computing Systems. 2015 Presented at: CHI'15; April 18-23, 2015; Seoul, Korea.
22. Bruckman A. Can educational be fun? In: Proceedings of the Game Developer's Conference. 1999 Presented at: GDC 1999; March 17, 1999; San Jose, CA.
23. Levy L, Lambeth A, Solomon R, Gandy M. Method in the madness: the design of games as valid and reliable scientific tools. In: Proceedings of the 13th International Conference on the Foundations of Digital Games. 2018 Presented at: FDG '18; August 7-10, 2018; Malmö, Sweden.
24. Groening C, Binnewies C. “Achievement unlocked!” - the impact of digital achievements as a gamification element on motivation and performance. Comput Hum Behav 2019 Aug;97:151-166. [doi: 10.1016/j.chb.2019.02.026]
25. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput Hum Behav 2017 Jun;71:525-534. [doi: 10.1016/j.chb.2015.08.048]
26. Jung JH, Schneider C, Valacich J. Enhancing the motivational affordance of information systems: the effects of real-time performance feedback and goal setting in group collaboration environments. Manag Sci 2010 Apr;56(4):724-742. [doi: 10.1287/mnsc.1090.1129]

27. Vanden Abeele V, Wouters J, Ghesquière P, Goeleven A, Geurts L. Game-based assessment of psycho-acoustic thresholds: not all games are equal!. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. 2015 Presented at: CHI PLAY '15; October 5-7, 2015; London, UK. [doi: 10.1145/2793107.2793132]
28. Delisle J, Braun CM. A context for normalizing impulsiveness at work for adults with attention deficit/hyperactivity disorder (combined type). Arch Clin Neuropsychol 2011 Nov;26(7):602-613. [doi: 10.1093/arclin/acr043]
29. The standards for educational and psychological testing. American Psychological Association. URL: https://www.apa.org/science/programs/testing/standards [accessed 2024-03-12]
30. Hedge C, Bompas A, Sumner P. Task reliability considerations in computational psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 2020 Sep;5(9):837-839. [doi: 10.1016/j.bpsc.2020.05.004] [Medline: 32605726]
31. Miranda AT, Palmer EM. Intrinsic motivation and attentional capture from gamelike features in a visual search task. Behav Res Methods 2014 Mar;46(1):159-172. [doi: 10.3758/s13428-013-0357-7] [Medline: 23835649]
32. Parsons S, Kruijt AW, Fox E. Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. Adv Methods Pract Psychol Sci 2019;2(4):378-395. [doi: 10.1177/2515245919879695]
33. Price RB, Kuckertz JM, Siegle GJ, Ladouceur CD, Silk JS, Ryan ND, et al. Empirical recommendations for improving the stability of the dot-probe task in clinical research. Psychol Assess 2015 Jun;27(2):365-376 [FREE Full text] [doi: 10.1037/pas0000036] [Medline: 25419646]
34. MacLeod CM. The Stroop task in cognitive research. In: Wenzel A, Rubin DC, editors. Cognitive Methods and Their Application to Clinical Research. Washington, DC: American Psychological Association; 2005.
35. Klein GS. Semantic power measured through the interference of words with color-naming. Am J Psychol 1964 Dec;77(4):576-588. [doi: 10.2307/1420768]
36. Tecce JJ, Dimartino M. Effects of heightened drive (shock) on performance in a tachistoscopic color-word interference task. Psychol Rep 1965 Feb;16:93-94. [doi: 10.2466/pr0.1965.16.1.93]
37. Dalrymple-Alford EC, Budayer B. Examination of some aspects of the stroop color-word test. Percept Mot Skills 1966 Dec;23(3):1211-1214. [doi: 10.2466/pms.1966.23.3f.1211] [Medline: 5972923]
38. MacLeod C. The Stroop task in clinical research. In: Wenzel A, Rubin DC, editors. Cognitive Methods and Their Application to Clinical Research. Washington, DC: American Psychological Association; 2005.
39. Groening C, Binnewies C. The more, the merrier? - how adding and removing game design elements impact motivation and performance in a gamification environment. Int J Hum Comput Interact 2021;37(12):1130-1150. [doi: 10.1080/10447318.2020.1870828]
40. Gómez-Tello MF, Rosetti MF, Galicia-Alvarado M, Maya C, Apiquian R. Neuropsychological screening with TOWI: performance in 6- to 12-year-old children. Appl Neuropsychol Child 2022;11(2):115-124. [doi: 10.1080/21622965.2020.1764357] [Medline: 32429822]
41. Hedge C, Powell G, Sumner P. The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. Behav Res Methods 2018 Jun;50(3):1166-1186 [FREE Full text] [doi: 10.3758/s13428-017-0935-1] [Medline: 28726177]
42. Tondello GF, Kappen DL, Ganaba M, Nacke LE. Gameful design heuristics: a gamification inspection tool. In: Proceedings of the 21st International Conference on Human-computer Interaction. 2019 Presented at: HCII 2019; July 26-31, 2019; Orlando, FL. [doi: 10.1007/978-3-030-22646-6_16]
43. Paolacci G, Chandler J. Inside the turk: understanding mechanical turk as a participant pool. Curr Dir Psychol Sci 2014;23(3):184-188. [doi: 10.1177/0963721414531598]
44. Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods 2012 Sep;17(3):437-455. [doi: 10.1037/a0028085] [Medline: 22506584]
45. Buchanan EM, Scofield JE. Methods to detect low quality data and its implication for psychological research. Behav Res Methods 2018 Dec;50(6):2586-2596. [doi: 10.3758/s13428-018-1035-6] [Medline: 29542063]
46. Ryan RM. Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. J Pers Soc Psychol 1982;43(3):450-461. [doi: 10.1037//0022-3514.43.3.450]
47. Derryberry D, Reed MA. Anxiety-related attentional biases and their regulation by attentional control. J Abnorm Psychol 2002 May;111(2):225-236. [doi: 10.1037//0021-843x.111.2.225] [Medline: 12003445]
48. Lakens D. Sample size justification. Collabra Psychol 2022;8(1):33267 [FREE Full text] [doi: 10.1525/collabra.33267]
49. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
50. Charalambides N. We recently went viral on TikTok - here's what we learned. Prolific. 2021 Aug 24. URL: https://www.prolific.com/resources/we-recently-went-viral-on-tiktok-heres-what-we-learned [accessed 2024-03-12]
51. Revelle W. Psych: procedures for psychological, psychometric, and personality research. The Comprehensive R Archive Network. 2024 Jan 18. URL: https://cran.r-project.org/web/packages/psych/index.html [accessed 2024-03-12]
52. Gallagher R, Kessler K, Bramham J, Dechant M, Friehs MA. A proof-of-concept study exploring the effects of impulsivity on a gamified version of the stop-signal task in children. Front Psychol 2023 Feb 9;14:1068229 [FREE Full text] [doi: 10.3389/fpsyg.2023.1068229] [Medline: 36844283]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
ICC: intraclass correlation coefficient

Edited by T Leung; submitted 29.06.23; peer-reviewed by X Cheng, Z Aghaei; comments to author 04.09.23; revised version received 29.11.23; accepted 31.01.24; published 10.04.24.

Please cite as:
Wiley K, Berger P, Friehs MA, Mandryk RL
Measuring the Reliability of a Gamified Stroop Task: Quantitative Experiment
JMIR Serious Games 2024;12:e50315
URL: https://games.jmir.org/2024/1/e50315
doi:10.2196/50315
PMID:38598265

Smartphone-Based Virtual and Augmented Reality Implicit Association Training (VARIAT) for Reducing Implicit Biases Toward Patients Among Health Care Providers: App Development and Pilot Testing

Jiabin Shen[1], PhD; Alex J Clinton[1], MA; Jeffrey Penka[2], MA, MEd; Megan E Gregory[3], PhD; Lindsey Sova[4], MPH; Sheryl Pfeil[5], MD; Jeremy Patterson[6], BA; Tensing Maa[7], MD

1
2
3
4
5
6
7

Corresponding Author:
Jiabin Shen, PhD

Abstract

Background: Implicit bias is as prevalent among health care professionals as among the wider population and is significantly associated with lower health care quality.

Objective: The study goal was to develop and evaluate the preliminary efficacy of an innovative mobile app, VARIAT (Virtual and Augmented Reality Implicit Association Training), to reduce implicit biases among Medicaid providers.

Methods: An interdisciplinary team developed 2 interactive case-based training modules for Medicaid providers focused on implicit bias related to race and socioeconomic status (SES) and sexual orientation and gender identity (SOGI), respectively. The simulations combine experiential learning, facilitated debriefing, and game-based educational strategies. Medicaid providers (n=18) participated in this pilot study. Outcomes were measured on 3 domains: training reactions, affective knowledge, and skill-based knowledge related to implicit biases in race/SES or SOGI.

Results: Participants reported high relevance of training to their job for both the race/SES module (mean score 4.75, SD 0.45) and SOGI module (mean score 4.67, SD 0.50). Significant improvement in skill-based knowledge for minimizing health disparities for lesbian, gay, bisexual, transgender, and queer patients was found after training (Cohen *d*=0.72; 95% CI −1.38 to −0.04).

Conclusions: This study developed an innovative smartphone-based implicit bias training program for Medicaid providers and conducted a pilot evaluation on the user experience and preliminary efficacy. Preliminary evidence showed positive satisfaction and preliminary efficacy of the intervention.

(JMIR Serious Games 2024;12:e51310) doi:10.2196/51310

KEYWORDS

implicit bias; health care; Medicaid; virtual reality; augmented reality; smartphone; mHealth; mobile app; innovative; implicit bias training program; sexual orientation; sexual orientations; gender identity; gender identities; gender preferences; gender preference; efficacy; health care providers; health care provider; socioeconomic; mobile application; training; XR; extended reality

Introduction

Implicit Bias in Health Care Settings

Defined as unconscious associations or negative evaluations of a person or group of people on the basis of nonrelevant characteristics [1], implicit biases have been found to be prevalent among the general population against "marginalized" groups such as those from minority racial, ethnic, or socioeconomic backgrounds [2]. Implicit biases, which present in health care settings as irrational and unconscious perceptions, stereotypes, or prejudices among health care providers when interacting with patients, are especially concerning [3]. Past research has found that implicit bias in health care settings is associated with a decrease in overall quality of care, with impacts including increased risk of misdiagnosis [4-7], inaccurate patient pain perception [8,9], differential treatment recommendations for patients who belong to sexual orientation or gender identity minority groups [10-12], and negative perceptions of patients from racial minority backgrounds [13-17]. Implicit biases may also exist during interactions between health care professionals, such as selection bias when choosing candidates for future health care practitioner residency [18], which may have wider implications for the quality and safety of patient care. Furthermore, such implicit biases have been found within adult and pediatric health care settings [19] across medical conditions including ADHD, asthma, cardiology, and child abuse, which could affect quality of care for these vulnerable populations [20-22].

Existing Efforts to Reduce Implicit Bias in Health Care Settings

In response, increasing efforts have been devoted to addressing the significant threat posed by implicit bias toward health care services and patient outcomes. The first type of interventional efforts focus on "environmental engineering," with the goal to minimize mechanisms in health care settings that may give rise to biased interactions between health care professionals and patients. One example of this type of intervention is the development and implementation of automatic patient care prompts through electronic portals, where computer algorithms are standardized for all patients regardless of sociodemographic backgrounds, attempting to reduce opportunities for human interference (beyond the algorithm development phase) [23]. A second type of intervention uses cognitive rehearsal to walk practitioners through potentially harmful scenarios to practice their ideal response; this has shown promise at changing health care practitioner behavior to reduce bullying and workplace turnover [24,25]. While not widely used in combatting health care bias explicitly, the methodology shows a clear avenue for its application to bias training.

A third type of intervention, which will also be the focus of this study, attempts to develop educational programs with the goal of improving knowledge and awareness of implicit bias among medical students or health care professionals, which can range from traditional educational seminars to experience-oriented storytelling, to highlight the importance of patient perspectives in daily practice [26-28]. Such efforts have so far yielded positive results where health care professionals were found to become more aware of their own biases and have resulted in improved communication between health care professionals and marginalized patient groups [29,30].

Application of Augmented Reality–Based Medical Training

Despite the promising results from educational programs in the existing literature, one limitation in existing approaches for implicit bias training is the lack of immersive learning experiences that may provide optimal learning outcomes and behavior changes. As a cutting-edge technology that prioritizes experiential learning, virtual reality (VR) and augmented reality (AR) could provide an ideal solution with immersive learning experiences for implicit bias training. For example, one recent study examined biases during interactions between virtual health care providers and virtual patients for medical triage training. Regardless of the skin tone of the avatar (ie, the health care provider), it took participants more time to initiate assistance and they were more likely to make errors when triaging dark-skinned virtual patients compared to light-skinned virtual patients [31].

AR, as a more recent member of the x-reality technologies, is posed to offer an even better learning experience that combines the immersion provided by VR and tailored customization that adapts to users' dynamic environments. Adoption of AR in medical education has been found in a wide range of medical branches from surgery (eg, laparoscopic procedure training) to anatomy [32]. Furthermore, because AR-based training is readily available on consumer-grade mobile devices such as smartphones and tablets, its mobility provides medical professionals with remote accessibility to training content regardless of their physical location (this advantage has been further acknowledged during the COVID-19 pandemic) [33]. However, despite the increasing adoption of AR in medical training, a recent systematic review has found little evidence on the availability of AR-based implicit bias training among health care professionals in the literature [33].

This Study

To address this important gap, this study aimed to develop a mobile training program, VARIAT (Virtual and Augmented Reality Implicit Association Training), specifically for improving the awareness of implicit biases among health care providers when interacting with patients in daily practice. The design considerations for developing this novel AR-based implicit bias training program are described, followed by a preliminary examination of initial user feasibility and learning outcomes, including user reactions; relevance to practice; and changes in knowledge, attitudes, and behavioral skills related to implicit bias before and after receiving the training program.

Methods

Designing the VARIAT Program

Overview of Technical Design Considerations

The VARIAT program focused on delivering an immersive, interactive learning experience to the broadest possible audience in self-contained segments, allowing users to complete the training over time and a variety of sessions while retaining their progress across sessions. When building the 3D worlds for delivery on the broadest number and sizes of mobile devices, simplified, realistic, and familiar spaces were built, including offices, lobbies, and examination rooms where the learner could experience the simulations. Characters in the world were designed with exaggerated cartoon features to provide visual distinction with skin tone, hair, size, outfits, and accessories, while minimizing unnecessary details and maximizing ease of recognition for interaction on mobile-sized screens. The approach to world design addressed design and performance considerations, allowing production of additional characters and scenarios without significant technical overhead in either the creation process or the learner's experience on their device. The dialogue and training content was presented via text.

Hardware Requirements and Considerations

One key goal of the VARIAT program was the need to maximize audience reach and minimize specialty equipment for learners to access the content. At the time of its development, most iPhones and Android devices had the cameras, accelerometers, and gyroscopes needed to provide users the ability to see into and navigate virtual worlds by simply holding up and moving their devices. When the VARIAT program was introduced, learners needed at least an iPhone 8 with iOS 13 or an Android device running Android 9 or higher.

Software Requirements and Considerations

The maturity of the mobile app environment offers many development tools and approaches for developing mobile apps. For the VARIAT program, the developers used the Unity game engine (Unity Technologies) for game content with ARKit (Apple Inc) and ARCore (Google LLC) for the augmented reality component and deployed both iOS and Android apps that were readily available in their respective app stores. Blender (Blender Foundation) was used for 3D modeling and animation, and Photoshop (Adobe) was used to create 2D assets.

The learner downloads the app from the Apple App Store or Google Play on their device, and their progress is maintained on the device with evaluation, progress, and study data synchronized as the learner completes various modules. When synchronized, the data are stored and managed using Google Big Query Workspace, which produces data feeds for training evaluators and researchers.

Overview of Content Design Considerations

Training Framework of the VARIAT App

The VARIAT game was designed based on the integration of evidence-based cognitive psychology with the latest simulation technologies, including VR (eg, a simulated experience of interacting with a virtual clinician-patient scenario that is vulnerable to implicit biases in a virtual environment using 6-degrees-of-freedom motion- and gesture-based interactions) and AR (eg, converting a user's physical environment into a clinic's waiting room for interactive experiences). The goal of the app is to improve awareness of implicit biases among Medicaid providers, to educate them on how these biases can lead to inequitable care, and to offer strategies and resources that may minimize health disparities. This mobile app can be installed on any Apple iOS and Android device and is designed to be completed in one sitting or in short segments.

The game consists of 2 distinct but interconnected modules, targeting implicit biases within medical settings toward patients from minority racial backgrounds, with low socioeconomic status (SES), or from a minority sexual orientation or gender identity (SOGI) group. Learners enter an AR-based interactive role-playing game, in which they encounter a series of 6 scenarios. Each scenario takes approximately 5 minutes to complete and is related to the specific implicit bias being addressed in that particular module (Race/SES or SOGI). Each scenario within the module is designed to address specific issues related to disparities in medical settings that relate to the overall theme of the module. The primary outcome measure of this training program is to evaluate providers' attitudes and beliefs on key concepts related to implicit biases and health disparities in a medical setting before and immediately following the training.

Race/SES Module

The first module within the VARIAT mobile game is the race/SES module, which consists of 3 scenarios dealing with issues of racial bias, transportation and food instability, and implicit bias. The first scenario in this module addresses issues of racial bias within a health care setting and prompts the user to consider how issues of racial and ethnic identity could impact treatment recommendations and the resulting care for patients of minority groups. The next scenario is designed to promote self-reflection on how socioeconomic factors like unreliable transportation or housing could affect a patient's ability to show up for health care visits or comply with care recommendations by medical providers. The last scenario is designed to help users understand how implicit bias from medical providers could impact patient perception and negatively impact patient care.

Each scenario contains prompts where the user is asked to make a choice about the "case" presented within the VARIAT AR game. The user is then given information about the scenario and resources for how to better understand the specific issues for each scenario with the goal to educate them on how to improve practitioner behavior as it relates to the theme of the module. A summary of the scenarios and objectives for this module can be found in Multimedia Appendix 1.

SOGI Module

The second module in the VARIAT mobile game is the SOGI module, which consists of 3 scenarios dealing with issues of SOGI implicit bias, inclusivity for patient care settings, and lesbian, gay, bisexual, transgender, queer (LGBTQ+) patient considerations. The first scenario helps portray the way that microaggressions and implicit biases in patient-provider communication can promote negative disparities in treatment for SOGI minority patients. The second scenario asks users to design their own patient waiting room and helps educate and guide users on what considerations should be made to ensure a medical setting is a safe and welcoming environment for SOGI minority patients. The last scenario in the module helps users recognize the harmful effects of biased behavior toward LGBTQ+ patients and offers space for self-reflection on how to reduce enacted bias for this patient group.

The scenarios in this module also contain prompts for users to answer to better assess their understanding of key concepts or takeaways from each scenario. The SOGI module places an increased emphasis on self-reflection as the scenarios are designed to help users draw parallels to their own experiences in medical practice through reflective exercises following the conclusion of each scenario within the module. A summary of the scenarios and objectives for this module can be seen in Multimedia Appendix 2.

User Workflow

On start-up, users are given some brief instructions on how to prepare themselves for immersion in the VARIAT AR game.

Users are then instructed to select any available module to start engaging with the content within. Once a module is selected, participants are placed in a virtual hospital setting and can check on the various patients within. When selected in the AR game, these patients display information on their illnesses and present the user with additional narratives about the patients from the "staff" in the AR game. Users are then given different decision options on what to do for each patient's individual case. After helping these patients, the user is provided with information and resources that relate to the content of the module. Unbiased choices "score" higher than choices that are considered to have been influenced by implicit biases toward marginalized patients. After completing the tasks in their module, the users are given a summary of their scores for that module with feedback on how to improve, and additional information to support that improvement relative to the context of their scenario. After completing a module, users are sent back to the home screen, where they can replay the same module or select a new module to explore. A depiction of the app layout, user experience, and scenario prompts is presented in Multimedia Appendix 3.

User Experience and Preliminary Efficacy of the VARIAT Program

Participants and Procedure

Eighteen clinicians (n=12 female) who were predominantly White (non-Hispanic) participated in the VARIAT training. Physicians comprised 8 of the 18 (44%) participants, and 12 of the 18 participants had more than 5 years' experience in health care. The most common workplace setting was hospitals, with private practices, health care system–affiliated clinics, and other workplace settings reported as well. Most participants estimated that Medicaid patients comprised more than 30% of their total caseload, with reported ages of patients seen varying between children, adults, and older adults. Demographic information is reported in Table 1 for the total number of participants (N=18), participants who participated only in the race/SES module (n=7), participants who participated in only the SOGI module (n=5), and participants who completed both the race/SES and SOGI modules (n=6). Participants were recruited through professional networks and were eligible for the study if they were Medicaid providers.

Table . Demographic information for the participants.

Variables		Overall (n=18), n (%)	Race/SES[a] only (n=7), n (%)	SOGI[b] only (n=5), n (%)	Race/SES and SOGI (n=6), n (%)
Race/ethnicity					
	White (non-Hispanic)	16 (88)	6 (85)	5 (100)	5 (83)
	Non–White[c]	2 (11)	1 (14)	0 (0)	1 (16)
Gender					
	Male	6 (33)	2 (28)	2 (40)	2 (33)
	Female	12 (66)	5 (71)	3 (60)	4 (66)
	Prefer not to say	0 (0)	0 (0)	0 (0)	0 (0)
Profession					
	Medical resident	3 (16)	2 (28)	0 (0)	1 (16)
	Nurse	0 (0)	0 (0)	0 (0)	0 (0)
	Fully credentialed physician	8 (44)	3 (42)	2 (40)	3 (50)
	Social worker	5 (27)	1 (14)	2 (40)	2 (33)
	Other	2 (11)	1 (14)	1 (20)	0 (0)
Work setting					
	Health care system–affiliated clinic	2 (11)	1 (14)	1 (20)	0 (0)
	Hospital	9 (50)	3 (42)	3 (60)	3 (50)
	Private practice	2 (11)	1 (14)	0 (0)	1 (16)
	Other	4 (22)	2 (28)	1 (20)	1 (16)
	Missing	1 (5)	0 (0)	0 (0)	1 (16)
Experience in work setting					
	Less than 1 year	0 (0)	0 (0)	0 (0)	0 (0)
	1-5 years	6 (33)	3 (42)	1 (20)	2 (33)
	6-10 years	3 (16)	2 (28)	1 (20)	0 (0)
	11-15 years	2 (11)	1 (14)	0 (0)	1 (16)
	16-20 years	0 (0)	0 (0)	0 (0)	0 (0)
	21-25 years	2 (11)	1 (14)	1 (20)	0 (0)
	26-30 years	2 (11)	0 (0)	1 (20)	1 (16)
	31 years or more	2 (11)	0 (0)	1 (20)	1 (16)
	Missing	1 (5)	0 (0)	0 (0)	1 (16)
Percentage of Medicaid patients seen					
	Less than or equal to 30%	4 (22)	2 (28)	1 (20)	1 (16)
	Greater than 30%	12 (66)	5 (71)	4 (80)	3 (50)
	I do not see Medicaid patients	1 (5)	0 (0)	0 (0)	1 (16)
	Missing	1 (5)	0 (0)	0 (0)	1 (16)
Age of patients[d]					
	Children	10 (55)	5 (71)	3 (60)	2 (33)
	Adults	11 (61)	0 (0)	2 (40)	2 (33)
	Older adults	8 (44)	5 (71)	2 (40)	1 (16)

Variables		Overall (n=18), n (%)	Race/SES[a] only (n=7), n (%)	SOGI[b] only (n=5), n (%)	Race/SES and SOGI (n=6), n (%)
	I do not see Medicaid patients	1 (5)	0 (0)	0 (0)	1 (16)

[a]SES: socioeconomic status.

[b]SOGI: sexual orientation and gender identity.

[c]Combined category.

[d]Multiple answers selected.

Questionnaires were administered to participants remotely through the VARIAT app, and data collection took place from March to June 2020.

Measures

User Experience Measures

Users' reactions to both the race/SES and SOGI modules of the VARIAT program were assessed by asking participants about their perception of the modules after the test. After experiencing the AR simulation, users were asked questions designed to test their engagement with the AR experience, such as if they felt a sense of "being there" in the AR experience or how real they found the AR experience to be. These answers were scored on a scale of 1 to 5, with higher scores indicating stronger agreement. Participants were also asked questions about how they might apply the AR experience to their job with questions such as "How do you think this training will help you on the job (Mark all that apply)?" with different response items to assess perceived benefits from the training. These items were scored using dichotomous coding for each option (0 for not applicable and 1 for applicable).

Preliminary Efficacy Measures

Training outcomes were reported through changes in affective knowledge and changes in skill-based knowledge measured by comparing pre-post test responses. Affective knowledge (items assessing how participants expect their perceptions to impact their patients) was measured by agreement with items that were adapted from the California Brief Multicultural Competence Scale [31]. Example items include the following: "I am aware of how my own values might affect my patients" or "I am aware of institutional barriers that affect patients." Skill-based knowledge was assessed differently for the race/SES and SOGI modules, with questions referring to each respective population focused on in the module. Race/SES skill-based knowledge was measured by rating participant agreement with the following internally developed statements: "I am confident that I can recognize the role that implicit bias plays in leading to inequitable care for patients of low socioeconomic status," and "I am confident that I can apply strategies and use resources to minimize health care disparities for patients with low socioeconomic status." SOGI skill-based knowledge was measured similarly, with "race/SES population" being replaced with "LGBTQ+ population" in the skills-based questions. Training outcomes were reported for each module separately, with the race/SES module (n=13) and SOGI module (n=11) consisting of all participants that completed each module. All measures were scored on a scale of 1 to 5, with higher scores representing stronger agreement.

Data Analysis Plan

All analyses were conducted using SPSS Statistics (version 27.0; IBM Corp). Demographic characteristics were described using frequencies and percentages for the categorical variables. Demographic characteristics were reported across 4 participant groupings: participants who took only the Race/SES module, participants who took only the SOGI module, participants who took both the Race/SES and SOGI modules, and an overall group of all unique participants.

After testing for normality using the Shapiro-Wilk test, the training reactions and pre-post skills and attitude outcome data were found to not be normally distributed ($P<.001$). As a result, we used nonparametric tests for analyzing these 2 outcome domains. For the usability data, we used the Wilcoxon signed-rank test to measure the continuous reaction items and report the mean and SD for participants who used both the race/SES module and the SOGI module. The categorical training reaction items were reported using frequencies and percentages. For analyzing the skills and attitudes outcome data, the Wilcoxon matched-pairs signed-rank test was used to report the mean, SD, effect size (Cohen *d*), and 95% CI for pre-post changes in scores. The scores for each module were analyzed separately for all participants who took each respective module. Given that some participants completed both modules (n=6), there is a small amount of overlap in participant representation across all reported outcome data. All data and study materials will be made available on request.

Ethical Considerations

The Ohio State University (OSU) Institutional Review Board has determined this study was exempt from review according to the Policy on Human Subjects Research of the OSU Human Research Protection Program.

Results

User Experience (Training Reactions)

For perception of the AR experience, participants who received training in the race/SES and SOGI modules reported similar ratings for the overall AR experience. Participants reported positive feelings of "being there" (race/SES module: mean score 4.62, SD 1.56; SOGI module: mean score 3.91, SD 1.97) and high relevance of the AR training to their respective jobs (race/SES module: mean score 4.75, SD 0.45; SOGI module: mean score 4.67, SD 0.50) across both modules. Participants across both modules perceived the AR experience as being "a little" realistic, with the SOGI participants reporting less realism on average (mean score 2.91, SD 1.64) compared to the race/SES participants (mean score 3.77, SD 1.83). For the reported intention to apply the AR experience to their jobs, only the participants who received training in the race/SES module responded to this item. On average, these participants reported that they were less likely to apply the AR experience to their jobs (mean score 2.31, SD 1.11).

Assessing the perceived benefits of the AR experience to the participants' jobs revealed that the race/SES and SOGI modules had some key differences in support across beliefs. Participants from both the race/SES and SOGI modules reported varying levels of positive agreement that the experience could improve their relationships with their patients (8/11, 73% SOGI participants; 8/13, 62% race/SES participants) and avoid undesirable events in patient care (8/11, 73% SOGI participants; 8/13, 62% SES participants). Conversely, 9 of 11 SOGI participants (82%) showed adequate agreement with the belief that the training would help improve tailored care and 7 of 11 participants (64%) believed that the training would improve patient satisfaction. For the race/SES participants, 7 of 13 (54%) showed moderate agreement with beliefs about improving tailored care, while 6 of 13 (46%) agreed that the training could improve patient satisfaction. The race/SES participants showed higher agreement with the belief that the module would improve their community resources (9/13, 69%) compared to the SOGI module participants (5/11, 45%). A detailed summary of user experience findings is reported in Table 2.

Table . User experience (training reaction) outcomes.

Variables		Race/SES[a] (n=13)	SOGI[b] (n=11)	Effect size (Cohen *d*)
Augmented reality experience scores, mean (SD)				
	Feeling of "being there"[c]	4.6 (1.6)	3.9 (2.0)	0.40
	Realism of augmented reality[c]	3.7 (1.8)	2.9 (1.6)	0.49
	Relevance to job[d]	4.8 (0.5)	4.7 (0.5)	0.17
	Intention to apply augmented reality experience[d]	2.3 (1.1)	N/A[e]	N/A
Participants reporting applicability to job, n (%) [f]				
	Improve relationship with patients	8 (62)	8 (73)	N/A
	Improve patient satisfaction	6 (46)	7 (64)	N/A
	Improve tailored care	7 (54)	9 (82)	N/A
	Avoid undesirable events	8 (62)	8 (73)	N/A
	Improve community resources	9 (69)	5 (45)	N/A
	Other benefit	0 (0)	1 (9)	N/A

[a]SES: socioeconomic status.

[b]SOGI: sexual orientation and gender identity.

[c]Measured on a scale from 1=not at all to 7=very much.

[d]Measured on a scale from 1=strongly disagree to 5=strongly agree.

[e]N/A: not applicable.

[f]Multiple answers selected.

Preliminary Efficacy

For the skills questions, there was no significant difference in pre-post scores assessing the changes in awareness of implicit bias for patients of varying race/SES groups (pre: mean 4.31, SD 0.48; post: mean 4.46, SD 0.52; *d*=0.22; 95% CI −0.77 to 0.33) or the ability to manage health disparities caused by race/SES group (pre: mean 3.85, SD 0.56; post: mean 4.31, SD 0.48; *d*=0.52; 95% CI −1.10 to 0.07) . This pattern was true for measuring awareness of implicit bias for LGBTQ+ patients (pre: mean 4.36, SD 0.51; post: mean 4.73, SD 0.47; *d*=0.54; 95% CI −1.16 to 0.11). For minimizing health disparities realted to LGBTQ+ status, there was a significant difference between pre- and posttest scores (pre: mean 3.91, SD 0.94; post: mean 4.64, SD 0.51; *d*=0.72; 95% CI −1.38 to −0.04) with participants scoring closer to "strongly agree" after experiencing the AR experience.

For the attitudinal questions, there were nonsignificant improvements in the race/SES module in assessing how personal values affected patients (pre: mean 3.92, SD 0.95; post: mean 4.31, SD 0.84; *d*=0.44; 95% CI −1.01 to 0.14), how institutional barriers affect patients (pre: mean 4.23, SD 0.60; post: mean 4.31, SD 0.48; *d*=0.12; 95% CI −0.66 to 0.43), and participants' ability to identify reactions based on stereotypes (pre: mean 4.15, SD 0.56; post: mean 4.38, SD 0.51; *d*=0.53; 95% CI −1.10 to 0.07). For the SOGI module, changes from pre to posttraining were also nonsignificant for all attitudinal items (pre: mean 4.18, SD 0.87; post: mean 4.45, SD 0.52; *d*=0.30; 95% CI −0.90

to 0.31), institutional barrier items (pre: mean 4.36, SD 0.67; post: mean 4.36, SD 0.67; *d*=0.00; 95% CI −0.59 to 0.59), and items related to identifying stereotypical reactions (pre: mean 4.36, SD 0.51; post: mean 4.36, SD 0.92; *d*=0.00; 95% CI −0.59 to 0.59). A summary of preliminary efficacy findings for each module can be found in Table 3.

Table . Race/socioeconomic status (SES) pre- and posttest skills and attitude outcomes (n=13).

Variables		Pretest score, mean (SD)	Posttest score, mean (SD)	Cohen *d* (95% CI)
Skills questions				
	Implicit bias (race/SES)	4.3 (0.5)	4.5 (0.5)	0.22 (–0.77 to 0.33)
	Minimize health disparities (race/SES)	3.9 (0.6)	4.3 (0.5)	0.52 (–1.10 to 0.07)
Attitudinal questions				
	How my values affect patients	3.9 (1.0)	4.3 (0.5)	0.44 (–1.01 to 0.14)
	How institutional barriers affect patients	4.2 (0.6)	4.3 (0.5)	0.12 (–0.66 to 0.43)
	Identify reactions based on stereotypes	4.2 (0.6)	4.4 (0.5)	0.53 (–1.10 to 0.07)

Discussion

This study developed a VR and AR implicit association training program for Medicaid providers based on cognitive psychology and the latest mobile simulation technologies. Designed to improve awareness of implicit biases related to patients' SES and sexual orientation/gender identity, learners are able to complete six 5-minute interactive role-playing scenarios on their smartphones. Results of pilot user experience research among 18 participants found adequate acceptability and preliminary efficacy (ie, a nonsignificant increase in most outcomes) of the VARIAT program. These findings are consistent with recent literature in cognitive psychology about the possible benefits of AR interventions for health care providers [34-37].

While researchers have spent the last 20 years attempting to reduce implicit bias [38-41], few attempts have been made to integrate the latest immersive technologies, such as AR and VR, with provider-level implicit bias training. For example, a recent meta-analysis of 492 interventions on implicit biases found only a handful of studies attempting to change implicit bias used any kind of VR or AR [42]. Narrowing down to implicit bias training in the health care setting, another recent literature review found few studies that focused on addressing bias at the provider level [43-46]. Therefore, while implicit bias in health care more broadly has been long recognized as a prominent issue [3], there is an important gap in research that develops technology-assisted training programs so that such programs can be more readily available for health care providers and so that implicit bias training can be received at a time and location that works best for them rather than having to attend in-person training sessions. The VARIAT program reported in this study addresses this critical literature gap by offering a convenient and publicly available program that can be integrated into medical training for health care professionals interacting with Medicaid patients, whose training may have important beneficial impacts on patients from disadvantaged backgrounds and those who experience reduced access to high quality of care due to multiple individual and societal barriers [47]. For example, the VARIAT program is brief and can be completed on a mobile device during "fragmented" time windows that fit within the often-chaotic work schedule of medical professionals. Therefore, medical institutions may consider integrating the VARIAT training as a regular refresh of lengthier and more comprehensive in-person or on-site bias-reduction training for their health care professional teams.

Furthermore, among the studies that focused on mitigating health care provider biases, few documented detailed feasibility and efficacy data [48-51]. This study is among the first in the literature to measure both positive provider reactions and efficacy outcomes at multiple levels, including user experience with AR, perceived utility in users' professional work, and perceived attitudes toward patients and skills in mitigating implicit biases at work. It was interesting to find that although the study participants perceived relatively high levels of immersion ("being there"), AR realism, and job relevance from the VARIAT training, they expressed low levels of intention to apply this experience to their daily work. One possible explanation for this discrepancy might be the challenges of translating learned knowledge to behavioral changes, as commonly seen in educational interventions, potentially due to the limited scenarios provided by the training compared to the broad variations in participants' own daily work experiences. The collection of both pre- and postintervention efficacy outcomes further allowed us to measure the potential interventional effects of each of the VARIAT training modules (race/SES and SOGI). However, it should be noted that this paper focused primarily on sharing with the scientific community the development processes and design considerations of a novel implicit bias training program for Medicaid providers. Therefore, caution should be applied when interpreting the preliminary results of this pilot user experience study.

Study Limitations

There are several important limitations to this study. First, the current iteration of the VARIAT program is being delivered on mobile devices. This training program might elicit different user experiences and efficacy outcomes should it be delivered on other platforms such as through an immersive VR headset. Second, the study sample for this user experience testing study was small and potentially unbalanced. Larger sample sizes and a more rigorous study design (eg, a randomized controlled trial) should be used in future research to formally evaluate the efficacy of the VARIAT program with sufficient statistical power and without inflating the type II error rate [52,53]. Third, the present version of the VARIAT program only consisted of 2 modular domains for implicit bias training, race/SES and SOGI, with only 3 training scenarios for each module due to limitations on study resources and team expertise. Further, although these modules were developed by an interdisciplinary team of clinicians and researchers, patient communities were not involved in the design process. Future research will invite patient advisory groups into the development and refinement process of additional modules and scenarios for VARIAT to provide training in more comprehensive implicit bias domains during clinician-patient interactions. Fourth, this study used only self-reported measures developed by the study team to assess the efficacy outcomes, which may not be able to accurately measure biases that are inherently "implicit." Future efficacy trials of the VARIAT program (and interventions alike) should incorporate validated implicit bias assessment tools such as the Implicit Association Test (IAT), which has been increasingly used by health care professionals in the existing literature [54]. Finally, several limitations of the study design should be noted. For example, this study did not restrict or record the number of times participants were allowed to undergo the training, which may have impacted usability and efficacy outcomes. Additionally, this study used an immediate pre-post training design. A more distant posttraining evaluation should be conducted to allow for examination of the impact of the modules on biases over time.

Conclusions

This study presents a novel intervention (VARIAT) that uses immersive mobile technology to improve awareness of implicit bias related to race/SES and SOGI among Medicaid providers. This publicly available training program has found a promising avenue for future research and practice in reducing implicit bias in health care workplaces. Future research should be conducted to formally evaluate the VARIAT program with large samples and implicit bias testing measures, as well as incorporate additional training domains to provide impactful benefits to both health care professionals and their patients.

Acknowledgments

This Medicaid Equity Simulation Project is funded by the Ohio Department of Medicaid and administered by the Ohio Colleges of Medicine Government Resource Center. The views expressed in this report are solely those of the authors and do not represent the views of the state of Ohio or federal Medicaid programs. Research efforts of the first (JS) and second (AJC) authors for this paper were partly supported by the Richard K. and Nancy L. Donahue Humanities Ethics Fellowship (University of Massachusetts Lowell).

Conflicts of Interest

J Penka is the founder and CEO and J Patterson is the cofounder of LittleSeed Inc. Both authors are members of the board of directors of LittleSeed Inc.

Multimedia Appendix 1
Race and socioeconomic status (SES) - module 1.
[DOCX File, 24 KB - games_v12i1e51310_app1.docx]

Multimedia Appendix 2
Sexual orientation gender identity (SOGI) - module 2.
[DOCX File, 23 KB - games_v12i1e51310_app2.docx]

Multimedia Appendix 3
User workflow.
[DOCX File, 959 KB - games_v12i1e51310_app3.docx]

References

1. Blair IV, Steiner JF, Havranek EP. Unconscious (implicit) bias and health disparities: where do we go from here? Perm J 2011;15(2):71-78. [doi: 10.7812/TPP/11.979] [Medline: 21841929]
2. Xu K, Nosek B, Greenwald A. Psychology data from the Race Implicit Association Test on the Project Implicit Demo website. J Open Psychol Data 2014 Mar 18;2(1):e3. [doi: 10.5334/jopd.ac]
3. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. BMC Med Ethics 2017 Mar 1;18(1):19. [doi: 10.1186/s12910-017-0179-8] [Medline: 28249596]

4. Bhat A, Mahajan V, Wolfe N. Implicit bias in stroke care: a recurring old problem in the rising incidence of young stroke. J Clin Neurosci 2021 Mar;85:27-35. [doi: 10.1016/j.jocn.2020.12.017] [Medline: 33581786]
5. Chapman EN, Kaatz A, Carnes M. Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities. J Gen Intern Med 2013 Nov;28(11):1504-1510. [doi: 10.1007/s11606-013-2441-1] [Medline: 23576243]
6. Dovidio JF, Eggly S, Albrecht TL, Hagiwara N, Penner L. Racial biases in medicine and healthcare disparities. TPM Test Psychom Methodol Appl Psychol 2016;23(4):489-510. [doi: 10.4473/TPM23.4.5]
7. Hall WJ, Chapman MV, Lee KM, et al. Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. Am J Public Health 2015 Dec;105(12):e60-e76. [doi: 10.2105/AJPH.2015.302903] [Medline: 26469668]
8. Saluja B, Bryant Z. How implicit bias contributes to racial disparities in maternal morbidity and mortality in the United States. J Womens Health (Larchmt) 2021 Feb;30(2):270-273. [doi: 10.1089/jwh.2020.8874] [Medline: 33237843]
9. Shen J, Carmichael A, Clinton AJ. A scoping review of research on potential impact of implicit bias in healthcare settings for children with acquired brain injuries. Rehabil Psychol 2024 Feb;69(1):36-44. [doi: 10.1037/rep0000519] [Medline: 37796578]
10. Lelutiu-Weinberger C, Clark KA, Pachankis JE. Mental health provider training to improve LGBTQ competence and reduce implicit and explicit bias: a randomized controlled trial of online and in-person delivery. Psychol Sex Orientat Gend Divers 2023 Dec;10(4):589-599. [doi: 10.1037/sgd0000560] [Medline: 38239562]
11. McDowell MJ, Goldhammer H, Potter JE, Keuroghlian AS. Strategies to mitigate clinician implicit bias against sexual and gender minority patients. Psychosomatics 2020;61(6):655-661. [doi: 10.1016/j.psym.2020.04.021] [Medline: 32641233]
12. Sabin JA, Riskind RG, Nosek BA. Health care providers' implicit and explicit attitudes toward lesbian women and gay men. Am J Public Health 2015 Sep;105(9):1831-1841. [doi: 10.2105/AJPH.2015.302631] [Medline: 26180976]
13. Wolsiefer KJ, Mehl M, Moskowitz GB, et al. Investigating the relationship between resident physician implicit bias and language use during a clinical encounter with Hispanic patients. Health Commun 2023 Jan;38(1):124-132. [doi: 10.1080/10410236.2021.1936756] [Medline: 34130567]
14. Ivers NN, Johnson DA, Rogers JL. The association between implicit racial bias and mindfulness in mental health practitioners. J Couns Dev 2021 Jan;99(1):11-23. [doi: 10.1002/jcad.12350]
15. Hagiwara N, Elston Lafata J, Mezuk B, Vrana SR, Fetters MD. Detecting implicit racial bias in provider communication behaviors to reduce disparities in healthcare: challenges, solutions, and future directions for provider communication training. Patient Educ Couns 2019 Sep;102(9):1738-1743. [doi: 10.1016/j.pec.2019.04.023] [Medline: 31036330]
16. Maina IW, Belton TD, Ginzberg S, Singh A, Johnson TJ. A decade of studying implicit racial/ethnic bias in healthcare providers using the implicit association test. Soc Sci Med 2018 Feb;199:219-229. [doi: 10.1016/j.socscimed.2017.05.009] [Medline: 28532892]
17. Sabin JA, Rivara FP, Greenwald AG. Physician implicit attitudes and stereotypes about race and quality of medical care. Med Care 2008 Jul;46(7):678-685. [doi: 10.1097/MLR.0b013e3181653d58] [Medline: 18580386]
18. Maxfield CM, Thorpe MP, Desser TS, et al. Awareness of implicit bias mitigates discrimination in radiology resident selection. Med Educ 2020 Jul;54(7):637-642. [doi: 10.1111/medu.14146] [Medline: 32119145]
19. Johnson TJ, Winger DG, Hickey RW, et al. Comparison of physician implicit racial bias toward adults versus children. Acad Pediatr 2017 Mar;17(2):120-126. [doi: 10.1016/j.acap.2016.08.010] [Medline: 27620844]
20. Sabin JA, Greenwald AG. The influence of implicit bias on treatment recommendations for 4 common pediatric conditions: pain, urinary tract infection, attention deficit hyperactivity disorder, and asthma. Am J Public Health 2012 May;102(5):988-995. [doi: 10.2105/AJPH.2011.300621] [Medline: 22420817]
21. Rojas M, Walker-Descartes I, Laraque-Arena D. An experimental study of implicit racial bias in recognition of child abuse. Am J Health Behav 2017 May 1;41(3):358-367. [doi: 10.5993/AJHB.41.3.15] [Medline: 28376980]
22. Amdani S, Bhimani SA, Boyle G, et al. Racial and ethnic disparities persist in the current era of pediatric heart transplantation. J Card Fail 2021 Sep;27(9):957-964. [doi: 10.1016/j.cardfail.2021.05.027] [Medline: 34139364]
23. Simon SL, Clay D, Chandrasekhar J, Duncan CL. Gender bias in pediatric psychology. Clin Pract Pediatr Psychol 2021;9(1):82-95. [doi: 10.1037/cpp0000378]
24. Seibel LM, Fehr FC, Sarwal S, Panchuk HJ. Impact of cognitive rehearsal training over time for new registered nurses. J Contin Educ Nurs 2022 May;53(5):203-211. [doi: 10.3928/00220124-20220414-02] [Medline: 35510923]
25. Kang J, Jeong YJ. Effects of a smartphone application for cognitive rehearsal intervention on workplace bullying and turnover intention among nurses. Int J Nurs Pract 2019 Dec;25(6):e12786. [doi: 10.1111/ijn.12786] [Medline: 31523905]
26. Teal CR, Shada RE, Gill AC, et al. When best intentions aren't enough: helping medical students develop strategies for managing bias about patients. J Gen Intern Med 2010 May;25 Suppl 2(Suppl 2):S115-S118. [doi: 10.1007/s11606-009-1243-y] [Medline: 20352504]
27. Gonzalez CM, Walker SA, Rodriguez N, Noah YS, Marantz PR. Implicit bias recognition and management in interpersonal encounters and the learning environment: a skills-based curriculum for medical students. MedEdPORTAL 2021;17:11168. [doi: 10.15766/mep_2374-8265.11168] [Medline: 34277934]
28. Sukhera J, Watling CJ, Gonzalez CM. Implicit bias in health professions: from recognition to transformation. Acad Med 2020 May;95(5):717-723. [doi: 10.1097/ACM.0000000000003173] [Medline: 31977339]

29. McMichael B, Nickel A, Duffy EA, et al. The impact of health equity coaching on patient's perceptions of cultural competency and communication in a pediatric emergency department: an intervention design. J Patient Exp 2019 Dec;6(4):257-264. [doi: 10.1177/2374373518798111] [Medline: 31853480]
30. Long A, Jennings J, Bademosi K, et al. Storytelling to improve healthcare worker understanding, beliefs, and practices related to LGBTQ + patients: a program evaluation. Eval Program Plann 2022 Feb;90:101979. [doi: 10.1016/j.evalprogplan.2021.101979] [Medline: 34275639]
31. Zipp SA, Craig SD. The impact of a user's biases on interactions with virtual humans and learning during virtual emergency management training. Educ Technol Res Dev 2019 Dec;67(6):1385-1404. [doi: 10.1007/s11423-019-09647-6]
32. Tang KS, Cheng DL, Mi E, Greenberg PB. Augmented reality in medical education: a systematic review. Can Med Educ J 2020 Mar;11(1):e81-e96. [doi: 10.36834/cmej.61705] [Medline: 32215146]
33. Dhar P, Rocks T, Samarasinghe RM, Stephenson G, Smith C. Augmented reality in medical education: students' experiences and learning outcomes. Med Educ Online 2021 Dec;26(1):1953953. [doi: 10.1080/10872981.2021.1953953] [Medline: 34259122]
34. Chiang DH, Huang CC, Cheng SC, et al. Immersive virtual reality (VR) training increases the self-efficacy of in-hospital healthcare providers and patient families regarding tracheostomy-related knowledge and care skills: a prospective pre-post study. Medicine (Baltimore) 2022 Jan 14;101(2):e28570. [doi: 10.1097/MD.0000000000028570] [Medline: 35029229]
35. Croghan IT, Hurt RT, Aakre CA, et al. Virtual reality for health care professionals during a pandemic: a pilot program. J Prim Care Community Health 2022;13:21501319221086716. [doi: 10.1177/21501319221086716] [Medline: 35352605]
36. Hayakawa J, Barrows J, See S, Schomberg J. Effects of classical music virtual reality on pediatric healthcare worker compassion fatigue. J Nurs Adm 2022 May 1;52(5):280-285. [doi: 10.1097/NNA.0000000000001148] [Medline: 35467593]
37. Beverly E, Hommema L, Coates K, et al. A tranquil virtual reality experience to reduce subjective stress among COVID-19 frontline healthcare workers. PLoS One 2022;17(2):e0262703. [doi: 10.1371/journal.pone.0262703] [Medline: 35139092]
38. Barnes-Holmes D, Barnes-Holmes Y, Power P, Hayden E, Milne R, Stewart I. Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit beliefs. Ir Psychol 2006;32(7):169-177 [FREE Full text]
39. Greenwald AG, Poehlman TA, Uhlmann EL, Banaji MR. Understanding and using the implicit association test: III. Meta-analysis of predictive validity. J Pers Soc Psychol 2009 Jul;97(1):17-41. [doi: 10.1037/a0015575] [Medline: 19586237]
40. Lai CK, Marini M, Lehr SA, et al. Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. J Exp Psychol Gen 2014 Aug;143(4):1765-1785. [doi: 10.1037/a0036260] [Medline: 24661055]
41. Whitford DK, Emerson AM. Empathy intervention to reduce implicit bias in pre-service teachers. Psychol Rep 2019 Apr;122(2):670-688. [doi: 10.1177/0033294118767435] [Medline: 29621945]
42. Forscher PS, Lai CK, Axt JR, et al. A meta-analysis of procedures to change implicit measures. J Pers Soc Psychol 2019 Sep;117(3):522-559. [doi: 10.1037/pspa0000160] [Medline: 31192631]
43. Ricks TN, Abbyad C, Polinard E. Undoing racism and mitigating bias among healthcare professionals: lessons learned during a systematic review. J Racial Ethn Health Disparities 2022 Oct;9(5):1990-2000. [doi: 10.1007/s40615-021-01137-x] [Medline: 34480317]
44. Gran-Ruaz S, Feliciano J, Bartlett A, Williams MT. Implicit racial bias across ethnoracial groups in Canada and the United States and Black mental health. Can Psychol 2022;63(4):608-622. [doi: 10.1037/cap0000323]
45. Vo JB, Gillman A, Mitchell K, Nolan TS. Health disparities: impact of health disparities and treatment decision-making biases on cancer adverse effects among Black cancer survivors. Clin J Oncol Nurs 2021 Oct 1;25(5):17-24. [doi: 10.1188/21.CJON.S1.17-24] [Medline: 34533532]
46. Newcomb AB, Rothberg S, Zewdie M, et al. Managing patient bias: teaching residents to navigate racism and bias in the workplace. J Surg Educ 2021;78(6):1791-1795. [doi: 10.1016/j.jsurg.2021.06.007] [Medline: 34284944]
47. Hsiang WR, Lukasiewicz A, Gentry M, et al. Medicaid patients have greater difficulty scheduling health care appointments compared with private insurance patients: a meta-analysis. Inquiry 2019;56:46958019838118. [doi: 10.1177/0046958019838118] [Medline: 30947608]
48. Alexis C, White AM, Thomas C, Poweski J. Unpacking reactions to white privilege among employees of an academic medical center. J Cult Divers 2019;26(1) [FREE Full text]
49. Crawford C. From the American Academy of Family Physicians: the Everyone Project unveils Implicit Bias Training Guide. Ann Fam Med 2020;18(2):182. [doi: 10.1370/afm.2525]
50. Nelson SC, Prasad S, Hackman HW. Training providers on issues of race and racism improve health care equity. Pediatr Blood Cancer 2015 May;62(5):915-917. [doi: 10.1002/pbc.25448] [Medline: 25683782]
51. Sherman MD, Ricco J, Nelson SC, Nezhad SJ, Prasad S. Implicit bias training in a residency program: aiming for enduring effects. Fam Med 2019 Sep;51(8):677-681. [doi: 10.22454/FamMed.2019.947255] [Medline: 31509218]
52. Guo Y, Logan HL, Glueck DH, Muller KE. Selecting a sample size for studies with repeated measures. BMC Med Res Methodol 2013 Jul 31;13:100. [doi: 10.1186/1471-2288-13-100] [Medline: 23902644]
53. Overall JE, Starbuck RR. Sample size estimation for randomized pre-post designs. J Psychiatr Res 1979;15(1):51-55. [doi: 10.1016/0022-3956(79)90006-2] [Medline: 430441]

54. Sukhera J, Wodzinski M, Rehman M, Gonzalez CM. The Implicit Association Test in health professions education: a meta-narrative review. Perspect Med Educ 2019 Oct;8(5):267-275. [doi: 10.1007/s40037-019-00533-8] [Medline: 31535290]

Abbreviations

AR: augmented reality
LGBTQ+: lesbian, gay, bisexual, transgender, queer
SES: socioeconomic status
SOGI: sexual orientation and gender identity
VARIAT: Virtual and Augmented Reality-based Implicit Association Training
VR: virtual reality

Edited by G Eysenbach, T Leung; submitted 29.07.23; peer-reviewed by E Chuang, O Joseph, JL Guo; revised version received 24.10.23; accepted 12.12.23; published 07.03.24.

Please cite as:
Shen J, Clinton AJ, Penka J, Gregory ME, Sova L, Pfeil S, Patterson J, Maa T
Smartphone-Based Virtual and Augmented Reality Implicit Association Training (VARIAT) for Reducing Implicit Biases Toward Patients Among Health Care Providers: App Development and Pilot Testing
JMIR Serious Games 2024;12:e51310
URL: https://games.jmir.org/2024/1/e51310
doi:10.2196/51310

Efficacy of a Virtual 3D Simulation–Based Digital Training Module for Building Dental Technology Students' Long-Term Competency in Removable Partial Denture Design: Prospective Cohort Study

KeXin Liu[1,2,3,*], BDS; YaQian Xu[1,2,3,*], BS; ChaoYi Ma[1,2,3], MMSc; Na Yu[1,2,3], MMSc; FaBing Tan[1,2,3], DDS; Yi Li[1,2,3], BS; YaXin Bai[1,2,3], BS; XiaoMing Fu[1,2,3], DDS; JiaWu Wan[4], BS; DongQi Fan[1,2,3], BDS; HuBin Yin[1,2,3], BS; MeiXi Chen[1,2,3], BS; HongJi Chen[1,2,3], BS; Lin Jiang[1,2,3], DDS; JinLin Song[1], DDS; Ping Ji[1,2,3], DDS; XiaoHan Zhao[5,*], MD; MengWei Pang[1,2,3,*], MDS

1

2

3

4

5

*these authors contributed equally

Corresponding Author:
MengWei Pang, MDS

Related Article:

This is a corrected version. See correction statement: https://games.jmir.org/2024/1/e63113

Abstract

Background: Removable partial denture (RPD) design is crucial to long-term success in dental treatment, but shortcomings in RPD design training and competency acquisition among dental students have persisted for decades. Digital production is increasing in prevalence in stomatology, and a digital RPD (D-RPD) module, under the framework of the certified Objective Manipulative Skill Examination of Dental Technicians (OMEDT) system reported in our previous work, may improve on existing RPD training models for students.

Objective: We aimed to determine the efficacy of a virtual 3D simulation–based progressive digital training module for RPD design compared to traditional training.

Methods: We developed a prospective cohort study including dental technology students at the Stomatology College of Chongqing Medical University. Cohort 1 received traditional RPD design training (7 wk). Cohort 2 received D-RPD module training based on text and 2D sketches (7 wk). Cohort 3 received D-RPD module pilot training based on text and 2D sketches (4 wk) and continued to receive training based on 3D virtual casts of real patients (3 wk). RPD design tests based on virtual casts were conducted at 1 month and 1 year after training. We collected RPD design scores and the time spent to perform each assessment.

Results: We collected the RPD design scores and the time spent to perform each assessment at 1 month and 1 year after training. The study recruited 109 students, including 58 (53.2%) female and 51 male (56.8%) students. Cohort 1 scored the lowest and cohort 3 scored the highest in both tests (cohorts 1-3 at 1 mo: mean score 65.8, SD 21.5; mean score 81.9, SD 6.88; and mean score 85.3, SD 8.55, respectively; *P*<.001; cohorts 1-3 at 1 y: mean score 60.3, SD 16.7; mean score 75.5, SD 3.90; and mean score 90.9, SD 4.3, respectively; *P*<.001). The difference between cohorts in the time spent was not statistically significant at 1 month (cohorts 1-3: mean 2407.8, SD 1370.3 s; mean 1835.0, SD 1329.2 s; and mean 1790.3, SD 1195.5 s, respectively; *P*=.06) but was statistically significant at 1 year (cohorts 1-3: mean 2049.16, SD 1099.0 s; mean 1857.33, SD 587.39 s; and mean 2524.3, SD 566.37 s, respectively; *P*<.001). Intracohort comparisons indicated that the differences in scores at 1 month and 1 year were not statistically significant for cohort 1 (95% CI –2.1 to 13.0; *P*=.16), while cohort 3 obtained significantly higher scores 1 year later (95% CI 2.5-8.7; *P*=.001), and cohort 2 obtained significantly lower scores 1 year later (95% CI –8.8 to –3.9; *P*<.001).

Conclusions: Cohort 3 obtained the highest score at both time points with retention of competency at 1 year, indicating that progressive D-RPD training including virtual 3D simulation facilitated improved competency in RPD design. The adoption of D-RPD training may benefit learning outcomes.

(JMIR Serious Games 2024;12:e46789) doi:10.2196/46789

KEYWORDS

removable partial denture; RPD; virtual simulation; dental technology; computer-aided design; CAD; clinical practice; efficacy; cohort study; digital training; training; dentistry; treatment; design; virtual; assessment

Introduction

The partially edentulous population is increasing because of increased life expectancy and an aging population [1]. Removable partial dentures (RPDs) possess the advantages of cost-effectiveness and needing a less invasive procedure compared to fixed and implant-retained restorations; thus, RPDs remain an attractive treatment option for partially edentulous patients [2].

The design of RPDs is a crucial technical step that greatly impacts the long-term success of dental treatment and warrants high standards due to the complex structure and variation in the oral morphology of individual patients [3,4]. Poor RPD design can exacerbate plaque retention, leading to gingivitis, periodontitis, and other oral diseases [5]. RPD design has traditionally been a complex subject to teach and learn [6]. Unfortunately, shortcomings in RPD design training and competency acquisition among dental students have persisted for decades [7]. The lack of student supervision by qualified instructors and progressive training patterns, as well as the absence of practice on real patients, have been found to be the main factors limiting successful training in RPD design [8,9]. The lack of competency in RPD design can hamper clinical practice among dentists, often leading to the assignment of the task to dental technicians. Dental technicians, however, lack direct observation of the oral soft and hard tissues of the patients. This factor can limit the quality of prosthesis design and can cause patient discomfort, resulting in additional repairs and medical disputes [10].

Using a pencil-drawn design of the RPD framework on a physical cast or a paper prescription has always been the classic approach for teaching RPD design in most dental schools [11]. However, this classic approach is marked by several constraints. The cumbersome processes used where teachers collect, rate, and hand out paper prescriptions can result in communication gaps and potential wastage of time [11]. The COVID-19 pandemic has further limited the availability of real, patient-based physical casts, thus eroding practice time for RPD design on patient models [12]. Although advances in the dental laboratory digital workflow facilitate the use of computer-aided design (CAD) and computer-aided manufacturing (CAM) in the fabrication of RPDs and communication between dentists and dental technicians [13], multiple surveys confirm that CAD/CAM RPD design courses continue to present significant barriers to widespread adoption in dental education settings due to the cost, lack of faculty, and lack of time available within the curriculum. Moreover, the education editions of commercial CAD software programs for dental laboratories remain expensive and require instructors proficient in CAD/CAM technology to facilitate teaching. Furthermore, the learning curve to master the skills of using commercial CAD software is steep and requires a long time commitment, which presents a problem in undergraduate dental education settings.

In our previous work, we reported a digital RPD (D-RPD) module under the framework of the certified Objective Manipulative Skill Examination of Dental Technicians (OMEDT) system, which is a free web-based application for computer-aided drawing and 2D sketch–based RPD design training for dental and dental technology students [14]. This prospective cohort study aimed to report a significant update to the D-RPD module and to further explore the optimal design of the D-RPD module for teaching. We specifically asked the following questions: (1) How can a progressive approach using case-based virtual 3D simulation be incorporated in a D-RPD design training module to better prepare students for the needs of practice? (2) What is the efficacy of digital training approaches in RPD design compared to traditional training? (3) Does a virtual 3D simulation–based progressive digital training module benefit long-term RPD design competency acquisition and retention?

Methods

Development of a Progressive D-RPD Module Incorporating Case-Based Virtual 3D Simulation

The virtual 3D simulation was based on casts from actual patients. In order to construct patient-based virtual casts, a desktop portable application, showModels, has been developed with the Unity engine and C++ version 11 and C# version 4.0. All clinical cases used in the RPD design training were collected from the Dental Technology Laboratory of the Stomatology Hospital of Chongqing Medical University. Virtual casts were constructed from physical plaster casts of clinical patients using LabScanner (E4; 3shape) and saved in the stereolithography file format using Format Converter (Autodesk; Delcam Exchange) to remove possible surface texture indicators. Since any prepared rest seats on a patient's physical casts may provide hints for RPD design, such rest seats on the virtual cast were filled using 3D reverse software (Geomagic Wrap; 3D System). The resultant virtual casts of real patients may be rotated or zoomed in and out to view the cast details, and the user may specify whether to display the maxillary or the mandibular cast (Multimedia Appendix 1).

Participants and Recruitment

Eligible participants (junior students majoring in dental technology) were recruited at the Stomatology College of Chongqing Medical University. The RPD design theory curriculum in dental technology was organized by the Stomatology Hospital of Chongqing Medical University. All participants provided signed informed consent. The prospective cohort study began in September 2020 and ended in September 2022.

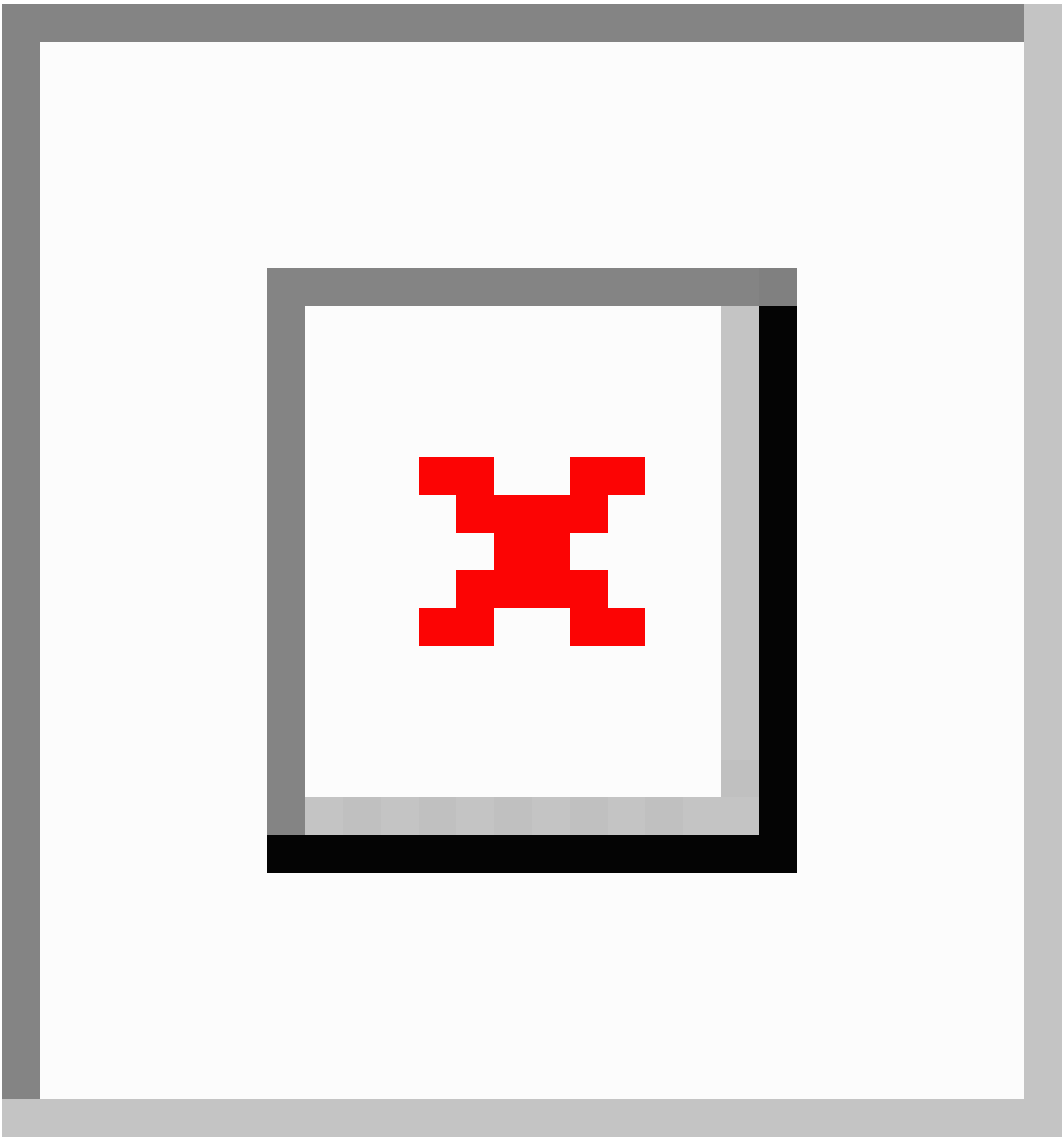
Ethical Considerations

The Research Ethics Committee of the Affiliated Hospital of Stomatology, Chongqing Medical University, approved this study protocol (COHS-REC-2022; LSNo. 096). Data reporting followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cohort studies.

Intervention Design

The study protocol and the participant flow diagram are depicted in Figure 1. A description of the training methods implemented is presented in Multimedia Appendix 2. In brief, after following the same RPD design theory curriculum, all participants were divided into 3 cohorts. Cohort 1 included 43 participants who received traditional RPD design training for 7 weeks. They received an RPD design task from the principal investigator each Monday, completed the RPD design using a paper prescription and red and blue pencil, and submitted it by Sunday. Cohort 2 included 36 participants who received D-RPD module training based on literal descriptions and 2D sketches for 7 weeks. They received literal case descriptions and 2D sketches for uniformly depicting missing teeth issued by the principal investigator in the D-RPD module every Monday, drew RPD designs using the D-RPD module, and submitted designs by Sunday. Cohort 3 included 30 participants who received D-RPD module training based on a literal description and 2D sketches for 4 weeks and continued to receive progressive instruction with the updated D-RPD module training based on the virtual casts of real patients for 3 weeks. The 7 RPD design tasks received by the 3 cohorts were all the same, and the types of dentition defects covered Kennedy classes I, II, III, and IV, with only some differences in presentation form. We set 1 month as the retention interval to avoid temporary effects from practice [15]. At 1 month and at 1 year after the training, RPD design tests using 3D virtual casts were administered using the updated D-RPD module and carried out for all of the cohorts. During the retention interval, participants' D-RPD module accounts were blocked to prevent participants from using the module for additional training. Within 1 year of completing their training, participants start a uniform dental laboratory internship, and the internship outline has uniform requirements for the design of RPDs with the same workload. For cohorts 1 and 2, a separate D-RPD module introductory session was held prior to the testing to ensure that the cohort could successfully complete the RPD design task using the updated D-RPD module.

Figure 1. The flowchart of this prospective cohort study. D-RPD: digital removable partial denture; RPD: removable partial denture.



Recruitment of the Expert Panel and Development of the Scoring Rubrics

The principal investigator recruited an expert panel to develop the scoring rubrics [16] (Table 1), and the exercises of all 3 cohorts were rated accordingly. The expert panel consisted of a dental technician experienced in the field of RPD manufacturing and a clinical prosthodontist recruited from the Stomatology Hospital of Chongqing Medical University. The expert panel was blinded to the cohort assignments, had not participated in the teaching of the participants, and did not know about the participants' major or nature of the intervention.

Table . The scoring rubric used to assess the removable partial denture design test task.

Scoring component	Met clinically acceptable criteria	Needs improvement	Clinically unacceptable
Case observation (20 points)	The missing tooth position was identified accurately and marked correctly on the drawing (20 points).	N/A[a]	The missing tooth position was identified inaccurately, marked incorrectly on the drawing, or both (0 points).
Design choices (40 points)	Design choices are ideal for the case (28-40 points).	Design choices have some flaws but are adequate (15-27 points). Examples include the following: no missing component; indirect retainer present but not in the optimal position; design choices do not violate biological principles; clasp choice adequate but not optimal for the case; inappropriate choice or extension of major connector; justified use of clasps/rests but excessive framework components.	Any missing component or inappropriate design for the case (0-14 points). Examples include the following: missing indirect retainer in a case requiring one; missing reciprocation; clasp choice inappropriate for situation; design choices violate biological principles; excessive and unjustified use of clasps/rests.
Drawing (20 points)	Drawing is ideal. Metal components are painted in blue and resin bases are in red (14-20 points).	Drawing has some flaws but is adequate (7-13 points). Examples include the following: components are represented by corresponding colors; minor inadequacy or inconsistency of spacing between components; components are occasionally not connected; the finish line is not drawn.	Drawing has major flaws (0-6 points). Examples include the following: components are not represented by corresponding colors; major inadequacy or inconsistency of spacing between components; component positioning significantly off optimal position; any component position that violates biomechanical design principles; components are frequently not connected; the finish line is not drawn.
Consistency with task description (10 points)	Exactly as described in the task description (8-10 points). Criteria include the following: clearly presents the requirements implied in the description, and the design is well aligned with the corresponding description; gives consideration to both aesthetics and function.	Some deviation from the task description, but it is acceptable (5-7 points). Examples include the following: conventional design carried out without addressing case-specific modifying factors or requirements listed in the task description; only function considered, consideration for aesthetics lacking.	Serious violation or deviation from the task description (0-4 points). Examples include the following: the design does not match the task description; lack of aesthetic and functional considerations.
Neatness and accuracy in presentation (10 points)	Neat and accurate, no inconsistencies between the table and drawing (8-10 points).	Some inaccuracy and neatness flaws, but it is adequate (5-7 points). Examples include the following: minor erasures; minor neatness issues but still legible.	Major inaccuracy and neatness flaws (0-4 points). Examples include the following: missing information; major neatness issues; writing not legible; any inconsistencies between the table and drawing.

[a]N/A: not applicable.

Data Collection

The main metrics collected were the time(s) to complete the RPD design exercise and the RPD design score (100 points) based on the scoring rubrics by the expert panel.

Statistical Analysis

Scores (ie, total points for each assessment) and time (ie, seconds needed to perform each assessment) were summarized descriptively as means and SDs, coefficients of variation (defined as the ratio of the SD to the mean), and IQRs [17]. Due to the small sample size, normality was tested using the Shapiro-Wilk test, and the homogeneity of variance was tested using the *F* test [18]. The results showed that the time and score data had a skewed distribution and heterogeneity of variance. Therefore, the nonparametric method was used to compare the data sets in this study. Since the Kruskal-Wallis test is widely used to determine whether 3 or more independent data sets are different on some variable of interest [19], it was used to compare the cohorts at the same time point (1 mo or 1 y later), using the 3 data sets in each analysis process. When the value of the Kruskal-Wallis statistic is calculated as statistically significant, it indicates that at least 1 of the compared groups is different from the others. Therefore, we chose the Bonferroni method for further analysis with pairwise multiple comparisons to locate the source of significance. As for in-cohort comparisons at different time points, the Wilcoxon matched-pairs test is a frequently used nonparametric test for paired data, especially for nonnormal data and categorical data, such as was present in this cohort study. Hypothesis tests were 2 sided with a significance threshold of P=.05. At the same time, when multiple sets of data are being processed and compared simultaneously, there is increased risk of a type I error, so to identify significant correlations, threshold levels of significance for correlation coefficients were adjusted for multiple comparisons; we used a set of κ correlation coefficients with Bonferroni correction to strictly control the occurrence of false positives (after Bonferroni correction, we used a significance threshold of *P*=.016) [20]. Statistical analysis was performed using SPSS (version 26.0; IBM Corp).

Results

This cohort study included 109 participants: 58 (53.2%) women and 51 (56.8%) men, with a mean age at the beginning the study in September 2020 of 22.5 (SD 0.7) years (Table 2). All 3 cohorts completed the experiment.

Table . Baseline characteristics of the participants in this cohort study. *P* values were estimated using the Kruskal-Wallis test.

Characteristics		Participants overall (n=109)	Cohort 1 (n=43)	Cohort 2 (n=36)	Cohort 3 (n=30)	*P* value
Gender, n (%)						.57
	Female	58 (53.2)	26 (60.5)	18 (50)	14 (46.7)	
	Male	51 (46.8)	17 (39.5)	18 (50)	16 (53.3)	
	Other	0 (0)	0 (0)	0 (0)	0 (0)	
Age (years), mean (SD)		22.5 (0.7)	22.3 (0.7)	22.6 (0.7)	22.5 (0.7)	.40

Intercohort Comparison of Performance

Scores

The scores of cohorts 1, 2, and 3 after 1 month showed a statistically significant difference (mean 65.8, SD 21.5; mean 81.9, SD 6.9; and mean 85.3, SD 8.6, respectively; *P*<.001). Pairwise comparisons showed that the mean score of cohort 1 was 16.1 points less than the mean score of cohort 2 (95% CI –23.0 to –9.0; *P*=.03) and 19.5 points less than that of cohort 3 (95% CI –26.7 to –12.2; *P*<.001), whereas the difference in scores between cohorts 2 and 3 was not statistically significant (95% CI –7.3 to 0.48; *P*=.29). At testing after 1 year, the scores of cohorts 1, 2, and 3 showed a statistically significant difference (mean 60.3, SD 16.7; mean 75.5, SD 3.9; and mean 90.9, SD 4.3, respectively; *P*<.001). Pairwise comparisons showed that the mean score of cohort 1 was 15.2 points less than that of cohort 2, but this was not significantly different (95% CI –20.5 to –9.9; *P*=.06). Meanwhile, the mean score for cohort 3 was 30.6 points higher than that of cohort 1 (95% CI –36.0 to –25.2; *P*<.001), and the mean score of cohort 3 was 15.4 points higher than that of cohort 2 (95% CI –17.4 to –17.3; *P*<.001); both represented a highly significant difference (Table 3 and Figure 2).

Table . Intercohort comparison of the score and time spent on removal partial denture (RPD) design tests conducted after 1 month and after 1 year; intracohort comparisons of the score and time spent between the 1 month and 1 year time points.

	After 1 month							After 1 year							
	Mean (SD)	Coefficient of variation	Quartile 1 (IQR; range)	*P* value[a]	Difference[b] (95% CI; *P* value[c])	Difference[d] (95% CI; *P* value)[e]	Difference[f] (95% CI; *P* value[g])	Mean (SD)	Coefficient of variation	Quartile 1 (IQR; range)	*P* value[a]	Difference[b] (95% CI; *P* value[c])	Difference[d] (95% CI; *P* value[e])	Difference[f] (95% CI; *P* value[g])	*P* value[h]
RPD design test score				<.001	–16.0 (–23.0 to –9.0; .003)	–19.5 (–26.7 to –12.2; <.001)	–3.4 (–7.3 to 0.48; .29)				<.001	–15.2 (–20.5 to –9.9; .006)	–30.6 (–36.0 to –25.2; <.001)	–17.4 (–17.4 to –17.3; <.001)	
Cohort 1	65.8 (21.5)	0.33	53.0 (31.00; 23-95)					60.3 (16.7)	0.28	50.0 (25.00; 20-95)					.16[i]
Cohort 2	81.9 (6.9)	0.08	80.0 (5.00; 60-95)					75.5 (3.9)	0.05	73.0 (5.00; 86-67)					<.001[i]
Cohort 3	85.3 (8.6)	0.10	80.8 (9.50; 60-98)					90.9 (4.3)		89.0 (5.25; 80-98)					.001[j]
Time spent, s				.06	572.9 (–33.8 to 1179.5; >.99)	567.5 (14.4 to 1120.6; >.99)	44.65 (–576.7 to 666.0; >.99)				<.001	191.8 (–195.7 to 579.3; >.99)	–475.2 (–868.3 to –82.0; <.001)	–667.0 (–951.6 to –382.4; .004)	
Cohort 1	2407.8 (1370.3)	0.57	1088.0 (2132.00; 293-5286)					2049.2 (1099.0)	0.54	1256.0 (1393.00; 364-4623)					.10[i]
Cohort 2	1835.0 (1329.2)	0.72	938.0 (1317.75; 65-5950)					1857.3 (587.4)	0.32	1617.5 (567.25; 593-3224)					.31[j]
Cohort 3	1790.3 (1195.5)	0.67	901.5 (1284.50; 553-4614)					2524.3 (566.4)	0.23	2380.8 (535.25; 837-3276)					.003[j]

[a]Kruskal-Wallis *H* test for differences in the score or time spent among the 3 cohorts.

[b]Cohort 1 vs cohort 2.

[c]Kruskal-Wallis *H* test for differences in the score or time spent between cohort 1 and cohort 2.

[d]Cohort 1 vs cohort 3.

[e]Kruskal-Wallis *H* test for differences in the score or time spent between cohort 1 and cohort 3.

[f]Cohort 2 vs cohort 3.

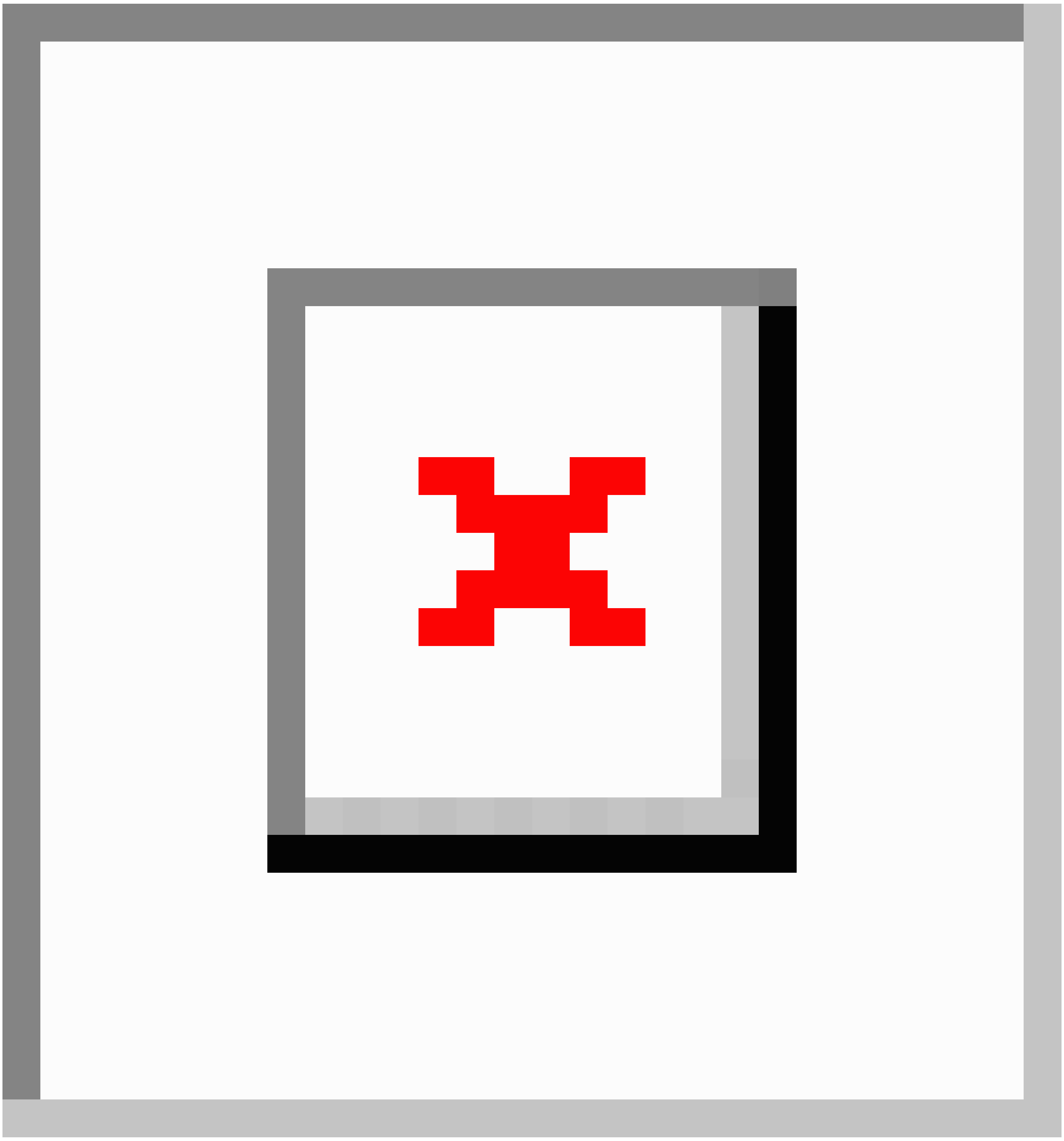
[g]Kruskal-Wallis *H* test for differences in the score or time spent between cohort 2 and cohort 3.

[h]*P* value for differences in both the score and time spent between the 1 month and 1 year time points.

[i]Paired 2-tailed *t* test for differences.

[j]Wilcoxon matched-pairs test for differences.

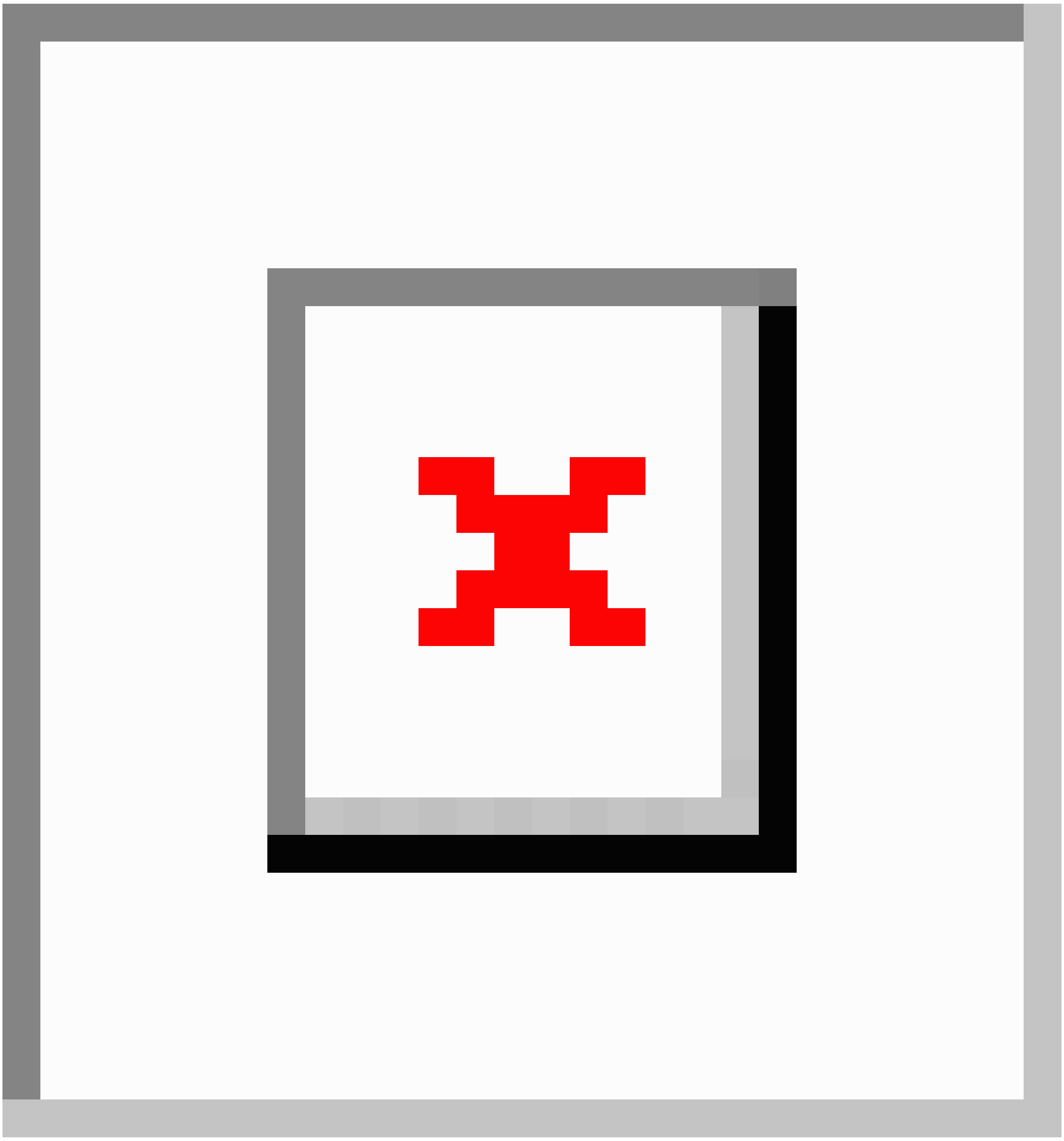
Figure 2. Intercohort comparison and intracohort comparison of the score of the removal partial denture (RPD) design test. **$P \leq .01$, ***$P \leq .001$.



Time Spent

No significant difference was noted in the time spent by the 3 cohorts on the test after 1 month (cohorts 1-3: mean 2407.8, SD 1370.3 s; mean 1835.0, SD 1329.2 s; and mean 1790.3, SD 1195.5 s, respectively; *P*=.06). Pairwise comparisons also did not show any significant differences (cohorts 1-3: 95% CI –33.8 to 1179.5; *P*>.99; 95% CI 14.4-1120.6; *P*>.99; 95% CI –576.7 to 666.0; *P*>.99, respectively). However, the mean time spent on the test after 1 year did show a statistically significant difference between the cohorts (cohorts 1-3: mean 2049.2, SD 1099.0; mean 1857.3, SD 587.4; and mean 2524.3, SD 566.4, respectively; *P*<.001). Pairwise comparisons showed that the mean time spent by cohort 1 was 745.1 seconds shorter than that by cohort 3 (95% CI –868.3 to –82.0; *P*<.001), and the mean time spent by cohort 2 was 667.0 seconds shorter than that by cohort 3 (95% CI –951.6 to –382.4; *P*=.004); both represent a statistically significant difference, while the difference between cohorts 1 and 2 was not statistically significant (95% CI –195.7 to 579.3; *P*>.99) (Table 3 and Figure 3).

Figure 3. Intercohort comparison and intracohort comparison of the time spent on the removal partial denture (RPD) design test. ns: not significant. **P≤.01, ***P≤.001.



Intracohort Comparison of Performance at Different Time Points

Scores

The difference in scores between the tests conducted 1 month and 1 year later for cohort 1 was not statistically significant (95% CI –2.1 to 13.0; *P*=.16). For cohort 2, the mean score obtained on the test conducted 1 month later was 6.4 points higher than that obtained 1 year later (95% CI 3.9-8.8; *P*<.001). For cohort 3, the mean score obtained on the test conducted 1 month later was 5.6 points less than that obtained on the test conducted 1 year later (95% CI –8.7 to –2.5; *P*=.001) (Table 3 and Figure 2).

Time Spent

The time spent by cohorts 1 and 2 on the tests conducted 1 month and 1 year later did not differ significantly (cohort 1: 95% CI –77.5 to 794.9; *P*=.10; cohort 2: 95% CI –372.5 to 327.9; *P*=.31). However, a significant difference was observed for cohort 3, where the time spent on the test conducted 1 month later was 734.0 seconds shorter than that conducted 1 year later (95% CI –1149.9 to –318.0; *P*=.003) (Table 3 and Figure 3).

Discussion

Principal Findings

Historically, the process of learning RPD design is a potentially difficult part of dental education [21]. It requires that dental students first acquire a knowledge base and then use critical thinking skills based on evidence to apply that knowledge to a wide variety of clinical patient care situations. This characteristic suggests that a case-based learning mode is the most appropriate approach for RPD design learning. Case-based learning requires the use of real patient cases and scenarios to reflect realistic patient care situations, and students are asked to draw from their established foundational knowledge to make decisions about problems they may encounter in practice [22]. Previous studies have confirmed the effectiveness of case-based learning in RPD design learning [23]. These studies have typically used text and 2D sketches to describe structured clinical cases, but enhanced digital techniques are gradually being applied to transition from simple presentation documents to computer-aided teaching [11,24-27]. Some studies have further developed decision support systems for RPD design based on clinical case libraries to help trainee dentists complete RPD design by providing cases with similar task requirements [28,29]. More recently, 3D virtual casts and CAD software have been introduced to align with clinical cases and currently prevalent dental laboratory digital workflows [30,31]. Nevertheless, several challenges limit the application of these findings. First, many studies have only addressed students in clinical dental programs, ignoring the dental technology student populace, who, as future dental technicians, are key stakeholders for any RPD design education. Second, in assessing the validity of a training program, most studies have investigated short-term effectiveness without considering the effect on long-term retention of skills, which is the most important for translation to future practice. Our research approach fills these gaps.

When investigating relatively permanent changes in learning, the experimental design needs to incorporate a retention interval, which refers to a period without further practice. Following this interval, assessments can be conducted to evaluate learning outcomes. The inclusion of retention intervals aims to eliminate transient effects resulting from practice, such as fatigue or motivational factors [32]. Existing research lacks discussion on how to determine the length of the retention interval. In this study, the retention interval was determined using a combination of experience, design of relevant literature, and course scheduling. One month after the end of training is the latest time the participants can schedule a test before entering the semester vacation. One year after the end of training is the latest time the participants can schedule a test before graduation. Both time nodes are supported by relevant literature studies [33-35]. Within 1 year after completing the training, participants participate in a uniform dental laboratory internship, and the internship outline has uniform requirements for the design of RPDs with the same workload. At the same time, the user accounts of the participants were blocked in the RPD module, preventing the participants from using the module for additional training. However, participants may use paper and pencil for additional practice since they have different expectations for work content after graduation. Therefore, confounding factors related to different amounts of practice are inevitable.

We noted that the scores obtained on the test after 1 month for cohort 1 was significantly lower than the scores for cohorts 2 and 3, who received the D-RPD intervention. This finding reflects the higher efficacy of the D-RPD digital training approach compared to traditional training at improving short-term performance in RPD design. In addition, cohort 2 scored less than cohort 3, which was provided with the 3D virtual cast–based progressive intervention, albeit with no statistically significant difference, which is consistent with the results of Mahrous et al [30]. This finding suggests no significant short-term benefits of progressive digital training incorporating 3D virtual simulation over digital training using 2D sketches and text alone. However, the scores obtained in the tests conducted 1 year later showed that cohort 3 displayed significantly improved performance in comparison with the other cohorts, thus demonstrating improved long-term outcomes of the progressive digital training approach. Of note, added tacit knowledge from clinical practice gained during the internship curriculum that commenced soon after the first test, where students had additional opportunities to learn and participate in the process of RPD design, could have contributed to such an effect. Such practice enriches the experiential learning of students by allowing for case-based learning and greater practice [36]. The D-RPD module with the 3D virtual simulation–based intervention for cohort 3 was aligned with routine clinical production models to a large extent, which possibly facilitated higher competency over a period of time in cohort 3. These findings are in contrast to our short-term observations and those of Mahrous et al [30]. For the “time spent” evaluation dimension, the differences between the 3 cohorts were not significant at the 1-month test. The complexity of the RPD design process itself could account for this finding. After 1 year, a significant difference was notable, and cohort 3 showed the longest mean time spent on the test. It is feasible that the participants in cohort 3 took more factors into account in the RPD design after undertaking clinical practice in the intervening period and that the D-RPD process with 3D virtual casts was the most consistent with clinical practice; therefore, this effect was produced over a longer period of time. The longer time taken by cohort 3 in estimating more factors and spending more time could also have contributed to their higher score over time. These results indicate that the use of D-RPD, especially when incorporating the use of 3D virtual casts in a progressive mode, may facilitate an improvement in the RPD design competency of students compared with the traditional RPD design training approach.

The mean scores of cohorts 1 and 2 were less after 1 year compared to the scores after 1 month, showing a certain degree of loss of competency over time. In contrast, a significant increase was noted in the mean score of cohort 3. Before entering clinical practice, the participants in this cohort were exposed to experiential learning through virtual simulation that had similarities with clinical work, which might have produced a synergistic effect on improving RPD design competency. It is especially noteworthy that the scores and time spent at the 2 time points by cohort 1 showed very large SDs, indicating high variability in RPD design competency among the cohort 1 students. The opposite was notable in cohorts 2 and 3, which may be related to better teacher supervision, which D-RPD can facilitate. Moreover, previous research has shown that D-RPD design training has advantages that can be partly attributed to improved tracking of students' learning progress and their timely interactions with trainers [37,38]. D-RPD allows teachers to check the progress of the RPD design tasks of the students, make efficient corrections, and provide more frequent feedback. It is evident that this digital teaching mode can facilitate greater student engagement and problem-based learning compared to traditional paper-based teaching. These findings also reinforce that the approach involving D-RPD design combined with 3D virtual casts can provide students with more effective teacher supervision, while offering them virtual experiential learning consistent with clinical activity.

The intervention mode for cohort 3 was similar to the clinical CAD/CAM digital denture design process, which can improve the quality and efficiency of prosthesis design and facilitate improved management of design schemes [39,40]. Intraoral scanning produces 3D virtual casts that can improve precision, and it is readily accepted by the patient compared to the traditional impression method, producing models with higher accuracy [41,42]. The digital workflow allows dental technicians to design directly on these models and to perform postprocessing of multiple scanning data [14,43], thus rendering the entire workflow efficient and convenient. However, despite the rapidly increasing adoption of digital workflows in dental practices worldwide, preclinical education in dentistry and dental technology is typically lagging at imparting the relevant skills to students [12,44,45]. Taken together with our earlier research, this work proposes and validates a progressive digital teaching module for RPD design training that incorporates 3D virtual simulation, demonstrates greater efficacy for a digital training approach compared to traditional training, and provides evidence that a virtual 3D simulation–based progressive digital training module can enhance long-term learning outcomes of RPD design training.

Limitations and Future Work

The limitations of this study include a small sample size, a single center for recruitment, and a lack of randomization, which may have led to unaccounted differences in the inherent learning ability of students and their existing competency prior to participation in the experiment. In future work, bias may be avoided by using a randomized controlled study design to provide stronger evidence for this training module. In addition, there is a lack of data regarding the effectiveness of this training module for clinical dentistry students. Further studies are merited to enable more widespread adoption of 3D virtual simulation–based digital training approaches in dental education.

Conclusions

In this cohort study, we report in detail a major update to the D-RPD module and the design of an intervention experiment to observe the effects of traditional training, D-RPD training, and additional 3D virtual simulation–based digital training on the RPD design competency of students. Based on the results, we propose an effective, progressive, digital 3D virtual simulation workflow–based training module for RPD design, and we have preliminarily verified the efficacy of this novel training approach for facilitating improvement and long-term retention of RPD design competency among dental technology students. This training module should be further extended to clinical dentistry students, randomized controlled experiments should be designed, and feedback from students and teachers should be collected to enable its further optimization and eventual inclusion in curricula.

Acknowledgments

KXL and YQX are co-first authors of this work, and XHZ and MWP are co-corresponding authors (XHZ can be reached at haoxiaohan_buaa@163.com for correspondence).

The engineers' group from Beijing Unidraw Virtual Reality Technology Research Institute Co. Ltd., led by XZ, helped develop the Objective Manipulative Skill Examination of Dental Technicians system, including the digital removable partial denture module. The Zhaozhi student team assisted in the early design and pre-experimental data collection for the digital removable partial denture module. This research was funded by the CQMU Program for Youth Innovation in Future Medicine (W0002 for LJ), and the Chongqing Municipal Teaching Reformation Fund (193070 for XF).

Authors' Contributions

MP designed the structure of the digital removable partial denture module; organized the cohort study; collected, analyzed and interpreted the data; obtained funding; and wrote and revised the manuscript for important intellectual content; KL drafted the manuscript and helped analyze and visualize the data; XZ helped complete the digital removable partial denture module algorithm and revised parts of the manuscript; PJ and XF helped obtain funding; CM, NY, FT, YL, and YB performed as the teaching expert group; YX and XF performed as the scoring expert group; JW, DF, and HY helped with the program development and provided technical support; MC recorded the multimedia appendices; and HC, LJ, and MC provided administrative and material support; JS, PJ, LJ, and MP supervised the entire process. MP and XZ had full access to all of the data in this study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Development of a progressive D-RPD module incorporating case-based virtual 3D simulation.
[MP4 File, 4446 KB - games_v12i1e46789_app1.mp4]

Multimedia Appendix 2
Display of the interventions of the three cohorts.
[MP4 File, 7529 KB - games_v12i1e46789_app2.mp4]

References

1. Kim JJ. Revisiting the removable partial denture. Dent Clin North Am 2019 Apr;63(2):263-278. [doi: 10.1016/j.cden.2018.11.007] [Medline: 30825990]
2. Choong EKM, Shu X, Leung KCM, Lo ECM. Oral health-related quality of life (OHRQoL) after rehabilitation with removable partial dentures (RPDs): A systematic review and meta-analysis. J Dent 2022 Dec;127:104351. [doi: 10.1016/j.jdent.2022.104351] [Medline: 36280004]
3. Bhathal M, Batra J, Attresh G, Sambyal S. A review on stresses-induced by removable partial dentures. Int J Contemp Dent Med Rev 2015:050115. [doi: 10.15713/ins.ijcdmr.56]
4. Becker CM, Kaiser DA, Goldfogel MH. Evolution of removable partial denture design. J Prosthodont 1994 Sep;3(3):158-166. [doi: 10.1111/j.1532-849x.1994.tb00147.x] [Medline: 7874258]
5. Zlatarić DK, Celebić A, Valentić-Peruzović M. The effect of removable partial dentures on periodontal health of abutment and non-abutment teeth. J Periodontol 2002 Feb;73(2):137-144. [doi: 10.1902/jop.2002.73.2.137] [Medline: 11895277]
6. Graser GN. Review of the literature: predoctoral removable prosthodontics education. J Prosthet Dent 1990;64(3):326-333. [doi: 10.1016/0022-3913(90)90016-6] [Medline: 2213617]
7. Christensen GJ. Removable prosthodontics: a forgotten part of dentistry. Alpha Omegan 2006 Mar;99(1):26-28. [doi: 10.1016/j.aodf.2006.02.028] [Medline: 16865850]
8. Lynch CD, Allen PF. The teaching of removable partial dentures in Ireland and the United Kingdom. Br Dent J 2007 Oct;203(8):E17. [doi: 10.1038/bdj.2007.581] [Medline: 17599079]
9. Dikbas I, Ozkurt Z, Kazazoglu E. Predoctoral prosthodontic curricula on removable partial dentures: survey of Turkish dental schools. J Dent Educ 2013 Jan;77(1):85-92. [doi: 10.1002/j.0022-0337.2013.77.1.tb05448.x] [Medline: 23314471]
10. Ali SA, Khalifa N, Alhajj MN. Communication between dentists and dental technicians during the fabrication of removable partial dentures in Khartoum State, Sudan. Acta Stomatol Croat 2018 Sep;52(3):246-253. [doi: 10.15644/asc52/3/8] [Medline: 30510300]
11. Mahrous A, El-Kerdani T. Teaching the design and fabrication of RPD frameworks with a digital workflow: a preclinical dental exercise. MedEdPORTAL 2020 Oct 30;16:11041. [doi: 10.15766/mep_2374-8265.11041] [Medline: 33150206]
12. Garcia-Hammaker SS, Snyder MD, Mendonça G, Saglik B. A novel approach to clinical simulation in removable partial denture treatment. J Dent Educ 2020 Jun 26. [doi: 10.1002/jdd.12297] [Medline: 32589768]
13. Barraclough O, Gray D, Ali Z, Nattress B. Modern partial dentures - part 1: novel manufacturing techniques. Br Dent J 2021 May;230(10):651-657. [doi: 10.1038/s41415-021-3070-4] [Medline: 34050296]
14. Pang M, Zhao X, Lu D, et al. Preliminary user evaluation of a new dental technology virtual simulation system: development and validation study. JMIR Serious Games 2022 Sep 12;10(3):e36079. [doi: 10.2196/36079] [Medline: 36094803]
15. Rolskov Bojsen S, Räder SBEW, Holst AG, et al. The acquisition and retention of ECG interpretation skills after a standardized web-based ECG tutorial–a randomised study. BMC Med Educ 2015 Mar 7;15(1):36. [doi: 10.1186/s12909-015-0319-0] [Medline: 25889642]
16. Carr AB, Brown DT. McCracken’s Removable Partial Prosthodontics: Elsevier Health Sciences; 2010.
17. Agresti A. Statistical Methods for the Social Sciences: Pearson; 2018.
18. Razali NM, Wah YB. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J Stat Model Analyt 2011;2(1):21-33 [FREE Full text]
19. MacFarland TW, Yates JM. Kruskal–Wallis H-test for oneway analysis of variance (ANOVA) by ranks. In: Introduction to Nonparametric Statistics for the Biological Sciences Using R: Springer International Publishing; 2016:177-211. [doi: 10.1007/978-3-319-30634-6]
20. Curtin F, Schulz P. Multiple correlations and Bonferroni’s correction. Biol Psychiatry 1998 Oct 15;44(8):775-777. [doi: 10.1016/s0006-3223(98)00043-2] [Medline: 9798082]
21. Basker RM, Harrison A, Davenport JC, Marshall JL. Partial denture design in general dental practice--10 years on. Br Dent J 1988 Oct 8;165(7):245-249. [doi: 10.1038/sj.bdj.4806577] [Medline: 3052549]
22. Nadershahi NA, Bender DJ, Beck L, Lyon C, Blaseio A. An overview of case–based and problem–based learning methodologies for dental education. J Dent Educ 2013 Oct;77(10):1300-1305. [doi: 10.1002/j.0022-0337.2013.77.10.tb05603.x] [Medline: 24098033]
23. Samuelson DB, Divaris K, De Kok IJ. Benefits of case-based versus traditional lecture-based instruction in a preclinical removable prosthodontics course. J Dent Educ 2017 Apr;81(4):387-394. [doi: 10.21815/JDE.016.005] [Medline: 28365602]

24. Lechner SK, Thomas GA, Bradshaw M. An interactive multimedia solution to learning removable partial denture design. J Prosthodont 1998 Sep;7(3):177-182. [doi: 10.1111/j.1532-849x.1998.tb00201.x] [Medline: 9807101]
25. Gibbard LL, Salajan F. A novel interactive online module in a traditional curriculum through a blended learning approach. Electron J e-Learn 2009;7(3):301-308 [FREE Full text]
26. Abdulhadi Al-Samawi L, Mohammed HA, Alsamawi A, Mahmood HL. A new digital system to generate and draw framework design of removable partial dentures. Presented at: ICERI2013, The 6th International Conference of Education, Research and Innovation; Nov 18-20, 2013; Seville, Spain.
27. Abdulhadi L, Mahmoud A, Al-Samawi H. Fully automated interactive digital system to design metal framework of removable partial dentures: system specifications and applications in teaching and patient management. In: EDULEARN20 Proceedings: 12th International Conference on Education and New Learning Technologies: International Academy of Technology, Education and Development; 2020:58-66. [doi: 10.21125/edulearn.2020.0037]
28. Chen Q, Wu J, Li S, Lyu P, Wang Y, Li M. An ontology-driven, case-based clinical decision support model for removable partial denture design. Sci Rep 2016 Jun 14;6:27855. [doi: 10.1038/srep27855] [Medline: 27297679]
29. Davenport JC, Hammond P, Fitzpatrick FJ. Computerized design of removable partial dentures: a knowledge-based system for the future. Dent Update 1993;20:221-226. [Medline: 8299844]
30. Mahrous A, Schneider GB, Holloway JA, Dawson DV. Enhancing student learning in removable partial denture design by using virtual three-dimensional models versus traditional two-dimensional drawings: a comparative study. J Prosthodont 2019 Oct;28(8):927-933. [doi: 10.1111/jopr.13099] [Medline: 31343801]
31. Mahrous A, Schneider GB. Enhancing student learning of removable prosthodontics using the latest advancements in virtual 3D modeling. J Prosthodont 2019 Apr;28(4):28-472. [doi: 10.1111/jopr.13044] [Medline: 30838728]
32. Sanli EA, Carnahan H. Long-term retention of skills in multi-day training contexts: a review of the literature. Int J Ind Ergon 2018 Jul;66:10-17. [doi: 10.1016/j.ergon.2018.02.001]
33. Kowalski C, Boulesteix AL, Harendza S. Effective methods to enhance medical students' cardioversion and transcutaneous cardiac pacing skills retention - a prospective controlled study. BMC Med Educ 2022 Jun 1;22(1):417. [doi: 10.1186/s12909-022-03495-4] [Medline: 35650577]
34. Lynch TG, Steele DJ, Johnson Palensky JE, Lacy NL, Duffy SW. Learning preferences, computer attitudes, and test performance with computer-aided instruction. Am J Surg 2001 Apr;181(4):368-371. [doi: 10.1016/s0002-9610(01)00575-x] [Medline: 11438276]
35. Hubert V, Duwat A, Deransy R, Mahjoub Y, Dupont H. Effect of simulation training on compliance with difficult airway management algorithms, technical ability, and skills retention for emergency cricothyrotomy. Anesthesiology 2014 Apr;120(4):999-1008. [doi: 10.1097/ALN.0000000000000138] [Medline: 24434303]
36. Eraut M. Non-formal learning and tacit knowledge in professional work. Brit J of Edu Psychol 2000 Mar;70(1):113-136. [doi: 10.1348/000709900158001] [Medline: 10765570]
37. Abdulhadi L, Mohammed H, Al-Samawi H. Assessment of the weak and powerful learning outcome of removable partial dentures using automated software as a universal benchmark tool. Presented at: 12th International Conference on Education and New Learning Technologies; Jul 6-7, 2020; Online. [doi: 10.21125/edulearn.2020.0036]
38. Abdulhadi Al-Samawi L, Mohammed H. Multiple teaching methods to enhance removable partial denture learning outcome and designing skill for undergraduate dental students. A personal experience. In: INTED2014 Proceedings: 8th International Technology, Education and Development Conference: International Academy of Technology, Education and Development; 2014:5881-5887 URL: https://library.iated.org/view/abdulhadi2014mul [accessed 2024-03-25]
39. Miyazaki T, Hotta Y, Kunii J, Kuriyama S, Tamaki Y. A review of dental CAD/CAM: current status and future perspectives from 20 years of experience. Dent Mater J 2009 Jan;28(1):44-56. [doi: 10.4012/dmj.28.44] [Medline: 19280967]
40. Arnold C, Hey J, Schweyen R, Setz JM. Accuracy of CAD-CAM-fabricated removable partial dentures. J Prosthet Dent 2018 Apr;119(4):586-592. [doi: 10.1016/j.prosdent.2017.04.017] [Medline: 28709674]
41. Siqueira R, Galli M, Chen Z, et al. Intraoral scanning reduces procedure time and improves patient comfort in fixed prosthodontics and implant dentistry: a systematic review. Clin Oral Investig 2021 Dec;25(12):6517-6531. [doi: 10.1007/s00784-021-04157-3] [Medline: 34568955]
42. Burzynski JA, Firestone AR, Beck FM, Fields HW, Deguchi T. Comparison of digital intraoral scanners and alginate impressions: time and patient satisfaction. Am J Orthod Dentofacial Orthop 2018 Apr;153(4):534-541. [doi: 10.1016/j.ajodo.2017.08.017] [Medline: 29602345]
43. Nilsson TA, Hedman LR, Ahlqvist JB. Dental student skill retention eight months after simulator-supported training in oral radiology. J Dent Educ 2011;75:679-684. [doi: 10.1002/j.0022-0337.2011.75.5.tb05094.x] [Medline: 21546602]
44. Machoy ME, Szyszka-Sommerfeld L, Vegh A, Gedrange T, Woźniak K. The ways of using machine learning in dentistry. Adv Clin Exp Med 2020 Mar;29(3):375-384. [doi: 10.17219/acem/115083] [Medline: 32207586]
45. Rodrigues P, Esteves A, Botelho J, et al. Usability, acceptance, and educational usefulness study of a new haptic operative dentistry virtual reality simulator. Comput Methods Programs Biomed 2022 Jun;221:106831. [doi: 10.1016/j.cmpb.2022.106831] [Medline: 35544961]

Abbreviations

CAD: computer-aided design
CAM: computer-aided manufacturing
D-RPD: digital removable partial denture
OMEDT: Objective Manipulative Skill Examination of Dental Technicians
RPD: removable partial denture
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A McDaniel; submitted 02.03.23; peer-reviewed by Z Zhou, WJ Chang; revised version received 29.01.24; accepted 21.02.24; published 05.04.24.

Please cite as:
Liu K, Xu Y, Ma C, Yu N, Tan F, Li Y, Bai Y, Fu X, Wan J, Fan D, Yin H, Chen M, Chen H, Jiang L, Song J, Ji P, Zhao X, Pang M
Efficacy of a Virtual 3D Simulation–Based Digital Training Module for Building Dental Technology Students' Long-Term Competency in Removable Partial Denture Design: Prospective Cohort Study
JMIR Serious Games 2024;12:e46789
URL: https://games.jmir.org/2024/1/e46789
doi:10.2196/46789

Original Paper

A Serious Game (“Fight With Virus”) for Preventing COVID-19 Health Rumors: Development and Experimental Study

Shuo Xiong[1], PhD; Long Zuo[2], PhD; Qiwei Chen[3], MA; Zhang Zeliang[4], PhD; Mohd Nor Akmal Khalid[4,5], PhD

[1]Philosophy and Social Sciences Laboratory of Big Data and National Communication Strategy, Huazhong University of Science and Technology, Wuhan, China

[2]School of Information Engineering, Chang’an University, Xi'an, China

[3]School of Journalism and Information Communication, Huazhong University of Science and Technology, Wuhan, China

[4]School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan

[5]School of Computer Science, Universiti Sains Malaysia, Georgetown, Malaysia

Corresponding Author:
Shuo Xiong, PhD
Philosophy and Social Sciences Laboratory of Big Data and National Communication Strategy
Huazhong University of Science and Technology
Building 6, No 1037, Luoyu Road
Wuhan, 430074
China
Phone: 86 15927188806
Email: xiongshuo@hust.edu.cn

Abstract

Background: Health rumors arbitrarily spread in mainstream social media on the internet. Health rumors emerged in China during the outbreak of COVID-19 in early 2020. Many midelders/elders (age over 40 years) who lived in Wuhan believed these rumors.

Objective: This study focused on designing a serious game as an experimental program to prevent and control health rumors. The focus of the study was explicitly on the context of the social networking service for midelders/elders.

Methods: This research involved 2 major parts: adopting the Transmission Control Protocol model for games and then, based on the model, designing a game named “Fight With Virus” as an experimental platform and developing a cognitive questionnaire with a 5-point Likert scale. The relevant variables for this experimental study were defined, and 10 hypotheses were proposed and tested with an empirical study. In total, 200 participants were selected for the experiments. By collecting relevant data in the experiments, we conducted statistical observations and comparative analysis to test whether the experimental hypotheses could be proved.

Results: We noted that compared to traditional media, serious games are more capable of inspiring interest in research participants toward their understanding of the knowledge and learning of health commonsense. In judging and recognizing the COVID-19 health rumor, the test group that used game education had a stronger ability regarding identification of the rumor and a higher accuracy rate of identification. Results showed that the more educated midelders/elders are, the more effective they are at using serious games.

Conclusions: Compared to traditional media, serious games can effectively improve midelders’/elders’ cognitive abilities while they face a health rumor. The gameplay effect is related to the individual’s age and educational background, while income and gender have no impact.

(JMIR Serious Games 2024;12:e45546) doi:10.2196/45546

KEYWORDS

serious game; COVID-19; health rumor; game communication; game TCP model; Transmission Control Protocol; gaming; misinformation; disinformation; rumor; health communication; false information; elder; older adult

Introduction

Background

In recent years, the arbitrary spreading of health rumors in mainstream social media on the internet has increasingly gained the attention of the public and raised concerns [1]. For health rumor researchers, a common concern is to propose a feasible and effective prevention and control program for current rampant rumors [2,3]. Furthermore, to prevent and control the spread of such rumors, it is necessary to strengthen the public's health knowledge to judge and identify the rumors [4]. The concept of a rumor involves a form of statement whose veracity cannot be quickly or ever confirmed. Generally, we have a "dream rumor" and a "bogie rumor," the former reflecting public desires and wished-for outcomes and the latter hiding some special purpose by somebody, both of them largely occurring during the early COVID-19 epidemic in Wuhan, China [4]. The traditional way of countering rumors often relies on media refutation, which can only be described as a "Band-Aid" solution. To fundamentally prevent and control COVID-19–related rumors and enhance the public's ability to resist them, we need to find a form of "information vaccine." Therefore, we chose serious games as a "vaccine" in this context.

The purpose of a serious game is to help people acquire knowledge by playing games. Serious games involve solving problems and studying via careful and thoughtful game ideas [5], while considering characteristics beyond gameplay (eg, purpose and scope [6]). In addition, game elements are used to improve information processing and identify relevant information, which is consistent with the purpose of health rumor prevention research [7]. COVID-19–related rumors are based on the content of serious games to experiment with health rumor prevention, mainly using the control variable method and the analysis-contrast method to apply serious game learning to health rumor prevention research. This paper explores how to help people acquire knowledge of health rumors and health commonsense from the prevention experiment using the relaxed approach of serious games [8].

Previously, serious games have provided a platform for education and business use. For instance, behavioral interventions can be carefully tested and designed to reduce risk-taking behaviors [9], where transmission risks and the usefulness of pandemic-like simulations were demonstrated in the laboratory to be safely and ethically comprehended at the initial state of a health crisis. In addition, other studies prove that serious games are used to accommodate informational and communication complexities in early warning disaster management to simulate and test how public information from social media is used in emergency operation centers to make (protective and communicative) decisions based on levels of trust, usefulness, and completeness [10,11]. Therefore, serious games as an "information vaccine" have certain feasibility, and this paper also explored this issue. Nevertheless, the prevention and control of health rumors have rarely been considered in the context of the social networking service (SNS) for elderly users.

Serious Games

Why are serious games chosen as a solution? Serious games refer to those electronic games whose main content is used for knowledge and skill development, professional training, and spreading culture. They are widely used in many fields. Compared to the limitations and congenital deficiencies of some communication models of traditional media, serious games have become an effective tool to address many social problems, because of their fast speed, wide range, and interactivity [12].

Abt [5] first defined the concept of serious games as follows: "These games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement." Later, Sawyer [8], in his white paper titled "Serious Game 2 Initiative," redefined the concept of a serious game as being an entertainment game with nonentertainment goals. Several variants of the concept have also been proposed. Michael and Chen [13] defined serious games as games that educate, train, and inform. Meanwhile, Zyda [14] defined serious games as a mental contest played with a computer following specific rules. This situation led some analysts to describe serious games as the next wave of technology-mediated learning [15]. Although there is no single definition of the serious game concept, all the proposed definitions convey the same idea: using games to teach or transmit something [16].

Serious games are present in many areas. Westera et al [17] argued that serious games open up many new opportunities for learning complex skills, especially in the education and training domains [12,18-20]. Moreover, Yusoff et al [21] and Crookall [22] argued that good computer games are an excellent example of modern educational theory and that establishing simulation-based serious games as a discipline is a crucial endeavor that could benefit many other related disciplines.

Some early studies were systematically outlined by Connolly et al [23]. For instance, Ziebarth et al [18] and Diehl et al [19] adopted serious games to develop a prototype for the training and education of health students. Some scholars have emphasized the role of serious games in highly specialized skill acquisition (ie, drilling operation [24], mitigation of student dropout [25], improving the command performance of pilots [26]) and education (ie, medical surgery) [27,28], while providing the means to influence cognition and motivational driver [29].

Serious games are also being applied to pass on knowledge or expertise, which can be adopted for various purposes (ie, rehabilitation, psychotherapy, and brain disorders [30-32]). Sometimes, a serious game can also be used to increase risk awareness in the working area of the manufacturing floor [33]. The review by Abd-Alrazaq et al [34] showed that tools such as serious games are usable but are not replaceable options for rehabilitation and clinical intervention where long-term effects are required. Another review by Krath et al [33] revealed that serious games have also incorporated many theoretical foundations relevant to 3 significant landscapes: behavior, learning, and affect-motivation.

In research related to midelders/elders, several studies have demonstrated the potential of serious games to promote physical activity among older adults. For example, a randomized controlled trial conducted by Fu et al [35] found that a 6-month program of exergaming (exercise using video games) significantly improves the physical function of older adults. Similarly, a study by Jiménez-Pavón et al [36] showed that exergaming increases physical activity and cognitive performance in older adults with mild cognitive impairment.

Serious games have also been used to enhance cognitive training and disease management among older adults. For instance, a study by Anguera et al [37] found that cognitive training through a video game improves cognitive control in older adults. Additionally, a systematic review by Loerzel et al [38] indicated that serious games have the potential to improve self-management and quality of life among older adults with chronic diseases.

In COVID-19–related research, several studies have investigated the potential of gamification and serious games in promoting physical activity during the COVID-19 pandemic. For example, a study by Hall et al [39] proposed a project at a hospital's senior health center in Canada to discuss how health care can be addressed using serious games among middle-aged and older adults during the pandemic [39]. The study found that the game was effective in increasing physical activity levels and improving self-efficacy. Lau et al [40] demonstrated the potential use of serious game to improve physical activity, cognitive training, and mental health among the aging population during COVID-19 in Hongkong.

Similarly, a study by Suppan et al [41] developed a serious game designed to promote safe behaviors for infection prevention and control (IPC), with a specific focus on COVID-19 among health care workers (HCWs) and other hospital employees. Another study by Ferreira et al [42] explored the potential of gamification in promoting hand hygiene among HCWs during the pandemic. The study found that the game was effective in increasing hand hygiene compliance among the participants [42].

Overall, gamification and serious games have emerged as a promising tool to promote physical activity and health and well-being during the COVID-19 pandemic. These technologies have the potential to support health promotion initiatives and encourage people to adopt healthy behaviors in a fun and engaging way. Therefore, we believe serious games can also solve the issue of COVID-19–related rumors that existed among Chinese midelders/elders.

Health Rumor Analysis

Zhang et al [1] investigated all 453 features of health rumor data collected from a definitive online reference in China. A logistic regression model was adopted to determine the contribution of such features to true and false health rumors. There were measurable differences between true and false health rumors, where the length of a headline or statement and the presence of pictures were negatively correlated with the probability that a rumor was true. Meanwhile, a rumor was more likely to be true if it contained elements such as numbers, source cues, and hyperlinks. They also found that the dread health rumor is more likely to be true than a wishful one. Meanwhile, Chua and Banerjee [4] conducted a study on health rumors from 2015 to 2018. Users' trust in online health rumors was investigated using 2 factors: length and presence of an image. Additionally, 2 types of rumors were studied: pipe-dream rumor, which offers hope, and bogie rumors, which instill fear. A total of 102 people participated in the experiment, where the finding suggested that pipe-dream rumors are trusted when they are short and do not contain images, while bogie rumors are trusted when they are long and contain images.

Subsequently, Chua and Banerjee [3] investigated the role of epistemic belief in affecting internet users' decision to share online health rumors. The study focused on the characteristics of rumors—true or false, textual or pictorial, dread or wishful—shaping the decision-making among epistemologically naive and robust users separately. The study showed that epistemologically naive individuals are likelier to share online health rumors than epistemologically robust individuals. In addition, epistemologically robust participants were more likely to share textual rumors than pictorial ones. However, there were no differences between true and false rumors (or between dread and wishful rumors) among either epistemologically naive or robust participants. Meanwhile, Wu [43] modeled factors that predicted fake news sharing during the COVID-19 health crisis. Results showed that informational dependency and social dependency engender both positive and negative cognitive states, namely perceived information timeliness, perceived socialization, and social overload, which then invoke positive and negative affects. Considering that SNS dependency affects information-seeking behavior, it is important for individuals to be exposed to as much accurate information as possible and to build up rational communication against the spread of false rumors.

Ji et al [44] explored factors that influence people's engagement in scientific rumormongering of genetically modified (GM) food on the Chinese social media platform Sina Weibo at both the group and the individual level. In total, 9070 posts about GM food were obtained from 1 million users. Analysis using logistic regression of the effect of peer influence did not find that users would depend on their friendship network to spread rumors. Instead, results revealed that people with negative attitudes toward GM food and who are social media extroverts (ie, celebrities) are more likely to spread rumors. In contrast, social reputation did not influence the spread of rumors, overwhelming the voices of the scientific community and negatively influencing public attitudes and behaviors.

Meanwhile, Hui et al [45] conducted a study on the spread mechanism of rumors on social network platforms during COVID-19 and considered education as a control measure against the spread of rumors. A novel epidemic-like model was established to characterize the spread of rumors based on 2 dimensions of users (age and time), susceptibility based on education classes, control strategies to effectively restrain rumor propagation, and numerical simulations to verify the main theoretical results. The study concluded that improving education levels and conducting short-term online education are essential strategies for effectively controlling rumor spread. In addition, Pulido et al [46] focused on the social impact of research to identify types of false health information shared on social media (Reddit, Facebook, and Twitter) using the application of social impact in social media (SISM) methodology. The results indicated that messages focusing on fake health information are primarily aggressive, while those based on the evidence of social impact are respectful and transformative, and deliberation contexts promoted on social media overcome false health information. The findings provide insights into how public health initiatives can support the presence and interactions of evidence as an effective strategy to combat fake news.

A study by Kim and Kim [47] investigated the misinformation belief produced in the context of COVID-19 via 2 main factors: risk perception (psychometric paradigm) and communication. It was found that perceived risk and stigma positively impact belief in fake news, while source credibility and the quantity of information reduce it. Meanwhile, among communication factors, source credibility and the quantity of information reduce belief in fake news, while the credibility of information sources increases it. In addition, Zhao et al [48] used features of online health misinformation that were classified into central level (including topic features) and peripheral level (including linguistic features, sentiment features, and user behavioral features) to propose a health misinformation detection model using the elaboration likelihood model (ELM). Based on a data set collected from a real online health community (because of the lack of a labeled data set), the model correctly detected about 85% of health misinformation. Furthermore, the findings demonstrated the efficacy of behavioral features in health misinformation detection and offered suggestions for misinformation detection by integrating the features of messages and message creators. In COVID-19–related fake news research, Wang and Huang [49] found that although an official denial can initially reduce citizens' belief in unconfirmed information, later when the denial is revealed to be false, the citizens will have lower levels of belief, not just in the current denial, but also in the government's future denials of similar rumors. Moreover, the negative lasting effects will carry over to satisfaction with the authorities in the related policy area.

COVID-19 Background

This paper was initially written in 2020, and the experiment was conducted in the period from February to March 2020. Therefore, many things changed from then up to the Omicron strain of COVID-19. As such, we acknowledge that this paper has time constraints; however, the research still provides some valuable inspiration and conclusions on game studies, media development, and health care. Since the COVID-19 pandemic broke out in December 2019, the related health rumors also began to wreak havoc on the internet.

Rumor prevention is difficult in the case of rumors that rely on propaganda, and the educational means of traditional media are ineffective due to the lack of interaction and the complexity of information. On the internet, especially the midelders/elders were in a state of panic and information-blind obedience [50].

In China, an SNS group existed, in addition to many WeChat groups, similar to Discord and Facebook in the West. Therefore, we could easily find a target sample for our newly established experimental community, where any questions could be communicated at any time. Our experiment was conducted in early 2020, and some people could answer the questionnaire face to face, while others could not because of the lockdown. Therefore, some respondents were sent offline paper questionnaires, and we also requested them to fill in the online questionnaire. The Chinese midelders/elders were comfortable playing the game on their cell phones, so they easily believed the health rumor that the information communication channel is too fast. Some of them whom we could not meet face to face were contacted over a video call, and we confirmed their age and other personal information clearly to ensure accuracy in the experiment.

The original survey, questionnaire, and serious game are in Chinese, convenient for our non–English-speaking respondents, and all the concepts in this paper are the translated version. This means we just translated the statistical data and labels; during the experiment, there was no translation, and we followed the same steps for all the scales.

Elderly WeChat Users in China: Original Survey

According to our data collected in the original survey, the contemporary middle-aged and older adults, especially those aged 40-60 years, have a high frequency of use of WeChat; the number of elderly WeChat users with frequent use accounts for 66.09% of the total. According to interviews at different levels of the questionnaire survey process, middle-aged and elderly users of WeChat are aged from 50 to 65 years. They are also familiar with using the WeChat "circle of friends" function and other social media platforms (eg, TikTok). They often record their daily lives and travel through videos and pictures. Generally, this user group is also active in online social group chats, and their frequency of using online social media is no less than that of some young user groups. For example, 48.85% of middle-aged and elderly WeChat users said they occasionally read health information on WeChat, and only 16.67% said they had never received health information forwarded by relatives and friends (Figure 1).

The survey on the acquisition and dissemination of health information by elderly WeChat users was the focus of this study. Most people do not have the habit of reading health information regularly. It can be seen from the data that this depends to some extent on the frequency of obtaining information. People read health information only when it is forwarded to them by relatives and friends or when relevant health public accounts push this information or when it is in the form of characteristic health information news, as shown in Figure 2. Regarding access to health information, 62.07% of the respondents received health information from their WeChat friends. In addition, 83.33% of the respondents had the experience of forwarding health information to their children or parents, and 57.47% of those forwarded health information to their WeChat friends. Most respondents felt that the original intention of forwarding health information was to help others with a positive attitude.

Figure 1. Reading frequency of SNS health information by Chinese midelders/elders. SNS: social networking service.

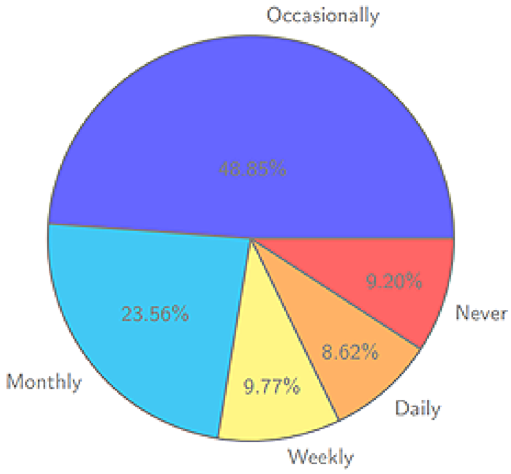
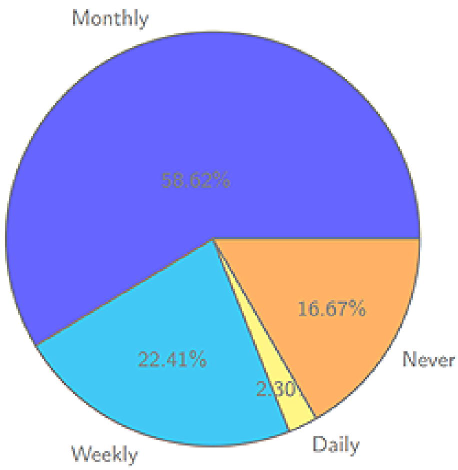


Figure 2. Receiving frequency of SNS health information by Chinese midelders/elders. SNS: social networking service.



However, many midelders/elders received health information without any judgment and recognition and then spread the information with a “good intention” motive, which is also why the health rumor issue is rampant. The data also show that the failure to recognize and identify health rumors is more likely to be the reason than the motive for spreading them. In addition, according to the questionnaire, 86.78% of elderly WeChat users trusted health information forwarded by relatives and friends and 45.4% considered it very trustworthy. The trustworthiness of health information forwarded by colleagues wa 82.18%. These data show that WeChat has become a hotbed for health rumors among the midelders/elders.

Therefore, this paper used a serious game as a tool to test the effect of game media on the prevention of health rumors. Compared with other media, the serious game had a special communication model and effect that could improve this situation (see the *Results* section for more details). Therefore, using the COVID-19 pneumonia rumor was suitable as the target and content of the serious game, involving not only the elderly closely related to COVID-19 pneumonia but also COVID-19 rumor communication relying on WeChat. Finally, the number of health rumors that emerged during the COVID-19 epidemic was enormous, and enough rumor cases could be collected for experimentation.

Methods

The Transmission Control Protocol Model of the Game

There are many theoretical models concerning the communication effect of games as media [51]. The computer networking concept was adopted as the inspiration for this research based on the idea of engineering. Two main protocols exist in network communication: Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) [51]. TCP originated in the initial network implementation, complementing Internet Protocol (IP). TCP provides reliable, ordered, and error-checked delivery of a stream of bytes between applications running on hosts communicating via an IP network. Major internet applications, such as the World Wide Web, email, remote administration, and file transfer, rely on TCP because of the 3-way handshake mechanism (Figure 3) [52].

Having introduced the logical mechanism of UDP and TCP from the technical perspective of communication, we can see that all media communication models are suited to TCP and UDP (2 computer network theories). UDP uses a simple connectionless communication model, just from the information source to the information sink. For example, the newspaper provides information to readers without any interaction (request and response). However, for all current media, only games match the TCP model (Figure 4). In traditional media, no matter the newspaper, broadcaster, or television program, the audience only receives information; the UDP model does not have a feedback process, the timeliness is good, but transmission is unstable. As a result, users can refuse to accept information or hardly notice useful content. Therefore, serious games can help society to address health rumor issues. In this paper, we proved the effect on the health communication area [53].

Figure 3. Three-way handshaking in TCP. ACK: acknowledge; RTT: round-trip time; SYN: synchronize; TCP: Transmission Control Protocol.

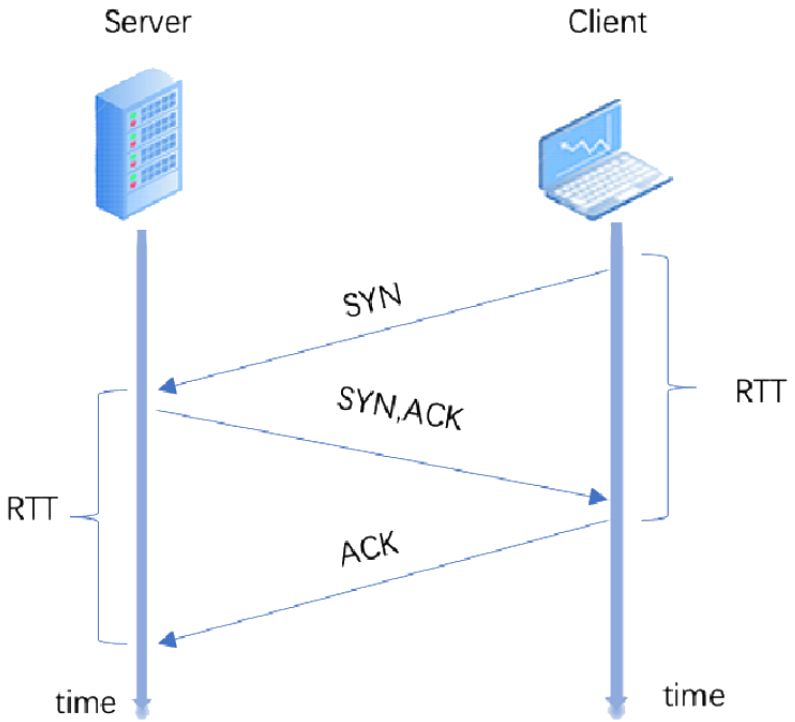
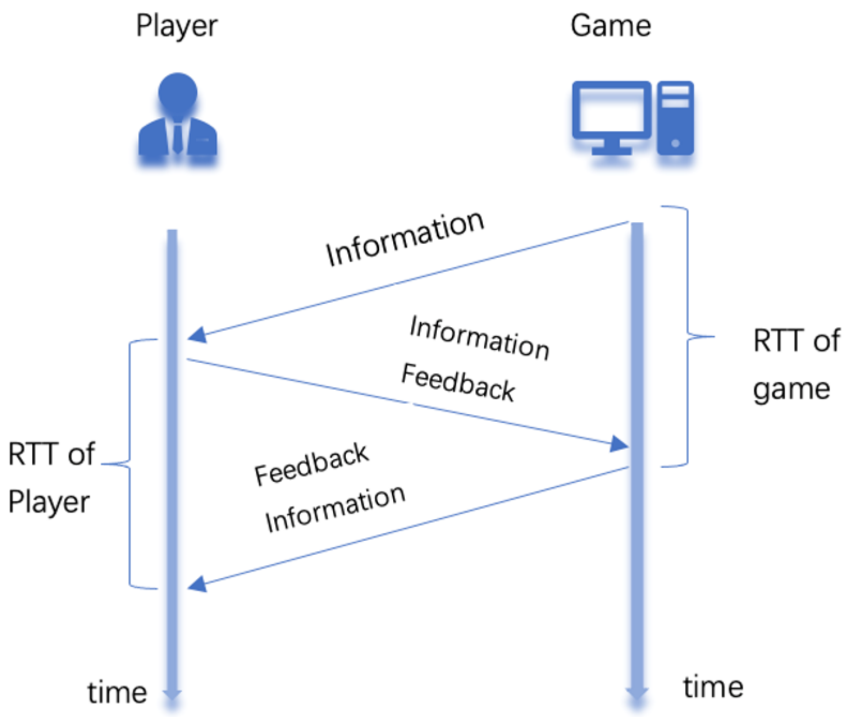


Figure 4. Serious game of TCP. RTT: round-trip time; TCP: Transmission Control Protocol.



Study Design

This study investigated the prevention and control of health rumors in WeChat, as most elderly WeChat users are concerned about health information and are negatively affected by health rumors. Here, the term "elderly" in our paper is a macroscopic definition: it is not only a physiological age classification but also a description of the psychology or state. In China, people who believe a health rumor via the SNS in the age range of 40-60 years (midelders/elders) were considered. We recruited 200 midelders/elders in Tongren City, Guizhou Province, China, which did not have a serious spread of COVID-19 in early 2020. The participants got together for dancing and training in the city plaza, and then, we requested them to attend our game experiment.

The experimental program was constructed in 2 parts. The first part was developing a serious game based on the content of health rumors and health commonsense; we named it "Fight With Virus." The purpose was to apply this in a health rumor prevention experiment. The second part was developing a cognitive questionnaire with the theme of COVID-19 health rumor, with a 5-point Likert scale, which aimed to compare and analyze the prevention effects of traditional and serious game learning models on health rumors. Baishya and Samalia [54] extended the unified theory of acceptance and use of technology (UTAUT) into UTAUT2, incorporating 3 constructs into the original UTAUT: hedonic motivation, price value, and habit. Individual differences (ie, age, gender, and experience) were hypothesized to moderate the effects of these constructs on behavioral intention and technology use, thus affecting their learning of new technologies. Therefore, according to several past studies based on the UTAUT2 model [55,56], this study adopted the UTAUT2 model to analyze the effect of the serious game. We modified and added variables, which were analyzed using IBM SPSS Amos and IBM SPSS Statistics on factors influencing health information use and dissemination. On this basis, a suitable serious game experiment scheme was built.

The experiment was conducted in 4 steps. In steps 1 and 2, we selected the target participants (midelders/elders), while in steps 3 and 4, we designed the game for the experiment.

Step 1

The construction of the experimental program based on serious games and experimental research needed to be based on a full understanding of the use and dissemination of health information by the research participants. We analyzed the health information needs of the research participants, the frequency and channels of use and the dissemination of health information, and their ability to identify and judge health rumors.

Step 2

To investigate the phenomenon of the dissemination of health information in WeChat's midelder/elder user groups, we used a questionnaire designed in 3 parts: The first part involved a survey to collect personal information, such as gender, age, place of residence, income level, and education. The second part was a survey on the habit of using WeChat. The third part mainly involved the frequency, channel, and motivation of users to obtain and forward health information.

At the same time, 30 health rumor judgment questions were attached to this survey questionnaire, and respondents were asked to judge whether they were correct or incorrect. Through the correct rate of health rumor judgment, we determined the trust level and ability of the respondents to identify health rumors. We also popularized the 30 relevant health rumors, with the hope to popularize the degree of health rumor knowledge and also to strengthen the respondents' ability to recognize information. The questionnaire is shown in Multimedia Appendix 1.

Based on the cognitive ability determined through the questionnaire, 200 participants were selected and asked for their willingness to play the serious game.

Step 3

Based on the use and dissemination of health information by the research participants, the theme of the health rumor learning content was selected and a serious game experimental scheme suitable for this group was constructed through the design and production of serious game content. Considering the experimental length of the serious game and the understanding and acceptance level of the participants, the video game mode of a multiline plot was not applicable for our research, so a single-line plot and scenario was used in the design of the game.

The learning content of the serious game is mainly based on the theme of "a personal day," and the content of the game plot is a person's life from morning to afternoon, in the form of a single storyline. An explanation is provided at the beginning of the game to accurately communicate the theme, rules, and intent of the game to the players. In the learning content of the game, information such as health rumors and general knowledge about COVID-19 was selected, as shown in Table 1, and based on the selected content, failure/passing conditions were set for the game, which involved "risk of infection" and "psychological stress." Different scenarios are set up in a day's life, and questions are set up to interact with the game players to promote and increase the knowledge of COVID-19-related rumors in this interactive learning serious game. The main line design is shown in Figure 5, and the game logic is shown in Figure 6.

In this study, to achieve the effect of the serious game and the purpose of health rumor prevention, a feedback link of the serious game–based health rumor control prevention experiment was important. The feedback link was mainly achieved by setting up a feedback mechanism, which reflected the understanding of the research participants (players) of the game content (COVID-19 health rumor); by setting up the feedback mechanism, interactivity with the research participants could also be strengthened. At the same time, the feedback data were used to reflect the learning effect of the serious game.

The feedback mechanism of the serious game–based health rumor prevention experiment was implemented in the following 3 parts.

- The first part was to communicate the theme, rules, and intention of the game to the research participants by means of game instructions at the beginning of the game. This is an important step to quickly integrate the player into the learning process of a serious game and to let the player know what they will do next in the game.

- The second part was realized in the textual feedback of the gameplay process, where the player was provided with choices through interactive video scenarios, and instant feedback was provided. Instant feedback is an important part of the overall feedback process, which needs to be clearly communicated to the player. It is necessary to clearly communicate to the player whether their choices are correct and to strengthen the knowledge of health rumors and general health. The textual feedback content of the game process is shown inTable 2.
- The last part was to provide feedback after the player passed or failed in the game. At the end of the serious game, based on the player's overall understanding of COVID-19 health rumors and health knowledge, the feedback can strengthen the player's knowledge of health rumors.

The serious game created in this study used a COVID-19 health rumor as the learning content (see Figures 7 and 8). To achieve the purpose of preventing and controlling health rumors, a textual feedback mechanism was designed, involving 4 infection risks and 7 psychological stress settings. These were assigned to game failure or passing conditions, as shown in Figures 9 and 10. The game data reflected the performance of the research participant (player) in the game, with the settings shown in Table 3 to cater to the experiment's needs. The specific game data value settings and game passing/failure conditions are shown in Figure 11. Specifically, the story is as follows: The protagonist, a young person, suddenly finds themselves caught up in the COVID-19 pandemic in early 2020 in their city. Various pieces of information related to COVID-19 start to emerge around the protagonist, causing a massive explosion of fear and panic, particularly among many elderly people who turn to social media for information. They begin to demand that the protagonist follow their advice on preventing the pandemic. The goal of the game is to distinguish between real health knowledge and rumors throughout the daily life story, to use accurate knowledge to save the elderly citizens who are in a state of panic, and to slow down the spread of the virus. In the end, the game outcome is judged based on the actions and choices made by the protagonist.

Table 1. Selection of health rumors/health facts for the serious game content.

Time of day and COVID-19 health rumor/health fact	Setting
Morning	
R[a]1: Drinking plenty of boiled water at 60°C can prevent COVID-19.	Infection risk
R2: Domesticated dogs and cats can also spread COVID-19.	Psychological stress
R3: Putting used masks in a sterilizer can continue to provide protection and use.	Infection risk
The correct way to wear a mask (not an option).	General health knowledge
R4: Going out with ginger slices in the mouth can prevent COVID-19.	Psychological stress
R5: The government will use military aircraft to spread disinfectants in the sky.	Psychological stress
Noon	
R6: You should keep more than 1 m distance from strangers when you go out in times of an epidemic.	Infection risk
R7: Eye-to-eye contact may transmit COVID-19.	Psychological stress
R8: Shuanghuanglian Oral Liquid can effectively inhibit the COVID-19 virus.	Psychological stress
Afternoon	
R9: Disinfection is required for items after returning home from outside.	Infection risk
R10: High temperature can kill the virus, so hot blow-drying and hot water bathing can inhibit it.	Psychological stress
R11: Do not eat fish; pickled fish made from grass carp can transmit COVID-19.	Psychological stress

[a]R: rumor.

Figure 5. The game process. R: rumor.

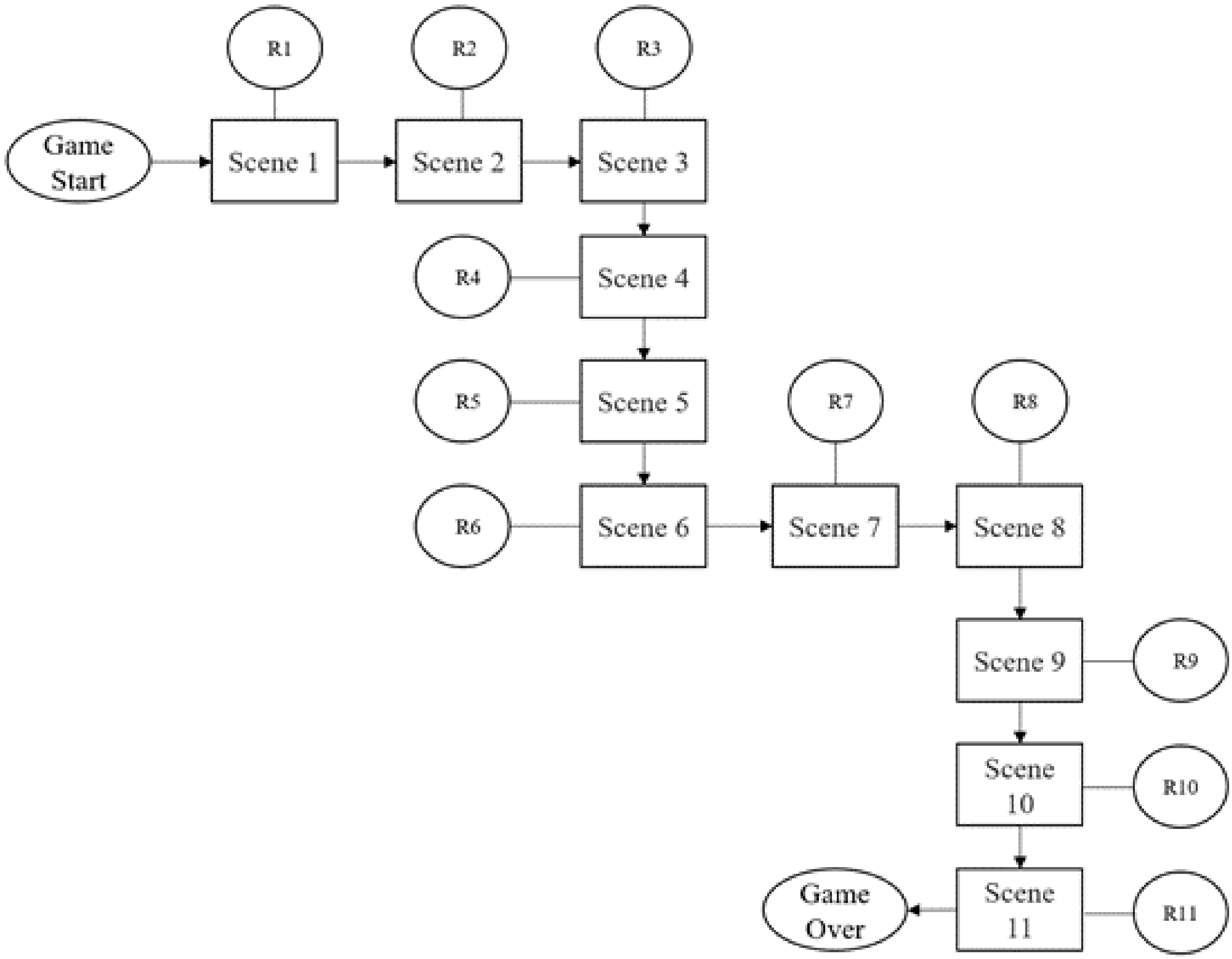


Figure 6. The game judgment logic. N: no; Y: yes.

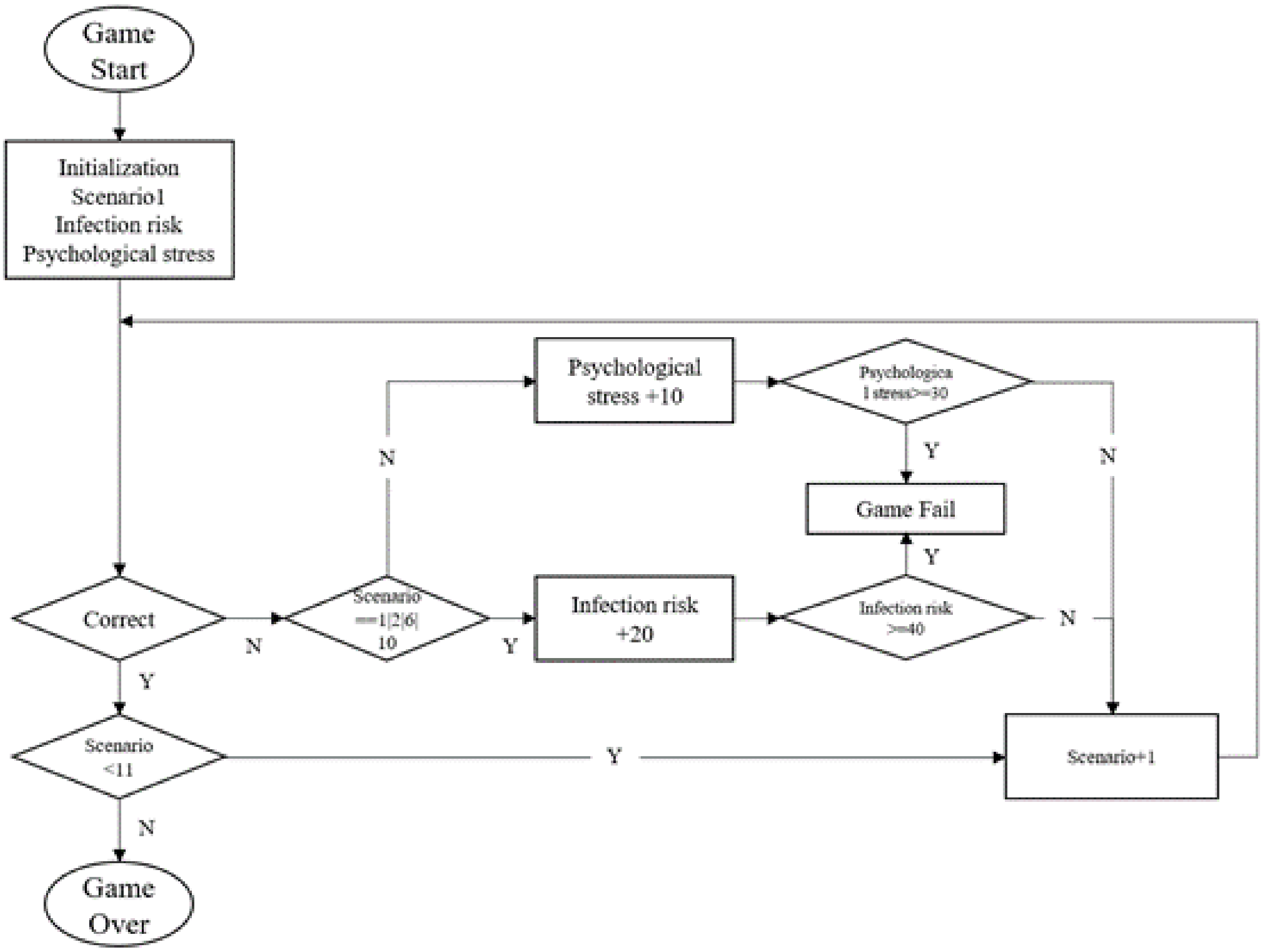


Table 2. Text feedback during gameplay.

Rumor number	Textual feedback (explanation and education) of scenario options during gameplay
R[a]1	Drinking water does not help, and scalding the mucous membrane of the mouth with hot water can increase the risk of infection.
R2	There is no evidence that COVID-19 can be transmitted to domesticated dogs and cats.
R3	Masks that have been used many times do not work to isolate droplets.
—[b]	Graphic feedback: follow the 3 steps (regulations) to wear the mask correctly.
R4	Ginger does not work to prevent the COVID-19.
R5	This is a rumor. There are no military aircraft to spread disinfectants in the sky; in addition, the local government has no right to do that. (This raises the players' sense of alertness and achieves the purpose of public education.)
R6	COVID-19 is spread via droplets and contact, and close contact increases the risk of infection.
R7	This is a rumor. The virus is transmitted through bodily fluids, droplets, and aerosols, not through the eyes. (This re-explains the mode of transmission of the COVID-19 virus.)
R8	Clearly inform that this is not yet clear information and should not be followed blindly.
R9	This is true. You should do that. (This provides possible contact transmission and health information on sterilization.)
R10	Dizziness and other symptoms can occur if you are exposed to bath bombs for too long or take a hot bath for too long.
R11	COVID-19 only infects mammals. Fish do not transmit COVID-19.

[a]R: rumor.

[b]Not applicable.

Figure 7. The experimental serious game's gameplay content 1 (Chinese version). Question: "Hi boy, do you know where one can buy Shuanghuanglian Oral Liquid (a Chinese medicine)? I hear it is useful for COVID-19 treatment!" Answer options: (A) "Really? I also want to buy some." (B) "We do not know the drug's action, so do not drink it by yourself!".

Figure 8. The experimental serious game's gameplay content 2 (Chinese version). Question: "Please come back to home soon; the government will use military aircraft to spray disinfectants!" Answer options: (A) "Really? I'm leaving right now." (B) "Fake news, Mom!".

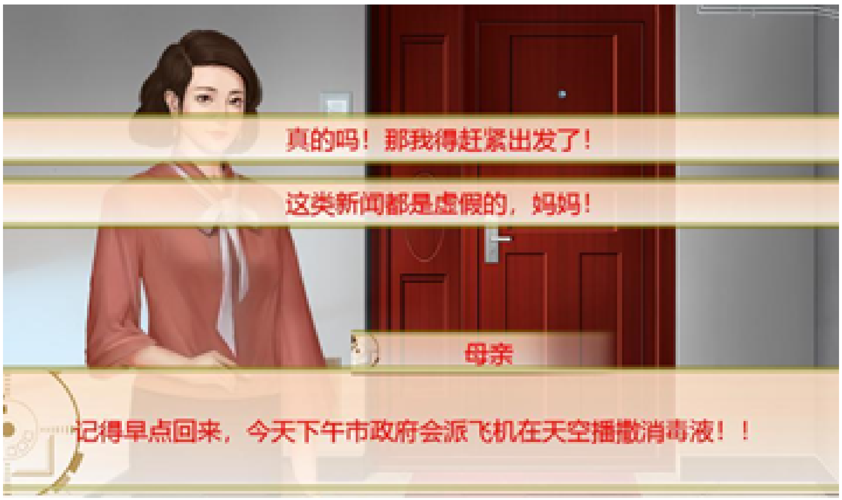


Figure 9. The experimental serious game: game over (Chinese version). Meaning: "The psychological pressure is 10, the risk of infection is 40, you are in a high-risk situation!".

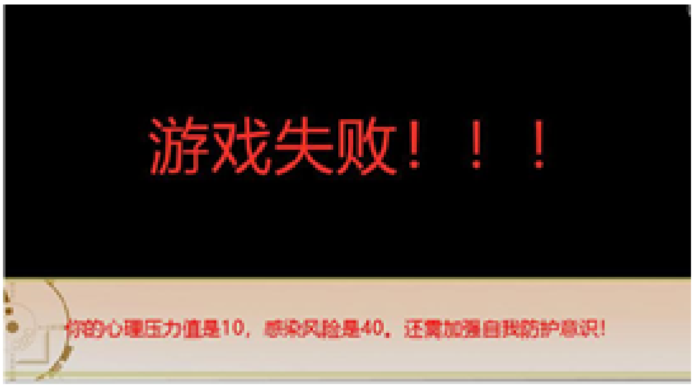
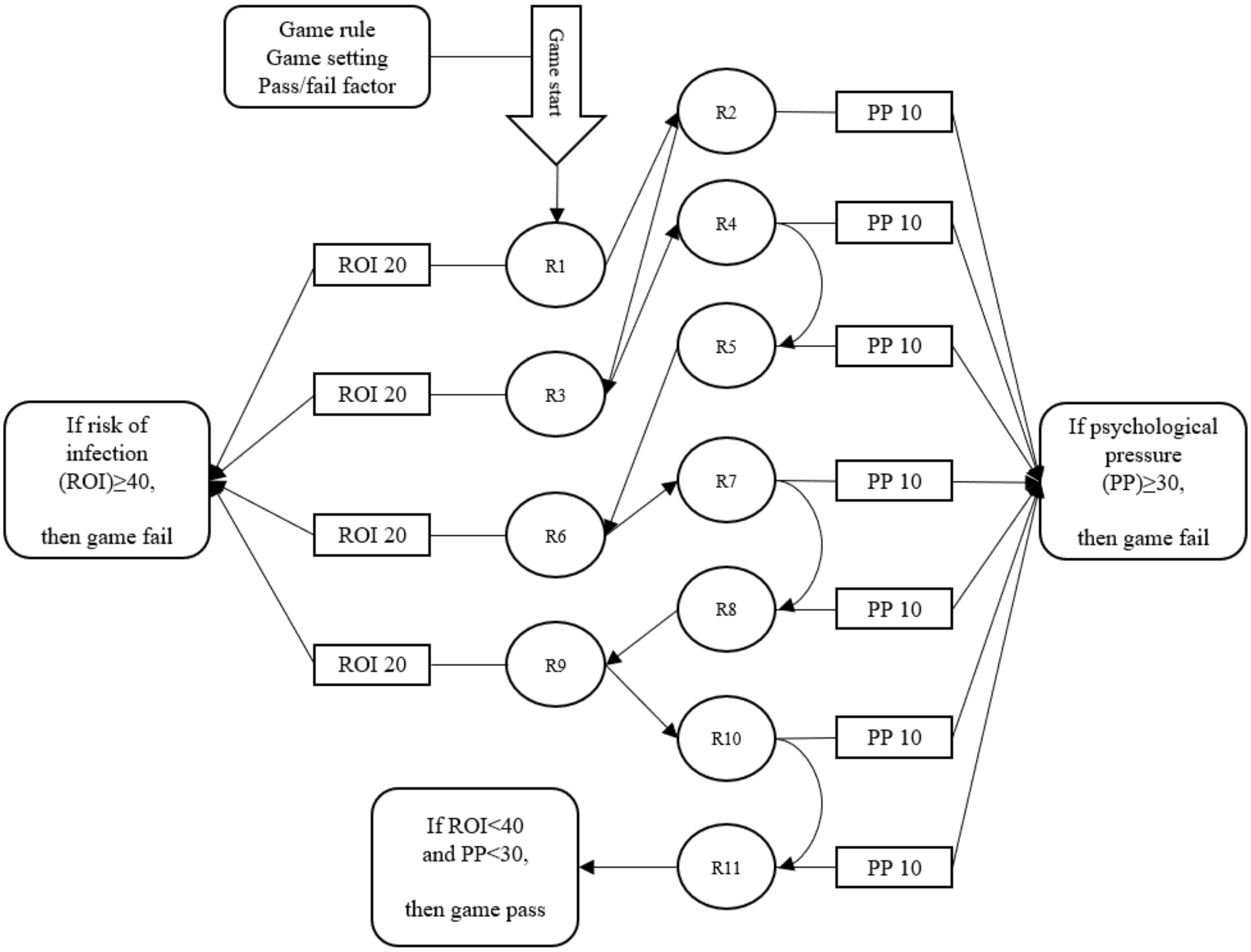


Figure 10. The preventive knowledge statement after the game, explaining how to wear a mask in 3 steps.

Table 3. Game data and game rating settings.

Game score	Error choice	Game failure	Game round
A	≤3	0	1
B	4-6	1	2
C	7-9	2	3
D	≥10	≥3	≥4

Figure 11. Game-related data values and failure/passing condition-setting thinking diagram.



Step 4

Finally, the data and feedback of the participants were obtained, and the effect of the serious game on health rumor prevention was analyzed through the data and feedback to determine whether serious games are useful to prevent health rumors.

Data Collected

This study compared and analyzed the differences between acquiring and understanding health rumor information through the learning modes of serious games and traditional media. A total of 100 people were selected to participate in the serious game experiment (G1 group), while 100 people who did not participate in the serious game experiment only studied by traditional media (G2 group). To ensure the objectivity of the controlled experiment, the educational background of the 200 participants was investigated before the formal study while keeping the 2 groups as similar as possible in terms of gender and age, as Table 4 shows. Next, we sent the testing questionnaire related to health commonsense and health rumors to G1 and G2.

Table 4. Demographic information of groups G1 and G2.

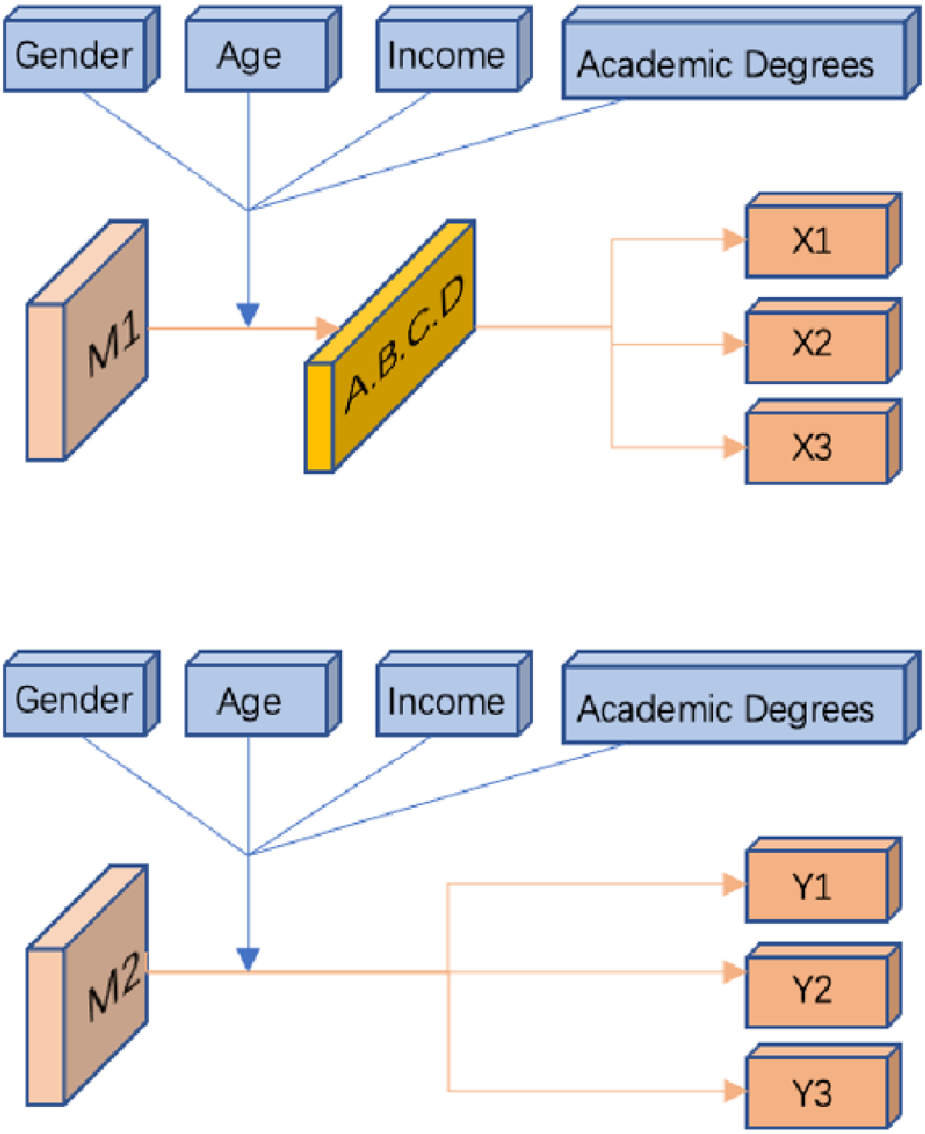
Characteristics	G1 (n=100), n (%)	G2 (n=100), n (%)
Gender		
Male	42 (42)	40 (40)
Female	58 (58)	60 (60)
Age (years)		
40-45	24 (24)	27 (27)
46-50	30 (30)	25 (25)
51-55	19 (19)	17 (17)
56-60	18 (18)	21 (21)
≥61	9 (9)	10 (10)
Educational background		
Less than high school	15 (15)	15 (15)
High school/technical secondary school	50 (50)	50 (50)
Junior college	20 (20)	20 (20)
Bachelor's degree and higher	15 (15)	15 (15)
Income		
Low	29 (29)	24 (24)
Average	48 (48)	55 (55)
High	23 (23)	21 (21)

Variables and Hypotheses

Based on the framework of serious games, a questionnaire was designed that included the following variable definitions:

- Independent variables: The learning mode refers to the way that information is obtained and knowledge learned; the variables were M1 (learning through serious games) and M2 (learning through traditional and new media). For gender, age, income, and education, in the preparation stage of the research, education was used as the main grouping basis, and the age distribution and gender ratio of the 2 groups of experimental objects were kept consistent.
- Intermediary variables: Game data for serious game experiments (A/B/C/D) refer to the player's performance data during the game, including the number of selection errors, the number of game failures, and the number of game passings. Personal performance was divided into 4 mediating variables: excellent (A), good (B), medium (C), and poor (D).
- Dependent variables: The variables were cognitive questionnaire (1) overall correct response rate of judgment and recognition of the COVID-19 health rumor (X1 for G1, Y1 for G2 [G2 did not participate in the serious game experiment]), (2) correct rate of judgment and recognition the COVID-19 health rumor part 1 (X2 for G1 [COVID-19 health rumor not included in the serious game experiment], Y2 for G2), and (3) correct rate of judgment and identification of the COVID-19 health rumor part 2 (X3 for G1, Y3 for G2 [COVID-19 pneumonia rumor included in the traditional media experiment for G2]).
- Intervening variables: The comprehension, cognitive level, learning ability, learning interest, and information attention of the G1 group affected the outcome of the dependent variables to a certain extent. The specific influencing relationship between various variables is shown in Figure 12.

Figure 12. Relationship between experimental study variables.



The following 10 hypotheses were proposed in this experimental study:

- Hypothesis 1 (H1): Serious game experiments can help research participants acquire and understand health rumor knowledge and health commonsense.
- H2: The serious game learning mode is more capable of inspiring the interest of research participants in their understanding of the knowledge acquired and their learning of health commonsense compared to the traditional learning mode.
- H3: The serious game learning mode is more impactful than the traditional learning mode.
- H4: In judging and recognizing the COVID-19 health rumor, G1 has a stronger judgment ability than G2 and a higher accuracy in identifying the rumor in the serious game experiment.
- H5: In judging the COVID-19 health rumor, for the rumor not included in the serious game experiment, without the influence of M1, the ability of G1 and G2 is not much different, and the accuracy rate of identifying the COVID-19 health rumor is roughly the same for both groups. The manifestation in the variable is X2≈Y2.
- H6: In judging and recognizing the COVID-19 health rumor, G1 has an overall stronger judgment ability than G2 and a higher accuracy rate of identifying the COVID-19 health rumor. The specific manifestation in the variable is X3>Y3.
- H7: Gender affects G1's and G2's judgment and recognition of the COVID-19 health rumor.
- H8: Age affects G1's and G2's judgment and recognition of the COVID-19 health rumor.
- H9: Income affects G1's and G2's judgment and recognition of the COVID-19 health rumor.
- H10: Academic qualifications affect G1's and G2's judgment and recognition of the COVID-19 health rumor.

Ethical Considerations

According to the guidelines of the People's Republic of China [57], this study met the conditions for exemption from ethical review.

Results

Analysis of Data Collected

According to the collected game data, 36% (72/200) of players received A, 41% (82/200) received B, 18% (36/200) received C, and the remaining 5% (10/200) received D. The accuracy rate of the judgment and recognition of the COVID-19 health rumor and health commonsense in the cognitive questionnaire of G1 and G2 groups are tabulated in Table 5. The overall accuracy was 84% for G1 and 78% for G2, with an average of 81%. The relationship of the parameters X1-X3 (G1) and Y1-Y3 (G2) are shown in Figures 13 and 14, respectively.

Table 5. Cognitive questionnaire data (judgment and recognition of rumor knowledge and health commonsense).[a]

Question number	Accuracy G1 (%)	Accuracy G2 (%)	Average accuracy (%), (G1+G2)/2
1	72	76	74
2	93	98	96
3	75	79	77
4	77	76	77
5	64	63	64
6	87	91	89
7	77	85	81
8	72	70	71
9	67	63	65
10	80	74	77
11	67	64	66
12	80	88	84
13	88	96	92
14	75	71	73
15	81	83	82
16	91	95	93
17	68	63	66
18	80	83	82
19	58	56	57
20	97	71	84
21	99	61	80
22	98	65	82
23	99	92	96
24	91	81	86
25	100	92	96
26	100	89	95
27	90	76	83
28	92	85	89
29	92	67	80
30	99	95	97

[a]All the correct rate values are kept to integer bits.

Figure 13. Comparison of corresponding dependent variables (G1).

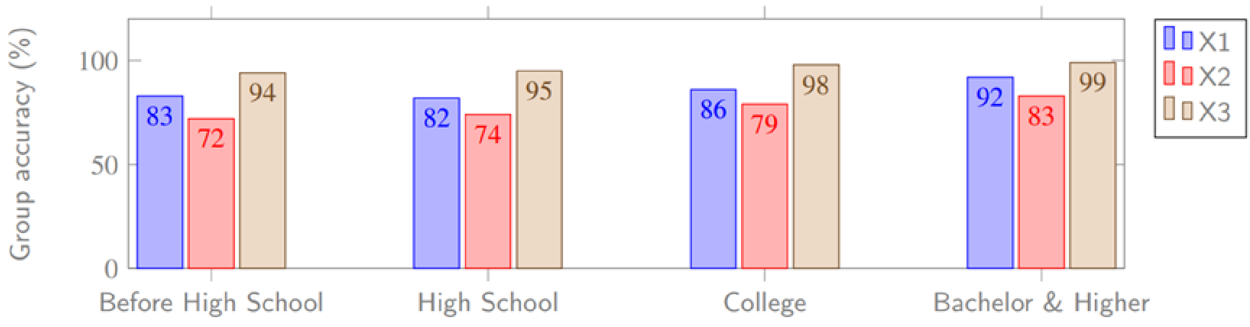
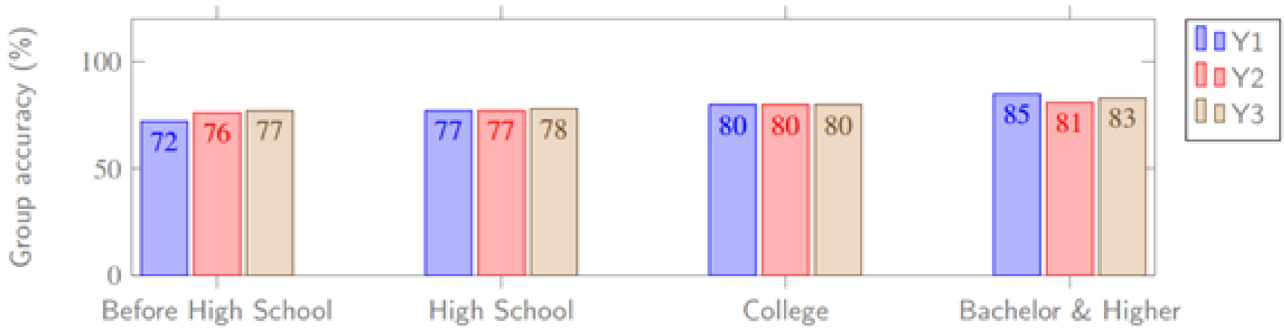


Figure 14. Comparison of two groups and corresponding dependent variables (G2).



According to the data, the overall average correct rate of judgment and recognition of G1 and G2 was 84% and 78%, respectively. In addition, the related data collected showed that 76% of the participants believed that the serious game learning mode is more interesting than the traditional learning mode. This finding shows that G1 has good interest in games. Furthermore, 60% of the participants thought that the game learning mode is more helpful than the traditional learning mode, and 65% thought that the serious game learning mode makes a more profound impression. Therefore, from the perspective of the selection of participants, the serious game learning mode is more interesting, helpful, and impactful than the traditional learning mode. Therefore, H2 and H3 hold.

Analysis of Dependent Variables (X, Y) and Intermediary Variables (A, B, C, D)

According to the data of the dependent variables in Table 6, we could not directly prove that the impact of M1 on X1 was greater than that of M2 on Y1. At the same time, when the knowledge of health rumors and health commonsense was not included in the game content, the correct rate of judgment and recognition of G1 and G2 was almost the same, and even G1 had a relatively lower rate. However, X3 was 96%, which is much higher than Y3 (79%), and X3 exceeded X2 by up to 23 percentage points. This condition implies that the impact of M1 on X1 is greater than the impact of M2 on Y1. It not only shows that G1 had relatively strong learning ability but also that after the serious game learning model experiment, there was a significant positive effect on the accuracy rate of the judgment and recognition accuracy of the COVID-19 health rumor and health commonsense. At the same time, X1>Y1, X2≈Y2, and X3 >Y3. Therefore, H4, H5, and H6 are established.

The intermediary variables A, B, C, and D were sequentially observed, corresponding to the dependent variables X1, X2, and X3. It can be clearly seen in Table 7 that A-X1>B-X1>C-X1>D-X1, A-X2>B-X2>C-X2>D-X2, and A-X3>B-X3>C-X3>D-X3. The dependent variables corresponding to the intermediary variables showed a decreasing trend from A to D, indicating that M1 affected X1 and X3 through A, B, C, and D and the degree of influence was in the order of A>B>C>D. These data showed that the higher the average game score, the higher the correct rate of recognition and judgment. Therefore, combined with the previous analysis, H1 holds.

Table 6. Dependent variables X and Y.

Variable	Value (%)
X1[a]	84
X2[b]	76
X3[c]	96
Y1[d]	78
Y2[e]	78
Y3[f]	79

[a]X1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G1.
[b]X2: cognitive questionnaire correct rate of judgment and recognition part 1 for G1.
[c]X3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G1.
[d]Y1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G2.
[e]Y2: cognitive questionnaire correct rate of judgment and recognition part 1 for G2.
[f]Y3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G2.

Table 7. Intermediary variables (A, B, C, D) and dependent variables (X1-X3).

Intermediary variable	Dependent variables		
	X1[a] (%)	X2[b] (%)	X3[c] (%)
A	87	80	98
B	83	76	96
C	80	72	94
D	72	62	89

[a]X1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G1.
[b]X2: cognitive questionnaire correct rate of judgment and recognition part 1 for G1.
[c]X3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G1.

Analysis of Independent Variables

Gender

We grouped participants, ensuring the educational composition of the 2 groups was as consistent as possible. By observing and comparing the G1 independent variable (gender) and its corresponding intermediary variables (Table 8), we found that the game score reached A, where the male participants were better than the female participants but changed from B to D when the female participants were better than the male participants. This condition was especially true when the game score reached B, where the female participants were much better than the male participants. This situation may also be influenced by the unequal relationship of the overall gender. There was no gender difference in the numbers from game rating A to D.

Subsequently, by observing and comparing the G1 and G2 independent variable gender and the corresponding dependent variables (Table 9), we observed that regarding the dependent variables X1, X2, and X3, corresponding to the independent variable gender (female, male), the comparisons were female<male, female<male, and female>male, respectively. Regarding Y1, Y2, and Y3, corresponding to gender, the comparisons were female<male, female>male, and female<male, respectively. As such, there was no gender difference. Therefore, H7 does not hold.

Table 8. G1 gender and corresponding intermediary variables.

Intermediary variable	Gender		Comparison
	Female	Male	
A	16	20	Female<male
B	27	14	Female>male
C	12	6	Female>male
D	4	1	Female>male

Table 9. G1 and G2 gender and corresponding dependent variables.

Dependent variable	Gender		Comparison
	Female (%)	Male (%)	
X1[a]	83	84	Female<male
X2[b]	75	77	Female<male
X3[c]	97	95	Female>male
Y1[d]	76	81	Female<male
Y2[e]	79	77	Female>male
Y3[f]	71	88	Female<male

[a]X1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G1.

[b]X2: cognitive questionnaire correct rate of judgment and recognition part 1 for G1.

[c]X3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G1.

[d]Y1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G2.

[e]Y2: cognitive questionnaire correct rate of judgment and recognition part 1 for G2.

[f]Y3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G2.

Age

There were apparent differences in the age ranges between the 2 groups, as shown in Table 4. Therefore, random sampling in G1 and G2 was conducted, and 25 participants under 51 and 56 years old each were selected, with 50 participants in each group for comparative observation and analysis of the corresponding variable data.

First, by observing and comparing the high-age and low-age groups' independent and variable age groups and their corresponding intermediary variables (Table 10), we found that the number of people who achieved A and B game scores were all of low age. As a result, the number of people in the low-age group was greater than the number of people in the high-age group; among those with game scores C and D, the number of people in the high-age group was greater than the number of people in the low-age groups, indicating that to a certain extent, the independent variable age positively affects the intermediary variables A, B, C, and D. Second, by observing and comparing the high- and low-age groups of the G1 and G2 independent variable age with corresponding dependent variables (Table 11), the values of independent variables X1, X2, and X3 could be determined. The values of the low-age group were greater than those of the high-age group; the independent variables Y1, Y2, and Y3 also exhibited this behavior. Therefore, H8 holds.

Table 10. G1 and G2 targets of different ages.

Intermediary variable	Age	
	Low (n=25), n (%)	High (n=25), n (%)
A	8 (32)	13 (52)
B	6 (24)	8 (32)
C	8 (32)	3 (12)
D	3 (12)	1 (4)

Table 11. G1 and G2 age groups and corresponding dependent variables.

Dependent variable	Age	
	Low (%)	High (%)
X1[a]	76	87
X2[b]	69	80
X3[c]	88	98
Y1[d]	68	82
Y2[e]	66	80
Y3[f]	72	86

[a]X1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G1.
[b]X2: cognitive questionnaire correct rate of judgment and recognition part 1 for G1.
[c]X3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G1.
[d]Y1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G2.
[e]Y2: cognitive questionnaire correct rate of judgment and recognition part 1 for G2.
[f]Y3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G2.

Income

The relevant data collected are shown in Table 4. Nearly half of the participants in G1 and G2 believed that their income level was average, and the number of people who believed that their income was high or low was relatively small. By observing and comparing the high- and low-income subgroups in G1 and G2 with corresponding dependent variables, we found that the dependent variables corresponding to the 2 independent variable subgroups were not identical, as shown in Table 12. Therefore, H9 does not hold.

Table 12. High- and low-income groups and corresponding dependent variables of G1 and G2.

Dependent variable	Income	
	High (%)	Low (%)
X1[a]	84	84
X2[b]	82	78
X3[c]	88	94
Y1[d]	73	74
Y2[e]	69	72
Y3[f]	80	78

[a]X1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G1.
[b]X2: cognitive questionnaire correct rate of judgment and recognition part 1 for G1.
[c]X3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G1.
[d]Y1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G2.
[e]Y2: cognitive questionnaire correct rate of judgment and recognition part 1 for G2.
[f]Y3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G2.

Education

Based on the cognitive questionnaire, the independent variable education was divided into 4 segments, as shown in Table 4. First, 4 groups of the G1 independent variable education and corresponding intermediary variables were compared (Table 13). Through comparison and observation, we found that the higher the education level, the better the performance in the game, which demonstrates that the dependent variable education positively affects the intermediary variables. Second, the grouping and corresponding dependent variables of G1 and G2 based on academic qualifications are shown in Table 14. We found that education has a positive effect on the corresponding dependent variables. Therefore, H10 holds.

Finally, a thorough investigation of the selection tendency between the learning modes of serious games (M1) and traditional learning (M2) was conducted on G1 involving a comparison experiment of interest, help, and impression (Figure 15). The data showed that 76% (152/200) of the participants thought the serious game learning mode was more interesting than the traditional learning mode, indicating that G1 had a reasonable learning interest in games. Furthermore, 60%

(120/200) of the participants thought that the serious game learning mode was more helpful than the traditional learning mode, and 65% (130/200) thought that the serious game learning mode was more impressive than the traditional learning mode. Therefore, from the perspective of the selection tendency of the participants, the serious game learning mode is more interesting, helpful, and impressive than the traditional learning mode, which strengthens the hypotheses.

Table 13. Proportion of game scores in different education segments.

Segment	Intermediary variables			
	A (%)	B (%)	C (%)	D (%)
Less than high school	20	47	20	13
High school	28	44	20	8
College degree	45	40	15	0
Bachelor's degree and higher	59	23	13	0

Table 14. Education segments and corresponding dependent variables of G1 and G2.

Segment	Dependent variables					
	X1[a] (%)	X2[b] (%)	X3[c] (%)	Y1[d] (%)	Y2[e] (%)	Y3[f] (%)
Less than high school	83	72	94	72	76	77
High school	82	74	95	77	77	78
College degree	86	79	98	80	80	80
Bachelor's degree and higher	92	83	99	85	81	83

[a]X1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G1.

[b]X2: cognitive questionnaire correct rate of judgment and recognition part 1 for G1.

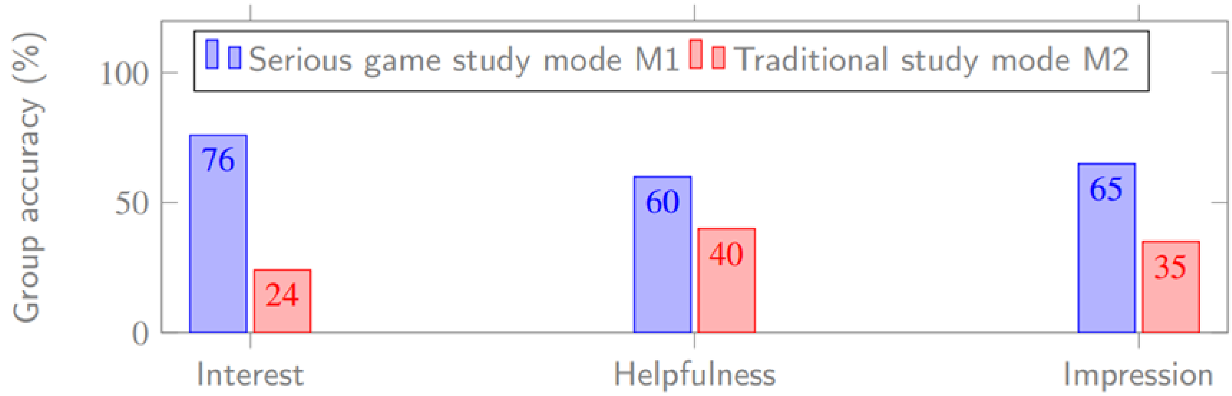
[c]X3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G1.

[d]Y1: cognitive questionnaire overall correct response rate of judgment and recognition of the COVID-19 health rumor for G2.

[e]Y2: cognitive questionnaire correct rate of judgment and recognition part 1 for G2.

[f]Y3: cognitive questionnaire correct rate of judgment and identification of the COVID-19 health rumor part 2 for G2.

Figure 15. Comparison between serious game learning mode and traditional learning mode.



Discussion

Principal Findings

Based on a self-made serious game, this paper investigated the health rumor phenomenon, and a study on the user behavior and willingness to disseminate health information among Chinese elderly WeChat users (SNS) was conducted during the early COVID-19 pandemic. After a survey, participants were chosen, and a COVID-19 health rumor was selected as the study content and the experimental platform with the self-made game was established. The UTAUT2 model was upgraded by adding parameters, several hypotheses were proposed, and a control experiment was designed. The experiment results show that the serious game is useful for health rumor prevention.

After collecting game data and the correct response rates of G1 and G2 in the cognitive questionnaire for the judgment and recognition of the COVID-19 health rumor, the game data and the cognitive questionnaire data were combined to determine the relationship between specific variables. Finally, the experimental hypotheses were tested and evaluated, proving that H1-H6, H8, and H10 hold, while H7 and H9 do not hold.

The findings affirm that serious games are a powerful tool to enhance learning and commonsense against health rumors in the context of elderly users of SNS. As Wu [43] argued, perceptions of rumor credibility affect the users' desire to find accurate information (*cognitive gratification*) because they use SNS for verifying the contents of rumors and for acquiring more knowledge and information. Equipping oneself with better knowledge and commonsense against health rumors could have a profound effect on the stability and harmony of society [58], minimize the chance of being misinformed [35], and help create effective control strategies against rumor spreading [45]. Furthermore, as serious games provide the means for people to receive direct feedback relative to their judgment of health rumors, using these games is considered a more humane and emotional approach [47]. In addition, it also provides a suitable channel for health care providers to increase awareness [49] since tackling COVID-19 requires everyone to follow medical advice. Based on the verification of our hypotheses, we found that the effect of serious games correlates with parameters such as education, which suggests that the future rumor management for the youth is perfectly suited to the use of serious games, especially in China, where the education level of the youth is much higher than that of the middle-aged and older populations.

Limitations

Given the seminal findings of this study, it has some limitations. First, the cognitive questionnaire was administered offline, and the midelder/elder participants were reluctant in terms of their willingness to cooperate with the research. As such, there was a risk of the sample distribution being uneven or biased. Second, strict epidemic prevention and control have geographically limited experimental samples. Third, the serious game design was restricted to the COVID-19 health information and had limited interactivity.

Conclusion

This experimental study on preventing new health rumors via serious games proves that the serious game learning mode can help research participants understand and learn about health and rumors. Furthermore, serious games make a more profound impression on people than traditional learning modes, while providing fertile ground for more comprehensive research in the future. In addition, serious games could provide suggestions and support in future research on rumor prevention and detection. In particular, the Chinese government ended the zero-COVID policy in December 2022, and many new health rumors related to the Omicron variant were found on the internet in China. This study could provide a method of challenging the new issue and the game could be updated for the current situation. More importantly, we discovered that serious games can act as an "informational vaccine" against rumors (if rumors are considered a kind of "informational virus or bacterium"), and in the future, we can conduct further research in this direction.

Acknowledgments

This study was supported by the Huazhong University of Science and Technology, Independent Innovation Fund (#450/5003450061).

Conflicts of Interest

None declared.

Multimedia Appendix 1
A priori questionnaire for health rumor recognition.
[DOCX File , 16 KB - games_v12i1e45546_app1.docx]

References

1. Zhang Z, Zhang Z, Li H. Predictors of the authenticity of internet health rumours. Health Info Libr J 2015 Sep 13;32(3):195-205 [FREE Full text] [doi: 10.1111/hir.12115] [Medline: 26268517]
2. Deng Z, Liu S. Understanding consumer health information-seeking behavior from the perspective of the risk perception attitude framework and social support in mobile social media websites. Int J Med Inform 2017 Sep;105:98-109. [doi: 10.1016/j.ijmedinf.2017.05.014] [Medline: 28750916]
3. Chua AY, Banerjee S. To share or not to share: the role of epistemic belief in online health rumors. Int J Med Inform 2017 Dec;108:36-41. [doi: 10.1016/j.ijmedinf.2017.08.010] [Medline: 29132629]
4. Chua AY, Banerjee S. Analyzing users' trust for online health rumors. 2015 Presented at: ICADL 2015: 17th International Conference on Asia-Pacific Digital Libraries; December 9-12, 2015; Seoul, Korea p. 9-12,33-38. [doi: 10.1007/978-3-319-27974-9_4]
5. Abt CC. Serious Games. Lanham, MD: University Press of America; 1987.
6. Djaouti D, Julian A, Jessel JP. Classifying serious games: the G/P/S model. In: Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches. Hershey, PA: IGI Global; 2011:118-136.

7. Ali F, Nair PK, Hussain K. An assessment of students' acceptance and usage of computer supported collaborative classrooms in hospitality and tourism schools. JoHLSTE 2016 Jun;18:51-60. [doi: 10.1016/j.jhlste.2016.03.002]
8. Sawyer B. Serious games: improving public policy through game-based learning and simulation. In: Foresight and Governance Project. Washington, DC: Woodrow Wilson International Center for Scholars; 2002.
9. Woike JK, Hafenbrädl S, Kanngiesser P, Hertwig R. The transmission game: testing behavioral interventions in a pandemic-like simulation. Sci Adv 2022 Feb 25;8(8):eabk0428 [FREE Full text] [doi: 10.1126/sciadv.abk0428] [Medline: 35213225]
10. Fox J, McKnight J, Sun Y, Maung D, Crawfis R. Using a serious game to communicate risk and minimize psychological distance regarding environmental pollution. Telemat Informat 2020 Mar;46:101320. [doi: 10.1016/j.tele.2019.101320]
11. Weyrich P, Ruin I, Terti G, Scolobig A. Using serious games to evaluate the potential of social media information in early warning disaster management. Int J Disaster Risk Reduct 2021 Apr;56:102053. [doi: 10.1016/j.ijdrr.2021.102053]
12. Ritterfeld Y, Cody M, Vorderer VP. Serious Games: Mechanisms and Effects. Oxfordshire, UK: Routledge; 2009.
13. Michael DR, Chen SL. Serious Games: Games That Educate, Train, and Inform. Cincinnati, OH: Muska & Lipman/Premier-Trade; 2005.
14. Zyda M. From visual simulation to virtual reality to games. Computer 2005 Sep;38(9):25-32. [doi: 10.1109/mc.2005.297]
15. Susi T, Johannesson M, Backlund P, Per B. Serious Games: An Overview. Skövde, Sweden: School of Humanities and Informatics, University of Skövde; 2007.
16. Djaouti D, Alvarez J, Jessel J, Rampnoux O. Origins of serious games. Serious Games Edutainm Appl 2011:25-43. [doi: 10.1007/978-1-4471-2161-9_3]
17. Westera W, Nadolski R, Hummel H, Wopereis I. Serious games for higher education: a framework for reducing design complexity. J Comput Assist Learn 2008 Sep 05;24(5):420-432. [doi: 10.1111/j.1365-2729.2008.00279.x]
18. Ziebarth S, Kizina A, Hoppe HU, Dini L. A serious game for training patient-centered medical interviews. 2014 Presented at: 2014 IEEE 14th International Conference on Advanced Learning Technologies; July 7-10, 2014; Athens, Greece. [doi: 10.1109/icalt.2014.69]
19. Diehl LA, Lehmann E, Souza RM, Alves JB, Esteves RZ, Gordan PA. A serious game prototype for education of medical doctors and students on insulin management for treatment of diabetes mellitus. 2011 Presented at: 2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH); November 16-18, 2011; Braga, Portugal. [doi: 10.1109/segah.2011.6165437]
20. Yi HL. Designing serious games to enhance political efficacy and critical thinking disposition for college students: the case of Taiwan. 2011 Presented at: 2011 Third International Conference on Games and Virtual Worlds for Serious Applications; May 4-6, 2011; Athens, Greece. [doi: 10.1109/vs-games.2011.29]
21. Yusoff A, Crowder R, Gilbert L, Wills G. A conceptual framework for serious games. 2009 Presented at: 2009 Ninth IEEE International Conference on Advanced Learning Technologies; July 15-17, 2009; Riga, Latvia. [doi: 10.1109/icalt.2009.19]
22. Crookall D. Serious games, debriefing, and simulation/gaming as a discipline. Simulat Gaming 2011 Jan 06;41(6):898-920. [doi: 10.1177/1046878110390784]
23. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. Comput Educ 2012 Sep;59(2):661-686. [doi: 10.1016/j.compedu.2012.03.004]
24. Brasil IS, Neto FMM, Chagas JFS, Monteiro R, Souza DFL, Bonates MF. An intelligent and persistent browser-based game for oil drilling operators training. 2011 Presented at: 2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH); November 16-18, 2011; Braga, Portugal. [doi: 10.1109/segah.2011.6165431]
25. Yessad A, Labat JM, Kermorvant F. Segae: a serious game authoring environment. 2010 Presented at: 2010 10th IEEE International Conference on Advanced Learning Technologies; July 5-7, 2010; Sousse, Tunisia. [doi: 10.1109/icalt.2010.153]
26. Laamarti F, Eid M, El Saddik A. An overview of serious games. Int J Comput Games Technol 2014;2014:1-15. [doi: 10.1155/2014/358152]
27. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012 Oct;99(10):1322-1330. [doi: 10.1002/bjs.8819] [Medline: 22961509]
28. Wattanasoontorn V, Boada I, García R, Sbert M. Serious games for health. Entertain Comput 2013 Dec;4(4):231-247. [doi: 10.1016/j.entcom.2013.09.002]
29. Wouters P, van Nimwegen C, van Oostendorp H, van der Spek ED. A meta-analysis of the cognitive and motivational effects of serious games. J Educ Psychol 2013 May;105(2):249-265. [doi: 10.1037/a0031311]
30. Burke JW, McNeill M, Charles D, Morrow P, Crosbie J, McDonough S. Serious games for upper limb rehabilitation following stroke. 2009 Presented at: 2009 Conference in Games and Virtual Worlds for Serious Applications; March 23-24, 2009; Coventry, UK. [doi: 10.1109/vs-games.2009.17]
31. Bartolomé NA, Zorrilla AM, Zapirain BG. A serious game to improve human relationships in patients with neuro-psychological disorders. 2010 Presented at: 2010 2nd International IEEE Consumer Electronics Society's Games Innovations Conference; December 21-23, 2010; Hong Kong, China. [doi: 10.1109/icegic.2010.5716881]
32. Bartolomé NA, Zorrilla AM, Zapirain BG. Can game-based therapies be trusted? Is game-based education effective? A systematic review of the serious games for health and education. 2011 Presented at: 2011 16th International Conference on Computer Games (CGAMES); July 27-30, 2011; Louisville, KY, USA. [doi: 10.1109/cgames.2011.6000353]

33. Krath J, Schürmann L, von Korflesch HF. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comput Hum Behav 2021 Dec;125:106963. [doi: 10.1016/j.chb.2021.106963]
34. Abd-Alrazaq A, Alajlani M, Alhuwail D, Toro CT, Giannicchi A, Ahmed A, et al. The effectiveness and safety of serious games for improving cognitive abilities among elderly people with cognitive impairment: systematic review and meta-analysis. JMIR Serious Games 2022 Mar 10;10(1):e34592 [FREE Full text] [doi: 10.2196/34592] [Medline: 35266877]
35. Fu AS, Gao KL, Tung AK, Tsang WW, Kwan MM. Effectiveness of exergaming training in reducing risk and incidence of falls in frail older adults with a history of falls. Arch Phys Med Rehabil 2015 Dec;96(12):2096-2102. [doi: 10.1016/j.apmr.2015.08.427] [Medline: 26360975]
36. Jiménez-Pavón D, Carbonell-Baeza A, Lavie CJ. Physical exercise as therapy to fight against the mental and physical consequences of COVID-19 quarantine: special focus in older people. Prog Cardiovasc Dis 2020 May;63(3):386-388 [FREE Full text] [doi: 10.1016/j.pcad.2020.03.009] [Medline: 32220590]
37. Anguera JA, Schachtner JN, Simon AJ, Volponi J, Javed S, Gallen CL, et al. Long-term maintenance of multitasking abilities following video game training in older adults. Neurobiol Aging 2021 Jul;103:22-30. [doi: 10.1016/j.neurobiolaging.2021.02.023] [Medline: 33789209]
38. Loerzel VW, Clochesy JM, Geddie PI. Using serious games to increase prevention and self-management of chemotherapy-induced nausea and vomiting in older adults with cancer. ONF 2020 Sep 1;47(5):567-576. [doi: 10.1188/20.onf.567-576]
39. Hall T, Kastner M, Woollard S, Ramdeyol C, Makarski J, Zhang Y, et al. Transcending distance in long-term care: can serious games increase resident resilience? In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Los Angeles, CA: SAGE Publications; 2021:1362-1366.
40. Lau SYJ, Balasankar G, Kenneth F. The impact of serious games on physical activity, cognitive training and mental health for the ageing population during COVID-19: implications and future trajectories. JoGHNP 2022:34710. [doi: 10.52872/001c.34710]
41. Suppan M, Catho G, Robalo Nunes T, Sauvan V, Perez M, Graf C, et al. A serious game designed to promote safe behaviors among health care workers during the COVID-19 pandemic: development of “Escape COVID-19”. JMIR Serious Games 2020 Dec 03;8(4):e24986 [FREE Full text] [doi: 10.2196/24986] [Medline: 33242312]
42. Ferreira MMN, Passos MAN, Ferraz CR. Nursing employing gamification for adhesion to hand hygienization in the fight against COVID–19. JRG Acad Stud Mag 2020;3(7):274-284 [FREE Full text] [doi: 10.5281/zenodo.4244859]
43. Wu M. What drives people to share misinformation on social media during the COVID-19 pandemic: a stimulus-organism-response perspective. Int J Environ Res Public Health 2022 Sep 17;19(18):11752 [FREE Full text] [doi: 10.3390/ijerph191811752] [Medline: 36142031]
44. Ji J, Chao N, Ding J. Rumormongering of genetically modified (GM) food on Chinese social network. Telemat Informat 2019 Apr;37:1-12. [doi: 10.1016/j.tele.2019.01.005]
45. Hui H, Zhou C, Lü X, Li J. Spread mechanism and control strategy of social network rumors under the influence of COVID-19. Nonlinear Dyn 2020 Aug 10;101(3):1933-1949 [FREE Full text] [doi: 10.1007/s11071-020-05842-w] [Medline: 32836821]
46. Pulido C, Ruiz-Eugenio L, Redondo-Sama G, Villarejo-Carballido B. A new application of social impact in social media for overcoming fake news in health. Int J Environ Res Public Health 2020 Apr 03;17(7):2430 [FREE Full text] [doi: 10.3390/ijerph17072430] [Medline: 32260048]
47. Kim S, Kim S. The crisis of public health and infodemic: analyzing belief structure of fake news about COVID-19 pandemic. Sustainability 2020 Nov 26;12(23):9904. [doi: 10.3390/su12239904]
48. Zhao Y, Da J, Yan J. Detecting health misinformation in online health communities: incorporating behavioral features into machine learning based approaches. Inf Proc Manag 2021 Jan;58(1):102390. [doi: 10.1016/j.ipm.2020.102390]
49. Wang C, Huang H. When “fake news” becomes real: the consequences of false government denials in an authoritarian country. Comparat Polit Stud 2020 Sep 14;54(5):753-778. [doi: 10.1177/0010414020957672]
50. Xu C, Zhang X, Wang Y. Mapping of health literacy and social panic via web search data during the COVID-19 public health emergency: infodemiological study. J Med Internet Res 2020 Jul 02;22(7):e18831 [FREE Full text] [doi: 10.2196/18831] [Medline: 32540844]
51. Mahmoodi Khaniabadi S, Javadpour A, Gheisari M, Zhang W, Liu Y, Sangaiah AK. An intelligent sustainable efficient transmission internet protocol to switch between User Datagram Protocol and Transmission Control Protocol in IoT computing. Expert Syst 2022 Sep 07;40(5):e13129. [doi: 10.1111/exsy.13129]
52. Xiong S. The structural model of communication research base on the computer network: two Chinese translations of the word ‘communication’. Digital Publ Res (Chinese) 2023;2(3):58-66.
53. Xiong S. Introduction of Ludology [in Chinese]. Wuhan, Hubei, China: Huazhong University of Science and Technology Press; 2023.
54. Baishya K, Samalia HV. Extending unified theory of acceptance and use of technology with perceived monetary value for smartphone adoption at the bottom of the pyramid. Int J Inf Manag 2020 Apr;51:102036. [doi: 10.1016/j.ijinfomgt.2019.11.004]

55. Herrero Á, San Martín H. Explaining the adoption of social networks sites for sharing user-generated content: a revision of the UTAUT2. Comput Hum Behav 2017 Jun;71:209-217. [doi: 10.1016/j.chb.2017.02.007]
56. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
57. Notice on issuing the ethical review measures for life sciences and medical research involving humans: national health science and education fa [2023] no. 4. China Government. URL: https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm [accessed 2024-01-29]
58. Wu Q, Zhao X, Zhou L, Wang Y, Yang Y. Minimizing the influence of dynamic rumors based on community structure. IJCS 2019 Sep 02;3(3):303-314. [doi: 10.1108/ijcs-09-2019-0025]

Abbreviations

GM: genetically modified
HCW: health care worker
IP: Internet Protocol
SNS: social networking service
TCP: Transmission Control Protocol
UDP: User Datagram Protocol
UTAUT: unified theory of acceptance and use of technology

Edited by T Rashid Soron; submitted 09.01.23; peer-reviewed by M Jordan-Marsh, X Zhu, Y Wu; comments to author 24.02.23; revised version received 20.04.23; accepted 30.12.23; published 26.02.24.

Please cite as:
Xiong S, Zuo L, Chen Q, Zeliang Z, Nor Akmal Khalid M
A Serious Game ("Fight With Virus") for Preventing COVID-19 Health Rumors: Development and Experimental Study
JMIR Serious Games 2024;12:e45546
URL: https://games.jmir.org/2024/1/e45546
doi:10.2196/45546
PMID:38407954

Original Paper

Pressure Ulcer Management Virtual Reality Simulation (PU-VRSim) for Novice Nurses: Mixed Methods Study

Soo Youn Jung[1], MS; Kyoung Ja Moon[1], PhD

College of Nursing, Keimyung University, Daegu, Republic of Korea

Corresponding Author:
Kyoung Ja Moon, PhD
College of Nursing
Keimyung University
1095 Dalgubeol-daero, Dalseo-gu
Daegu, 42601
Republic of Korea
Phone: 82 53 258 7663
Fax: 82 53 258 7616
Email: kjmoon2150@gmail.com

Abstract

Background: Pressure ulcers (PUs) are a common and serious complication in patients who are immobile in health care settings. Nurses play a fundamental role in the prevention of PUs; however, novice nurses lack experience in clinical situations. Virtual reality (VR) is highly conducive to clinical- and procedure-focused training because it facilitates simulations.

Objective: We aimed to explore the feasibility of a novel PU management VR simulation (PU-VRSim) program using a head-mounted display for novice nurses and to investigate how different types of learning materials (ie, VR or a video-based lecture) impact learning outcomes and experiences.

Methods: PU-VRSim was created in the Unity 3D platform. This mixed methods pilot quasi-experimental study included 35 novice nurses categorized into the experimental (n=18) and control (n=17) groups. The PU-VRSim program was applied using VR in the experimental group, whereas the control group received a video-based lecture. The PU knowledge test, critical thinking disposition measurement tool, and Korean version of the General Self-Efficacy Scale were assessed before and after the intervention in both groups. After the intervention, the experimental group was further assessed using the Clinical Judgment Rubric and interviewed to evaluate their experience with PU-VRSim.

Results: The results compared before and after the intervention showed significant improvements in PU knowledge in both the experimental group (*P*=.001) and control group (*P*=.005). There were no significant differences in self-efficacy and critical thinking in either group. The experimental group scored a mean of 3.23 (SD 0.44) points (accomplished) on clinical judgment, assessed using a 4-point scale. The experimental group interviews revealed that the VR simulation was realistic and helpful for learning about PU management.

Conclusions: The results revealed that PU-VRSim could improve novice nurses' learning of PU management in realistic environments. Further studies using VR for clinical training are recommended for novice nurses.

(JMIR Serious Games 2024;12:e53165) doi:10.2196/53165

KEYWORDS

virtual reality; nursing; simulation; virtual training; pressure ulcer; simulation training; nurse; clinician; health care worker; ulcer; hospital; health care center; PU-VRSim; mixed methods study; health professional; medical education; training; games; gamification; learning; decubitus ulcer

Introduction

Pressure ulcers (PUs) refer to the skin damage caused by ischemia of the skin, subcutaneous fat, and muscles due to a continuous blood circulation disorder in a compressed body [1]. PUs result in a reduction of oxygen and nutrition delivered to the cells, which can contribute to the development of cancer and cardiovascular disease, ultimately amounting to high medical expenses [1]. The incidence of PUs is estimated at 12% in hospitals, representing a common but important health problem because it leads to high nursing burdens, increased medical costs, and mortality [2]. PUs are among the global health indicators and are included in the standard of nursing [3]. Most PUs are preventable by maintaining proper skin integration, and prevention is considered as important as treatment [3,4].

Nurses have a great responsibility in the well-being and safety of patients [5]. In addition, nurses are required to perform appropriate nursing for the prevention and management of PUs. However, they experience difficulties in performing clinical nursing and providing the necessary care to patients [6]. Novice nurses are those with less than 3 years of working experience based on the Benner novice-to-expert model; although they can recognize the basic order in nursing and take decisions, it is generally more difficult for novice nurses to establish priorities [7]. Particularly, novice nurses who have completed the regular curriculum do not have sufficient opportunities for practice during their training and thus experience difficulties in adapting to a new environment and changes in roles in the clinical field, leading to stress and anxiety [8,9]. Therefore, a program that can help novice nurses adapt to the clinical environment is needed.

Virtual reality (VR) is characterized by interaction, immersion, and imagination, and has been increasingly used in the curriculum for nursing with great potential for course development [10]. VR using a head-mounted display (HMD) provides the learning experience of communication between medical staff and patients, as well as simulations of standardized and controlled situations [11]. VR use has been easily accepted by learners in various medical environments and plays an essential role in improving their performance [11,12]. In nursing education, VR has been used in areas such as cardiopulmonary resuscitation, respiratory nursing, and delivery nursing, as well as for improving professional knowledge, clinical reasoning skills, and learning satisfaction [13,14]. In addition, as a learning method, VR meets the expectations and learning styles of the new generation of young learners [14,15].

In nursing education, teaching methods have shifted from traditional lecture-style education to simulation education [15]. Lecture-style education is effective in terms of knowledge transfer to novice nurses. However, there is a limit of this approach in improving nursing work skills in hospitals where various problems can occur [16]. Video-based education for novice nurses is a time-efficient and economically effective method owing to the heavy workload and lack of physical time; however, this format often lacks an appropriate feedback system [17]. Simulation education provides educational opportunities for clinical practice without putting patients or others at risk, and learners have the advantage of safely learning from experience [18]. VR is a representative technology for simulation education [15]. VR simulation can be used by novice nurses freely, which has been shown to improve their knowledge, critical thinking, and self-efficacy [13,18], thereby helping them transform into professional clinical nurses.

Simulation creates a learning environment in which learners can experience intervention and treatment in a safe manner, and various educational theories and structural models can be applied to achieve effective learning results [19].

The Analysis, Design, Development, Implementation, and Evaluation (ADDIE) instructional design model [20] is an effective and efficient development model based on the five steps of analysis, design, development, implementation, and evaluation. Kolb's experiential learning theory [21] states that learning is achieved through the process of "active experimentation," starting with a "concrete experience," "reflective observation," and "abstract conceptualization." Through the concrete experience of simulation, learners make reflective observations and abstract conceptualizations by trying and practicing new techniques in a safe environment, and they perform active experimentation to understand the patient's situation in an actual clinical environment and provide appropriate nursing practice.

This study aimed to develop a nursing PU management VR simulation program (PU-VRSim) and assess the feasibility of the novel virtual program for novice nurses. Toward this end, we applied the ADDIE instructional design model and Kolb's experiential learning theory. The first objective was to assess the feasibility of implementing PU-VRSim for nursing education on PUs. The second objective was to compare the effects of the VR program and video-based lectures on PU knowledge, self-efficacy, and critical thinking, and confirm the level of clinical judgment and experience of participants after undergoing PU-VRSim. The main research questions included: (1) Is implementing PU-VRSim for nursing education feasible? (2) What is the effect of PU-VRSim compared with that of video-based lectures? and (3) What are the participants' experiences with PU-VRSim?

Methods

Design

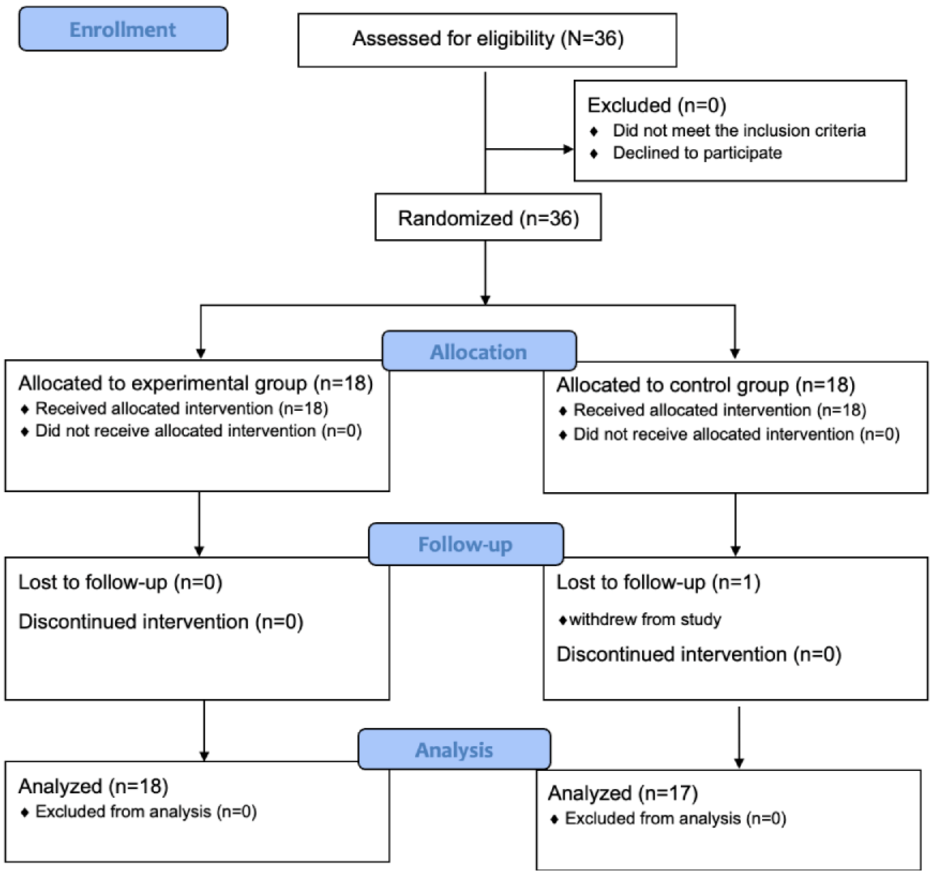
This study applied Kolb's experiential learning theory [21] based on the ADDIE model [20], which is a model of instructional design that was used to develop PU-VRSim for preventing and managing PUs using VR in a nursing education program. This was a mixed methods, pilot quasi-experimental study [22] including nurses with less than 2 years of clinical experience to confirm the effectiveness of PU-VRSim. PU-VRSim was created in the Unity 3D platform (Unity Technologies). Participants experienced the program through an HMD and hand controllers (HTC Corporation, VIVE pro).

Participants

For data collection, nurses with less than 2 years of clinical experience were notified of the purpose, period, conditions of participation, and benefits and disadvantages of participating in the study in the nurses' community bulletin boards. Recruitment for the preintervention survey, intervention, and postintervention survey was conducted through convenience sampling. Participants were categorized into the two groups based on the work schedule of the novice nurses, and participants were blinded to their group allocation. A total of 35 participants were recruited voluntarily from October 10 to December 31, 2022, with 18 assigned to the experimental group and 17 assigned to the control group from January 1 to March 31, 2023. In both groups, one researcher conducted a one-on-one survey and measured general characteristics, PU knowledge, critical thinking, and self-efficacy using preliminary questionnaires, which were sent to the two groups before implementing the program. Regarding the intervention, the experimental group participated in the PU-VRSim program in the simulation room, whereas the control group participated in a video-based lecture on the prevention and care of PUs. After the program, PU knowledge, critical thinking skills, and self-efficacy were measured in both groups. The effectiveness of the program was further assessed with participants in the experimental group via interviews and the Lasater Clinical Judgment Rubric (LCJR) (Figure 1).

The sample size required to compare variables between groups with the *t* test was calculated using the G*Power 3.1 program according to the method of Polit and Sherman [23], using a significance level of α=.05, effect size (*f*) of 0.80, and power (1–β) of .90 [24]. Considering that the sample size satisfying the above conditions was at least 16 people per group, 36 participants were selected, including 18 in the experimental group and 18 in the control group, prior to data collection. One participant in the control group dropped out of the study. Finally, 35 participants were included in the analysis.

Figure 1. Flow diagram of participants' enrollment, allocation, follow-up, and analysis.



Ethical Considerations

Data collection began after obtaining approval from the Institutional Review Board (40525-202204-HR-016-03) of Keimyung University in Daegu City for the protection of the research participants. The purpose of the study, procedures, guarantee of anonymity and confidentiality, and assurance that there are no consequences in case of withdrawal from the study were explained to the research participants, and they were allowed to respond to the questionnaire only when they agreed to participate in the research. The researchers conducted the preintervention survey, application of programs, and postintervention survey. All data collected during this study were anonymized. Participants were compensated for their contribution with a beverage coupon worth 10,000 KRW (~US $8) after the postintervention survey.

Instruments

PU Knowledge

The Pieper–Zulkowski pressure ulcer knowledge test (PZ-PUKT), a PU knowledge tool developed by Pieper and Zulkowski [25] and modified and supplemented by Park [26], was used in this study. The PZ-PUKT comprises 39 questions, including 19, 9, and 11 questions on PU stage confirmation, wound assessment, and dressing methods, respectively. Each question was answered "yes," "no," or "don't know," with 1 point for correct answers and 0 points for incorrect answers. The total score ranges from 0 to 39, with higher scores indicating greater knowledge of PUs. The Cronbach α value was 0.80 and 0.70 in the studies by Pieper and Zulkowski [25] and Park [26], respectively, and was 0.69 in our study.

Critical Thinking

The critical thinking disposition measurement tool developed by Yun [27] and modified and supplemented by Shin et al [28] was used for evaluating the impact of the intervention on critical thinking skills. This tool comprises 27 questions divided into 7 subdomains: intellectual passion/curiosity (5 questions), prudence (4 questions), confidence (4 questions), systemicity (3 questions), intellectual fairness (4 questions), healthy skepticism (4 questions), and objectivity (3 questions). Answers are rated on a scale of 1 point for "not so" to 5 points for "very much so"; a higher score indicates a stronger critical thinking disposition. The Cronbach α value was 0.84 in the studies of both Yun [27] and Shin et al [28] and was 0.83 in our study.

Self-Efficacy

The Korean version of the General Self-Efficacy Scale developed by Schwarzer and Jerusalem [29] and adapted by Schwarzer et al [30] was used to determine general self-efficacy. The Korean version of the General Self-Efficacy Scale comprises 10 questions rated on a 4-point Likert scale ranging from 10 to 40, with higher scores indicating higher self-efficacy. The Cronbach α value was 0.90 and 0.88 in the studies by Schwarzer and Jerusalem [29] and Schwarzer et al [30] and it was 0.86 in our study.

Clinical Judgment Rubric

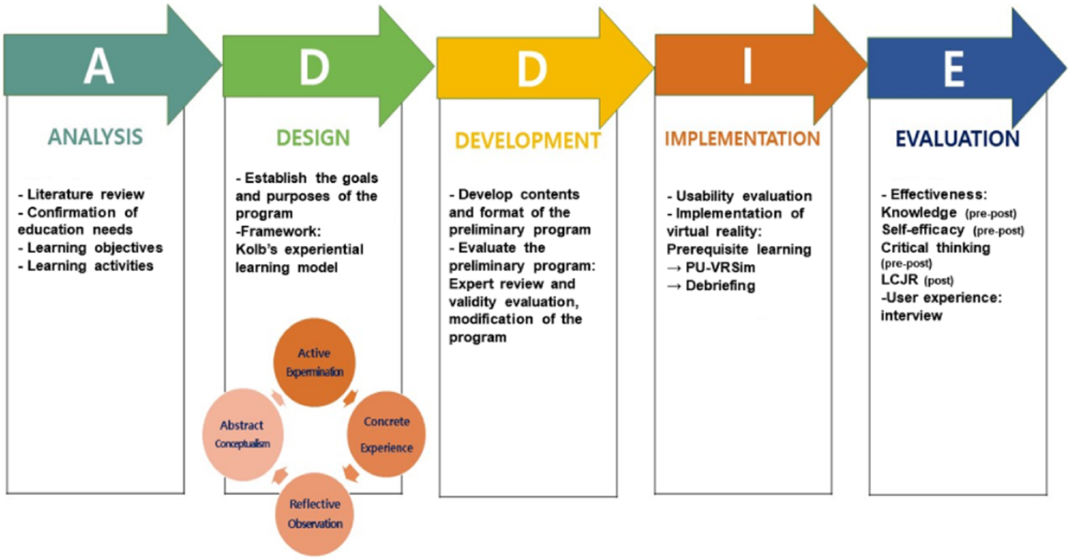
The LCJR, developed by Lasater [31], was used to evaluate the simulation experience. This rubric comprises 11 items based on the following four phases: noticing, interpreting, responding, and reflecting. The LCJR evaluates participants' performance as beginning (1 point), developing (2 points), accomplished (3 points), or exemplary (4 points). The total score ranges from 11 to 44, with a higher score indicating higher clinical judgment ability. The Cronbach α value was 0.83 in the study of Shin et al [32] and was 0.92 in our study.

Procedures

Development Overview

PU-VRSim was developed by applying Kolb's experiential learning theory to the ADDIE model (Figure 2).

Figure 2. PU-VRSim program development. LCJR: Lasater Clinical Judgment Rubric; PU-VRSim: pressure ulcer management virtual reality simulation.



Analysis Stage

The analysis stage identified learners' general and learning-related characteristics. Through a literature review [33], the importance of PU care, factors of PU occurrence, and prevention and management methods were confirmed, and the factors to be included in the development of the VR simulation were analyzed.

Design Stage

In the design stage, the teaching method for developing an effective educational program was determined. Kolb's experiential learning theory [21] was applied to the basic data collected during the analysis stage to determine the teaching method using the VR simulation (concrete experience) and debriefing (reflective observation). The design aimed to improve critical thinking, self-efficacy, and clinical judgment (abstract conceptualization). Using this approach, the learner would assist in performing PU prevention and nursing (active experimentation) well in actual patients.

Development Stage

In the development stage, a VR-based program was developed based on the educational topics selected in the analysis and design stages. A VR platform (Unity 3D, Unity Technologies) was constructed in collaboration with a professional company. A preliminary VR program was tested by five nurses with more than 5 years of clinical experience in scenarios and nursing care of patients with PUs. By checking and correcting errors in the VR program, operational problems were improved and addressed.

Implementation and Evaluation Stage

The implementation stage involved application and operation of the program completed in the development stage (Table 1).

Table 1. Core contents and images of the virtual reality system.

Main contents	Components	Image
Login	Registration of users	
Objectives	Identifying learning objectives: • Assess the degree of risk of developing PUs[a] in the patient. • Classify the PU stage of the patient. • Apply a proper dressing to the patient's PU. • Provide prevention education and care for PUs to the patient.	
Patient case	• Patient hospitalization history • Patient information: name, sex, age, diagnosis, past medical history, social history	
Patient information	Identifying data: vital signs, results of blood test, x-ray findings, physical exam, medication	
Nursing intervention		
Risk assessment for PU prevention	Braden scale score; a lower score indicates a higher risk of developing PUs.	
Assessment and evaluation of PUs	• Assessment of PUs: size and stage • Evaluation of PUs: writing a report	
Management of PUs	Dressing on PUs • Stage 1: film dressing • Stage 2: foam dressing	
Patient education	PU prevention and management education: • Skin care, urinary and fecal incontinence management • Support surfaces • Repositioning • Nutrition	
Final confirmation	• Running times • Feedback (recording)	

[a]PUs: Pressure ulcers

Assessments

Overview of Study Design

The program was used during the evaluation stage. PU knowledge, critical thinking, and self-efficacy in the experimental and control groups were measured once before the start of the study and then again after the program. For the experimental group only, assessment using the LCJR was performed after the program and the effect of PU-VRSim was evaluated through an interview.

Preintervention Survey

The preintervention survey of the experimental group was conducted from October 1 to November 30, 2022, and that of the control group was conducted from January 1 to March 1, 2023. After the participants signed a consent form to participate in the study, their PU knowledge, self-efficacy, and critical thinking were measured using structured questionnaires as described above.

Implementation

The experimental group received the VR simulation program, comprising a prebriefing session (15 minutes) where participants briefly learned about the definition, classification, prevention, and wound management of PUs. Participants were then exposed to PU-VRSim (10 minutes), including PU assessment, nursing care, and education to patients through VR. This was followed by a debriefing session (20 minutes), in which participants were assessed using the LCJR after the simulation.

The control group received a video-based lecture. The video format was selected to reduce the time burden on participants who work in shifts and to ensure safety from SARS-CoV-2 infection, according to the participants' hospital work. In total, 17 participants in the control group received lecture materials and a 30-minute video-based lecture on the definition, classification, prevention, and management of PUs.

Postintervention Survey

After the program, PU knowledge, self-efficacy, and critical thinking were assessed in both groups. The experimental group was further assessed using the LCJR and an interview was conducted to confirm their experience.

Interview

To discuss the experience of participating in PU-VRSim, which could not be verified using objective data, the participants were interviewed after the program. The interview included a self-introduction by the researcher and participant, recording of the interview, guaranteeing anonymity, and explaining that the research results were used only for research purposes and that the interviews were conducted with the participants' voluntary consent. One-on-one interviews were conducted in all cases in a quiet seminar room. Before conducting the interviews, the questions were drafted based on the purpose of the study and proceeded in the order of introduction, transition, and main questions, as shown in Textbox 1.

Textbox 1. Interview question structure.

Introduction question: Thank you for taking the time after work to participate in the virtual reality (VR) program. Can you briefly describe your feelings?

Transition question: Now, we would like to take the time to talk freely about the program's effectiveness and improvements.

Main question: What helped you with the program? What do you think about the content and methods of the VR program in which you participated? What do you need to improve or add to this program?

Data Analysis

The data collected in this study were analyzed using IBM SPSS 23.0, and a two-tailed test was performed at a significance level of .05. The normality of the dependent variable was verified using the Shapiro-Wilk test. The homogeneity of the data in the experimental and control groups was verified using the χ^2 test and independent *t* test. General characteristics and performance on the LCJR aspects of the participants were presented as means (SDs) and n (%), respectively. Wilcoxon signed rank and Mann-Whitney *U* tests were used to verify differences in PU knowledge, critical thinking, and self-efficacy between the experimental and control groups.

The data collected through the interviews were analyzed using an inductive approach, which is one of the content analysis methods suggested by Elo and Kyngäs [34]. For the data analysis, the researcher repeatedly read the transcripts of the interviews, interpreted the meaning of the key statements, and created categories by assigning titles. After data analysis, the authors discussed their interpretations to reach a consensus. Subsequently, the semantic units identified were grouped into higher-level categories, the properties were stated, and the keywords were derived by coding the contents accordingly.

Results

Feasibility of PU-VRSim

Our first objective was to assess the feasibility of implementing the PU-VRSim program for nursing education on PUs in the implementation and evaluation stages.

The general characteristics of the participants are presented in Table 2. The average age and work experience of the 35 novice nurses was 24.8 years and 14 months, respectively. The experimental group comprised 18 (100%) women, whereas the control group comprised 2 (12%) men and 15 (88%) women. We analyzed the homogeneity of the two groups in terms of general characteristics such as age, educational level, VR experience, and PU education experience; no significant difference was observed between the two groups (all *P*>.05) and thus homogeneity between the two groups was confirmed.

Table 2. Characteristics of participants (N=35).

Characteristics	Experimental group (n=18)	Control group (n=17)	χ^2 or t[a]	*P* value
Age (years), mean (SD)	24.11 (1.32)	25.53 (2.72)	–1.98	.06
Sex, n (%)			1.46	.16
Male	0 (0)	2 (12)		
Female	18 (100)	15 (88)		
Work experience (months), mean (SD)	14.06 (7.75)	14.65 (7.17)	–0.23	.82
Education, n (%)			1.852	.83
College	0 (0)	3 (18)		
University	18 (100)	14 (82)		
PU[b] education, n (%)			1.00	.33
Yes	17 (94)	17 (100)		
No	1 (6)	0 (0)		
VR[c] experience, n (%)			0.882	.38
Yes	3 (17)	7 (41)		
No	15 (83)	10 (59)		

[a]*df*=17 for the experimental group and 16 for the control group.

[b]PU: pressure ulcer.

[c]VR: virtual reality.

Effect of PU-VRSim on Outcomes

Our second objective addressed the effects of the VR intervention on PU knowledge, self-efficacy, critical thinking, and critical judgment.

As shown in Table 3, in the experimental group, the PU knowledge score increased by 2.88 points and the self-efficacy score increased by 0.56 points compared with those in the preintervention survey. In the control group, the PU knowledge score increased by 4.12 points, the critical thinking score increased by 4.0, and the self-efficacy score increased by 0.76 points. Each group showed significant improvements in PU knowledge after the intervention. However, there were no significant differences in critical thinking and self-efficacy in either group. There were no significant differences in the change in PU knowledge, critical thinking, and self-efficacy between the two groups.

The results for the clinical judgment assessment are summarized in Table 4. In the experimental group, after PU-VRSim, the overall clinical judgment of novice nurses was 3.23 points. When evaluated in the four phases to confirm whether all phases reached the level of “accomplished,” out of a total of 4 points, the mean scores for noticing, interpretation, responding, and reflecting were 3.27, 3.31, 3.32, and 2.91 points, respectively. The items “well-planned intervention/flexibility” and “skill proficiency” in the responding phase scored the highest, with 3.67 points, whereas “commitment improve” in the reflecting phase scored the lowest, with 2.78 points.

Table 3. Effect of the pressure ulcer management virtual reality simulation on outcomes.

Variable	Preintervention, mean (SD)	Postintervention, mean (SD)	*Z*	*P* value[a]	Difference		
					Mean (SD)	*Z*	*P* value[b]
Knowledge						–0.81	.42
Experimental (n=18)	24.39 (4.79)	27.28 (4.43)	–3.45	.001	2.88 (2.40)		
Control (n=17)	24.71 (5.42)	28.82 (3.41)	–2.78	.005	4.12 (4.74)		
Critical thinking						–1.51	.13
Experimental (n=18)	99.00 (10.70)	98.61 (8.51)	–0.13	.896	–0.39 (5.95)		
Control (n=17)	94.82 (7.24)	98.82 (9.36)	–1.40	.163	4.00 (9.01)		
Self-efficacy						–0.57	.57
Experimental (n=18)	29.11 (3.41)	29.67 (3.24)	–1.45	.148	0.56 (1.69)		
Control (n=17)	25.59 (4.18)	26.35 (3.98)	–1.16	.247	0.76 (3.40)		

[a]Wilcoxon signed rank test.
[b]Mann-Whitney *U* test.

Table 4. Clinical judgment scores.

Clinical judgment phase	Score, mean (SD)
Overall	3.23 (0.44)
Noticing	
Total	3.27 (0.45)
Focused observation	3.17 (0.51)
Recognizing deviations from expected patterns	3.06 (0.54)
Information seeking	3.61 (0.50)
Interpretation	
Total	3.31 (0.60)
Prioritizing data	3.28 (0.67)
Making sense of data	3.33 (0.59)
Responding	
Total	3.32 (0.46)
Clear communication	3.11 (0.58)
Well-planned intervention/flexibility	3.67 (0.49)
Being skillful	3.17 (0.62)
Reflecting	
Total	2.91 (0.62)
Evaluation/self-analysis	3.06 (0.64)
Commitment to improvement	2.78 (0.73)

Qualitative Outcomes of the PU-VRSim Experience Among Novice Nurses

Overview of Themes

For the third objective, the interviews were grouped into five main themes: (1) realistic VR scenarios, (2) helpfulness of VR learning, (3) usability, (4) satisfaction, and (5) limitations of VR equipment (Table 5).

Table 5. Qualitative outcomes of the pressure ulcer (PU) management virtual reality (VR) experience in novice nurses.

Themes and subthemes	Description
Realistic VR scenarios	
Real environment	There was a sense of presence in the hospital
Real experience	Actual practice in nursing on PUs
Helpfulness of VR learning	
Improving knowledge	Remember the concept and care of PUs well
Improving skills	Perform overall practice (assess-evaluate-manage) for PUs
Usability	
Safety	Safe from infection due to low contamination
Accessibility	Able to participate freely regardless of time and place
Satisfaction	
Joy and fun	It was more like playing games than learning
New experience	It was a new and interesting learning experience
Limitations of VR equipment	Inconvenient to use equipment

Theme 1: Realistic VR Scenarios

The participants gained practical experience through VR scenarios. The subtopics related to this were "real field" and "real experience."

> *When I experienced it myself, it was lively and felt like a real clinical environment.* [S1, S8, S10, S16]

> *It was realistic to be able to assess and manage wounds about which I learned from books using VR* [S4]

> *Although it is VR for pressure ulcer nursing, which is difficult to understand only through lectures, it was good to apply it as a direct action* [S11]

> *I only practiced with lying mannequins, but it felt more realistic when I experienced it with patients in VR like a real clinical environment.* [S13]

Theme 2: Helpfulness of VR Learning

The participants expressed that VR was helpful for learning. Subtopics related to this theme were "improvement of knowledge learning" and "improvement of skills."

> *I was able to learn the stages of PUs and the types of dressings.* [S6]

> *I was able to learn about PU care.* [S8]

> *When applying nursing care to patients with PUs through VR, it seems to be more memorable.* [S9, S15, S16]

> *I was able to confirm the PU classification concept.* [S12]

> *There were no bedsore patients in the ward where I worked, and through this virtual reality program, I was able to evaluate, intervene, and evaluate PUs.* [S2]

> *I was able to perform overall nursing activities for patients with PUs.* [S6]

Theme 3: Usability

The participants expressed the feeling of using VR learning as "safe" and "easy to access."

> *There were times when the mannequins were dirty in nursing practice, but it was nice that the virtual patients were not contaminated.* [S13]

> *It was nice to be able to participate without the burden of time and place.* [S5]

Theme 4: Satisfaction

The participants expressed their satisfaction of using VR learning as "pleasure," "fun," and "new experience." They showed interest in VR and experienced fun and enjoyment through learning.

> *It feels more like playing a game than learning something. Enjoyed it.* [S4]

> *It was my first time using virtual reality, and I was able to enjoy it. I want to try again.* [S5, S13, S17]

> *It was a new experience, and I enjoyed it.* [S8, S14, S17]

Theme 5: Limitation of VR Equipment

The participants expressed limitations in terms of the equipment used in the VR program.

> *When I put the equipment on my head, I took off my glasses and put it on, so it was difficult to read the text because my vision was not clear.* [S5]
>
> *It was a bit heavy to wear on my head.* [S8]
>
> *The preparation for running the program was complex and took a long time.* [S12]
>
> *It took a long time when the focus was not good, and the text looked blurry and the controller was not recognized well in VR.* [S13]

Discussion

Principal Findings

In this study, we developed PU-VRSim by applying the ADDIE model and Kolb's experiential learning theory [20,21]. PU-VRSim was designed as a PU prevention and nursing simulation program for novice nurses with less than 2 years of clinical experience. The participants of the PU-VRSim group showed significant improvements in PU knowledge. They reached the accomplished phase of clinical judgment. They commented that it was realistic and helpful for learning about PU management.

PU-VRSim was developed by applying an analysis-design-development-implementation-evaluation method according to the ADDIE model [20]. In the analysis stage, a literature review [33] confirmed that PU care was an important indicator of the quality of nursing services, which is becoming increasingly important [35]. PUs are caused by immobility, pressure, and friction. Factors to be included in education were analyzed by evaluating the methods for preventing and managing PUs through support surface management, position change, and dressing application to relieve pressure on the skin surface. Previous studies [36,37] have confirmed the improvement in PU knowledge and nursing performance of nurses through PU nursing education, thereby suggesting that continuous education on PUs for nurses is needed.

Kolb's experiential learning theory [21] applied at the design stage of PU-VRSim connects theory and practice in VR simulation. Through concrete experience and reflective observation, an abstract conceptualization of theories in realistic situations can help to acquire the knowledge and skills that can be used in real situations. Kolb's theory has also been applied in simulation education in various health fields [38], and VR provides learners with experience-based learning in a real environment by which the learners make decisions and take appropriate actions in real situations [10]. Through PU-VRSim, novice nurses freely apply the theoretical knowledge acquired through existing knowledge and prior learning materials to the process of solving problems encountered by patients in a safe virtual clinical environment. Ultimately, positive results can be expected by applying the improved nursing capabilities in actual clinical trials.

After novice nurses underwent the program, they showed improvement in PU knowledge and reached the "accomplished" stage of clinical judgment. PU knowledge scores increased on average in both the experimental and control groups after the educational program. This shows that the effect of knowledge transfer [16] can be confirmed via both traditional teaching and lectures and with the new VR simulation method. According to Kolb's experiential learning theory, during knowledge transfer via VR, learners can improve their knowledge by reapplying it through concrete experiences and reflections, as well as learn how to utilize what they have learned and gain new knowledge. PU knowledge is the basis of PU care, and professional nursing can be performed through critical thinking and improvement of clinical performance skills. Furthermore, clinical judgment is a particularly important skill in nursing that has also received recent attention. VR has a positive impact on clinical judgment in nursing education [39]. Therefore, positive effects and acquisition of new skills can be expected if field-tailored simulation [40] is applied to nurses to reproduce clinical situations.

Interview contents were analyzed to confirm the experiences of the novice nurses participating in the PU-VRSim program. The analysis revealed that the realistic scenarios of PU-VRSim help in learning, with usability of safety, easy accessibility, and satisfaction being expressed as positive experiences. However, inconveniences in using the equipment to implement VR programs were expressed as negative experiences. This is consistent with the findings of Adhikari et al [41] on the experience of VR programs in terms of acceptability, applicability, areas of improvement, and limitations. VR can be safely and repeatedly applied in situations that can be dangerous to patients; however, it is expensive and has usage limitations [42]. VR was deemed to be a safe and effective educational method for use during the COVID-19 pandemic [43]. The development of a VR program that reduces the inconvenience and cost burden of equipment is expected to increase the use of VR in nursing education.

Our participants confirmed that PU-VRSim was helpful for learning because it could be used repeatedly to access disease-focused nursing problems. There is a need for education about various clinical situations in which nurses can apply nursing interventions according to the situation and an overall assessment of the patient [44,45]; PU-VRSim reflects the clinical situation of PUs and requires nursing education. In addition, it was confirmed as a positive experience, suggesting that improvement in nursing knowledge and clinical performance ability as well as repetitive learning are possible through the promotion of spontaneous thinking and immediate feedback of learners, which were evaluated as advantages of simulation in previous studies [46,47]. These results confirm the possibility of using PU-VRSim as an educational program in clinical practice.

In nursing practice using mannequins during the COVID-19 pandemic, the participants expressed concerns about infection via contamination of the mannequins from multiple contacts. Using VR, they felt safe as the risk of infection could be avoided. In previous studies, VR simulation was suggested and used as a nonface-to-face practice method when clinical practice was not possible due to the prevalence of COVID-19 [43,48]. In addition, participants did not feel the burden of time and space when participating in VR education. This is an advantage of VR, in which one can experience the actual medical field using only computers and equipment. Furthermore, individual learning is possible; therefore, it can provide optimal learning to individual learners and help them overcome obstacles in the physical environment [49]. VR may be an appropriate training method for shift workers, because nurses who work shifts can access the education without experiencing a time burden.

Novice nurses in this study regarded VR education as a new experience and evaluated it as enjoyable. A VR learning environment enhances immersion and activates learners' imagination to simulate the real world [50]. The VR program is a teaching method that incorporates the latest technology and meets the learning needs of a new generation. Educational programs are being developed on various topics for nursing students and novice nurses, and the effects of enjoyment and fun have been confirmed [13]. Learning satisfaction through the enjoyable and fun VR improves learners' learning motivation and confidence, and they experience reduced fear in real situations [14]. Because enjoyment and fun in learning are factors that stimulate learning motivation and interest, gamification can be applied when developing programs so that learners can enjoy various experiences in the virtual world.

When implementing the VR program, participants had difficulty in using an HMD; in particular, participants wearing glasses experienced inconvenience when wearing the device along with their glasses. As in previous studies, most of the participants experienced technical difficulties [41]. A VR program is typically executed using a computer program, an HMD, and a controller; however, the HMD and controller devices do not recognize the participants' fine movements, making it difficult to proceed [42]. In the future, the development of a convenient version of the HMD, with a clear field of view, ease of wearing, and usability, may lead to an increase in the use of VR education.

PUs are common health problems in hospitals, and novice nurses experience difficulties in treating PUs. Education of nurses has been regarded as an integral component of PU prevention [51]. VR is an ideal educational technology, and the number of educational programs applying VR in nursing education has been increasing recently [52]. For the VR education program, we confirmed the improvement in knowledge of the participants through the experience in prevention and nursing interventions for patients with PUs. Improvement in clinical performance can be expected with improved knowledge. In addition, the novice nurses in this study expressed satisfaction with VR education as a new experience and a safe learning method. Considering the limitations of VR equipment, it is necessary to develop and utilize a popular simulation program that is more user-friendly and can be manipulated easily. Based on this study, we suggest the development of VR nursing education programs focusing on the educational needs of novice nurses and including new technology such as artificial intelligence with the development of technology.

Conclusion

The PU-VRSim program developed in this study was found to be effective in improving novice nurses' knowledge of PUs and was positively evaluated as a pleasant experience conducive to learning in an actual hospital-like environment. Therefore, PU-VRSim can be used as an effective educational method for novice nurses, as well as for nursing students and clinical nurses. In addition, a synergistic effect can be expected when the content used incorporates various software programs, including VR simulation programs.

Limitations exist in understanding and generalizing the effects of nonrandomized control-group experiments targeting novice nurses. To supplement this, we propose a follow-up study that applies the PU-VRSim program to nursing students and clinical nurses, as well as a randomized control group experimental study of novice nurses. All participants in the experimental group were women; therefore, we propose a further study with a more heterogeneous group of participants by including male and female nurses. In addition, we suggest the development of a field-tailored VR simulation for health professionals, including novice nurses, and study of its educational effect. Finally, developing a program for VR simulation is expensive and wearing an HMD when implementing the program is uncomfortable. We propose the development of software and VR simulations using technology such as smartphone apps, which are inexpensive, comfortable, and easy to use. In conclusion, we propose the continuous development and improvement of VR nursing education programs for novice nurses applying new technologies.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NFR) funded by the Ministry of Education (2022R1I1A3066118).

Conflicts of Interest

None declared.

References

1. European Pressure Ulcer Advisory Panel (EPUAP), National Pressure Injury Advisory Panel (NPIAP), Pan Pacific Pressure Injury Alliance (PPPIA). Prevention and treatment of pressure ulcers/injuries: quick reference guide. 2019. URL: https://www.hauoratairawhiti.org.nz/assets/Documents/Pressure-injuries/Quick-reference-guide-PPPIA-2019.pdf [accessed 2024-05-16]

2. Borojeny LA, Albatineh A, Hasanpour Dehkordi A, Ghanei Gheshlagh R. The incidence of pressure ulcers and its associations in different wards of the hospital: a systematic review and meta-analysis. Int J Prev Med 2020;11:171 [FREE Full text] [doi: 10.4103/ijpvm.IJPVM_182_19] [Medline: 33312480]
3. Hart S, Bergquist S, Gajewski B, Dunton N. Reliability testing of the National Database of Nursing Quality Indicators pressure ulcer indicator. J Nurs Care Qual 2006;21(3):256-265. [doi: 10.1097/00001786-200607000-00011] [Medline: 16816607]
4. Moore Z, Patton D. Risk assessment tools for the prevention of pressure ulcers. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD006471 [FREE Full text] [doi: 10.1002/14651858.CD006471.pub4] [Medline: 30702158]
5. Berkow S, Virkstis K, Stewart J, Conway L. Assessing new graduate nurse performance. J Nurs Adm 2008 Nov;38(11):468-474. [doi: 10.1097/01.NNA.0000339477.50219.06] [Medline: 18997551]
6. Saintsing D, Gibson L, Pennington A. The novice nurse and clinical decision-making: how to avoid errors. J Nurs Manag 2011 Apr;19(3):354-359. [doi: 10.1111/j.1365-2834.2011.01248.x] [Medline: 21507106]
7. Benner P. From Novice to Expert: Excellence and Power in Clinical Nursing Practice. Menlo Park, CA: Addison-Wesley Publishing Company; 1984.
8. Blomberg K, Isaksson A, Allvin R, Bisholt B, Ewertsson M, Kullén Engström A, et al. Work stress among newly graduated nurses in relation to workplace and clinical group supervision. J Nurs Manag 2016 Jan 25;24(1):80-87. [doi: 10.1111/jonm.12274] [Medline: 25421164]
9. Powers K, Herron EK, Pagel J. Nurse preceptor role in new graduate nurses' transition to practice. Dimens Crit Care Nurs 2019;38(3):131-136. [doi: 10.1097/DCC.0000000000000354] [Medline: 30946120]
10. Pottle J. Virtual reality and the transformation of medical education. Future Healthc J 2019 Oct;6(3):181-185 [FREE Full text] [doi: 10.7861/fhj.2019-0036] [Medline: 31660522]
11. Keller MS, Park HJ, Cunningham ME, Fouladian JE, Chen M, Spiegel BMR. Public perceptions regarding use of virtual reality in health care: a social media content analysis using Facebook. J Med Internet Res 2017 Dec 19;19(12):e419 [FREE Full text] [doi: 10.2196/jmir.7467] [Medline: 29258975]
12. Samadbeik M, Yaaghobi D, Bastani P, Abhari S, Rezaee R, Garavand A. The applications of virtual reality technology in medical groups teaching. J Adv Med Educ Prof 2018 Jul;6(3):123-129 [FREE Full text] [Medline: 30013996]
13. Chen F, Leng Y, Ge J, Wang D, Li C, Chen B, et al. Effectiveness of virtual reality in nursing education: meta-analysis. J Med Internet Res 2020 Sep 15;22(9):e18290 [FREE Full text] [doi: 10.2196/18290] [Medline: 32930664]
14. Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical virtual simulation in nursing education: randomized controlled trial. J Med Internet Res 2019 Mar 18;21(3):e11529 [FREE Full text] [doi: 10.2196/11529] [Medline: 30882355]
15. Chen C, Chen S. Effects of 2D wedge design as a wayfinding facilitator in a 3D virtual environment. J Soc Info Display 2015 Jun 30;23(1):27-35. [doi: 10.1002/jsid.304]
16. Mansoory MS, Khazaei MR, Azizi SM, Niromand E. Comparison of the effectiveness of lecture instruction and virtual reality-based serious gaming instruction on the medical students' learning outcome about approach to coma. BMC Med Educ 2021 Jun 15;21(1):347 [FREE Full text] [doi: 10.1186/s12909-021-02771-z] [Medline: 34130676]
17. Kim Y, Son E. A study on satisfaction of new nurse orientation program applying CIPP Evaluation Model focusing on affiliated hospitals of an educational foundation in South Korea. J Korea Acad Indust Coop Soc 2017;18(9):226-235. [doi: 10.5762/KAIS.2017.18.9.226]
18. Chang S, Kwon E, Kwon YO, Kwon HK. The effects of simulation training for new graduate critical care nurses on knowledge, self-efficacy, and performance ability of emergency situations at intensive care unit. J Korean Acad Adult Nurs 2010:375-383.
19. Taylor L. Educational theories and instructional design models. Their place in simulation. Nurs Educ Res Southern Health 2004:1-6.
20. Richey R, Seels B. Defining a field: a case study of the development of the 1994 definition of instructional technology. Educ Media Technol Yearbook 1994;20:2-17.
21. Kolbe D. Experiential learning: Experience as the Source of Learning and Development. Upper Saddle River, NJ: FT Press; 1983.
22. Almeida F. Strategies to perform a mixed methods study. Eur J Educ Stud 2018;5(1):137-151. [doi: 10.5281/zenodo.1406214]
23. Polit DF, Sherman RE. Statistical power in nursing research. Nurs Res 1990;39(6):365-370. [doi: 10.1097/00006199-199011000-00010]
24. Ahn MK, Lee CM. Development and effects of head-mounted display-based home-visits virtual reality simulation program for nursing students. J Korean Acad Nurs 2021 Aug;51(4):465-477. [doi: 10.4040/jkan.21051] [Medline: 34497255]
25. Pieper B, Zulkowski K. The Pieper-Zulkowski pressure ulcer knowledge test. Adv Skin Wound Care 2014 Sep;27(9):413-419. [doi: 10.1097/01.ASW.0000453210.21330.00] [Medline: 25133343]
26. Park O. Effectiveness of Case-Centered Education Program based on Nursing Protocol for Pressure Injury Stages. Dissertation. Seoul, Korea: Korea University; 2018.
27. Yoon J. Development of an Instrument for the Measurement of Critical Thinking Disposition: In Nursing. Dissertation. Seoul, Korea: The Catholic University of Korea; 2004.

28. Shin H, Park CG, Kim H. Validation of Yoon's Critical Thinking Disposition Instrument. Asian Nurs Res (Korean Soc Nurs Sci) 2015 Dec;9(4):342-348 [FREE Full text] [doi: 10.1016/j.anr.2015.10.004] [Medline: 26724244]
29. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio Causal and Control Beliefs. London: GL Assessment; 1995:35-37.
30. Schwarzer R, Bäßler J, Kwiatek P, Schröder K, Zhang JX. The assessment of optimistic self‐beliefs: comparison of the German, Spanish, and Chinese versions of the General Self‐efficacy Scale. Appl Psychol 2008 Jan 22;46(1):69-88. [doi: 10.1111/j.1464-0597.1997.tb01096.x]
31. Lasater K. Clinical judgment development: using simulation to create an assessment rubric. J Nurs Educ 2007 Nov;46(11):496-503. [doi: 10.3928/01484834-20071101-04] [Medline: 18019107]
32. Shin H, Park CG, Shim K. The Korean version of the Lasater Clinical Judgment Rubric: a validation study. Nurse Educ Today 2015 Jan;35(1):68-72. [doi: 10.1016/j.nedt.2014.06.009] [Medline: 25034956]
33. Jung SY, Park M, Moon KJ. Effectiveness of devices for prevention and treatment of pressure ulcers: a scoping review. Korean J Adult Nurs 2022;34(2):123. [doi: 10.7475/kjan.2022.34.2.123]
34. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr 18;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
35. Jo E, Kim H, Lee S. A study on the knowledge of nurses, performance and preventive practice of pressure ulcer among nurses in long-term care hospital. J Korea Content Assoc 2015 Aug 28;15(8):356-365. [doi: 10.5392/jkca.2015.15.08.356]
36. Beeckman D, Schoonhoven L, Boucqué H, Van Maele G, Defloor T. Pressure ulcers: e-learning to improve classification by nurses and nursing students. J Clin Nurs 2008 Jul 12;17(13):1697-1707. [doi: 10.1111/j.1365-2702.2007.02200.x] [Medline: 18592624]
37. Sardari M, Esmaeili R, Ravesh N, Nasiri M. The impact of pressure ulcer training program on nurses' performance. J Adv Pharmacy Educ 2019;3:145-149 [FREE Full text]
38. Wijnen-Meijer M, Brandhuber T, Schneider A, Berberat P. Implementing Kolb´s experiential learning cycle by linking real experience, case-based discussion and simulation. J Med Educ Curric Dev 2022 May 12;9:23821205221091511 [FREE Full text] [doi: 10.1177/23821205221091511] [Medline: 35592131]
39. Vihos J, Chute A, Carlson S, Buro K, Velupillai N, Currie T. Virtual health assessment laboratory course delivery and nursing student clinical judgment: a mixed-methods exploratory study. Nurse Educ 2022;47(3):E51-E56. [doi: 10.1097/NNE.0000000000001173] [Medline: 35148304]
40. Chernikova O, Heitzmann N, Stadler M, Holzberger D, Seidel T, Fischer F. Simulation-based learning in higher education: a meta-analysis. Rev Educ Res 2020 Jun 15;90(4):499-541. [doi: 10.3102/0034654320933544]
41. Adhikari R, Kydonaki C, Lawrie J, O'Reilly M, Ballantyne B, Whitehorn J, et al. A mixed-methods feasibility study to assess the acceptability and applicability of immersive virtual reality sepsis game as an adjunct to nursing education. Nurse Educ Today 2021 Aug;103:104944. [doi: 10.1016/j.nedt.2021.104944] [Medline: 34015677]
42. Kavanagh S, Luxton-Reilly A, Wuensche B, Plimmer B. A systematic review of virtual reality in education. Themes Sci Technol Educ 2017;10(2):85-119.
43. Sadek O, Baldwin F, Gray R, Khayyat N, Fotis T. Impact of virtual and augmented reality on quality of medical education during the COVID-19 pandemic: a systematic review. J Grad Med Educ 2023 Jun;15(3):328-338 [FREE Full text] [doi: 10.4300/JGME-D-22-00594.1] [Medline: 37363680]
44. Lee SH, Kim J, Jung D, Kang SJ. Educational needs for new graduate nurses in Korea. Nurse Educ Pract 2019 Jan;34:167-172. [doi: 10.1016/j.nepr.2018.11.021] [Medline: 30553233]
45. Kim S, Choi Y. Nursing competency and educational needs for clinical practice of Korean nurses. Nurse Educ Pract 2019 Jan;34:43-47. [doi: 10.1016/j.nepr.2018.11.002] [Medline: 30445293]
46. Kim Y, Kim WJ, Min HY. Nursing students' experiences in virtual simulation practice. J Korean Acad Soc Nurs Educ 2020 May;26(2):198-207. [doi: 10.5977/jkasne.2020.26.2.198]
47. Kim M, Kim S, Lee WS. Effects of a virtual reality simulation and a blended simulation of care for pediatric patient with asthma. Child Health Nurs Res 2019 Oct 31;25(4):496-506 [FREE Full text] [doi: 10.4094/chnr.2019.25.4.496] [Medline: 35004441]
48. Leaver C, Stanley J, Goodwin Veenema T. Impact of the COVID-19 pandemic on the future of nursing education. Acad Med 2022 Mar 01;97(3S):S82-S89 [FREE Full text] [doi: 10.1097/ACM.0000000000004528] [Medline: 34789661]
49. Jumani A, Siddique W, Laghari A, Abro A, Khan AA. Virtual reality and augmented reality for education. In: Tiwari R, Duhan N, Mittal M, Anand A, Khan MA, editors. Multimedia Computing Systems and Virtual Reality. Boca Raton, FL: CRC Press; 2022:189-210.
50. Huang H, Rauch U, Liaw S. Investigating learners' attitudes toward virtual reality learning environments: based on a constructivist approach. Comput Educ 2010 Nov;55(3):1171-1182. [doi: 10.1016/j.compedu.2010.05.014]
51. Porter-Armstrong A, Moore Z, Bradbury I, McDonough S. Education of healthcare professionals for preventing pressure ulcers. Cochrane Database Syst Rev 2018 May 25;5(5):CD011620 [FREE Full text] [doi: 10.1002/14651858.CD011620.pub2] [Medline: 29800486]
52. Pedram S, Kennedy G, Sanzone S. Toward the validation of VR-HMDs for medical education: a systematic literature review. Virtual Real 2023 May 13;27(3):1-26 [FREE Full text] [doi: 10.1007/s10055-023-00802-2] [Medline: 37360815]

Abbreviations

ADDIE: Analysis, Design, Development, Implementation, and Evaluation
HMD: head-mounted device
LCJR: Lasater Clinical Judgment Rubric
PU: pressure ulcer
PU-VRSim: pressure ulcer management virtual reality simulation
PZ-PUKT: Pieper-Zulkowski pressure ulcer knowledge test
VR: virtual reality

Edited by A McDaniel; submitted 27.09.23; peer-reviewed by D Lerner, M Brown, V Girishan Prabhu; comments to author 01.02.24; revised version received 26.03.24; accepted 15.04.24; published 24.06.24.

Please cite as:
Jung SY, Moon KJ
Pressure Ulcer Management Virtual Reality Simulation (PU-VRSim) for Novice Nurses: Mixed Methods Study
JMIR Serious Games 2024;12:e53165
URL: https://games.jmir.org/2024/1/e53165
doi:10.2196/53165
PMID:38913417

Review

Use and Design of Virtual Reality–Supported Learning Scenarios in the Vocational Qualification of Nursing Professionals: Scoping Review

Jenny-Victoria Steindorff[1*], BSc, MSc; Lisa-Marie Redlich[2*], BSc; Denny Paulicke[1,3*], Prof Dr; Patrick Jahn[1*], Prof Dr

[1]Health Service Research Working Group, Acute Care, Department of Internal Medicine, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[2]Interdisciplinary Center for Health Sciences, Institute of Health and Nursing Science, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[3]Department of Medical Pedagogy, Akkon University of Human Sciences, Berlin, Germany

[*]all authors contributed equally

Corresponding Author:
Jenny-Victoria Steindorff, BSc, MSc
Health Service Research Working Group, Acute Care, Department of Internal Medicine
Faculty of Medicine, University Medicine Halle (Saale)
Martin-Luther-University Halle-Wittenberg
Magdeburger Straße 12
Halle (Saale), 06112
Germany
Phone: 49 345 557 4164
Email: jenny.steindorff@uk-halle.de

Abstract

Background: Numerous reviews advocate using virtual reality (VR) in educational contexts. This medium allows learners to test experiences in realistic environments. Virtually supported scenarios offer a safe and motivating way to explore, practice, and consolidate nursing skills in rare and critical nursing tasks. This is also cited as one of the reasons why VR can significantly increase the knowledge acquisition of nursing students. Nevertheless, studies are limited in their significance owing to the chosen design. Despite great interest, this results in a low level of confidence in VR as a curricular teaching method for nursing education. Therefore, defining concrete design and didactic-methodological parameters that support teachers in the use and implementation of VR is more relevant.

Objective: This scoping review aims to provide an overview of significant design aspects for VR scenario conception and its transfer to generalist nursing education to generate value for the development of teaching scenarios and their sustainable implementation in teaching.

Methods: A comprehensive literature search was performed using the MEDLINE (via PubMed) and CINAHL databases, and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist was applied. The search was conducted from May to July 2022, using a specific search principle corresponding to the focus and the growing study corpus. A previously defined "population, concept, and context" scheme was employed as the basis for the double-blind review of all relevant international German and English publications released up to May 1, 2022.

Results: In accordance with the predefined selection procedure, 22 publications were identified. The identified aspects aided in the development of design, didactic, and research recommendations. The intuitive operation of realistically designed VR scenarios, which are standardized, reliable, and modifiable, as well as clear instructions and specific multimodal feedback functions were described positively. The same applied to the linear structure of the sequences with graduated demands and high image quality for increased immersion with low sensory overload. Changes in perspectives, multiuser options, dialogs, and recording functions can contribute to an interactive care practice. On the research side, it is advisable to define VR terminologies. In addition to considering larger samples, varying settings, and financial issues, it is recommended to conduct long-term studies on knowledge acquisition or improved patient outcomes.

Conclusions: VR scenarios offer high potential in the context of nursing education if teachers and learners develop them co-creatively according to design features and implement them by means of a well-conceived concept. VR enables trainees to develop practical skills continuously in a standardized way. In addition, its deployment supports the sensitization of trainees to digital nursing technologies and the expansion of their digital skills in a practical setting. Furthermore, it allows sustainability issues to be addressed.

(JMIR Serious Games 2024;12:e53356) doi:10.2196/53356

KEYWORDS

scoping review; generalist nursing education; digital teaching formats; virtual reality application; co-creation

Introduction

Background

Germany and the German health care system are facing enormous challenges. In addition to an increase in the need for care, demographic change is also leading to a blatant and increasing shortage of skilled nursing staff. Thus, the care gap in Germany is growing in all areas of care, and the need for nursing staff will rise alarmingly to 500,000 by 2035 [1]. However, in addition to the quantitative needs, the complexity of care is also increasing. A multimodal approach that considers other solutions in addition to human resources is required to counter the care crisis effectively [2]. Increasing technologization and digitalization in the health care sector can not only provide relief and additional security, but also strengthen the availability of current and person-centered (specialist) knowledge and skills in training and further education [3]. Accordingly, the educational pathways for the health care system in Germany are also changing. On the one hand, the academization of nursing education is being discussed and implemented in model study programs, while on the other hand, the new curricular orientation provides for generalist nursing education [4].

One recommendation therefore advocates transformative learning approaches that enable trainees and existing nurses to deal constructively and reflectively with the changing processes of an increasingly complex care reality.

Virtual Reality as a Transformative Learning Approach in Nursing Education

Gradually, more educational institutions of health care are making use of virtually supported teaching-learning scenarios [4], as they represent a suitable medium to train or support the skills of health care professionals.

One of the potentials is attributed to the immersive effect [5]. With the help of virtual reality (VR), users immerse themselves in a computer-generated synthetic environment [6] and perceive it via the senses of sight and hearing and increasingly via movement and touch as well. Therefore, immersion refers to the objective degree of sensory reality fidelity. According to Milgram et al [7], the degree of reality representation, which originally referred primarily to visual representation and has since been extended to haptic or acoustic experiences, can be classified on a “virtuality continuum” between the extremes of reality and VR.

As there is no standardized usage of the term VR yet, we applied the description of immersive virtual reality simulation (iVRS) according to Shorey et al [8] as a working definition, which was decisive for determining the study inclusion and exclusion criteria as follows: “The virtual world [also “virtual reality,” authors’ note] is a 3D computer environment that provides users with interactive experiences of an alternate reality in which they are avatars who can move, sense, touch, and act upon simulated objects that appear real [9]. There are 2 variations of virtual worlds, namely, desktop virtual reality simulation (dVRS) and iVRS [10]. dVRS, also known as non-iVRS, is where users interact with an environment displayed on a computer monitor using a mouse, keyboard, touchscreen, or joystick [11]. In contrast, iVRS provides a complete simulated environment where the user is equipped with several sensory output devices such as a head-mounted device, stereoscopic unit, audio device, and haptic device [10]. It involves a higher degree of interactivity compared to dVRS — by blocking out many visual elements of the real-world environment and inducing sensory stimuli that correspond with the virtual environment, it enables the user to immerse in the virtual environment [12].”

Nevertheless, both these varieties use the principles of interaction and user participation in addition to the characteristic of graduated immersion [13]. This characteristic enables nursing trainees to experience both routines and the complexity of rare or dangerous care tasks in an activating but safe and motivating environment [14].

In the last decade, various international studies have investigated the application of VR for educational purposes in nursing. As an interesting complement to traditional teaching methods, the use of VR to improve the teaching of basic nursing skills, communication, or teamwork [15] has increased. Here, above all, the possibility of conveying abstract and complicated content is used, as one’s actions and their effects are brought into focus and the learning content is perceived as more attractive [13] and is addressed via several sensory channels in parallel [16]. Beyond this, the procedure for learning and acquiring skills and competencies, which trainees can repeat as often as needed, promotes neuronal linkage [17,18] and the resulting confidence in action. In following this approach, ways of translating theoretical knowledge into practical skills and abilities emerge [19]. However, learners and teachers have described this theory-practice transfer as critical and inadequate if only conventional teaching methods are used [13]. This can lead to not only inadequate care but also dropouts from training as trainees demonstrate an excessive demand for the learning content and its transfer to concrete practical requirements [20], especially since the number and regularity of patient contacts during training are often insufficient.

Accordingly, technology-supported teaching-learning arrangements can provide multiple services as follows:

1. They can take up the changing range of professional tasks in nursing, depict them, and teach the competencies required for this in a situational and interactive way in a safe learning setting or support the acquisition by opening up opportunities for self-observation and self-reflection [5], particularly for complex action situations that occur rather rarely in care practice and cannot be guaranteed or practiced in the training phases.
2. They can increase the intrinsic motivation to learn and the attractiveness of training [18] and can make it more effective [21].
3. They can indirectly fulfill the demand for the inclusion of digital-related competence requirements in curricula [22].

In order to establish a connection between educational and care contexts and thus provide educational value, digital technologies should be used as a learning medium in a reflected and justified manner. This makes it more relevant to define concrete design and didactic-methodological parameters that support teachers in the use and implementation of VR in their teaching.

Despite the increasing number of publications on VR as a learning medium in the educational context of health care, there is still a lack of a merger between best practice experiences and recommendations for targeted use and specific design in generalist nursing education. To our knowledge, this didactic-methodological approach to VR-supported nursing education has not been applied yet. Based on this, our scoping review is intended to contribute to showing the potentials and indications of VR as a specifically selected and supplementary teaching-learning medium and to reveal the needs of this distinctive target group for an efficient design.

Study Objectives

The aim of this comprehensive literature review is to compile the findings and best practice examples of projects on VR-supported educational processes in nursing that have already been completed or are still in progress. With the help of this exploratory overview of the currently available evidence, it should be possible to make statements and recommendations as to which design aspects are relevant for the conception and use of didactically and methodologically significant virtually supported teaching-learning scenarios in the professional qualification of nursing specialists and to what extent these can be transferred to basic nursing training.

Methods

Overview

This scoping review, based on the JBI methodology [23], has obtained and mapped an overview of previous and current international research projects [24], and it is as broad and in-depth as possible [23]. With the help of the procedure described by Arksey and O'Malley [25], which comprises the steps of searching for and identifying relevant studies; selecting them; presenting the data; and compiling, summarizing, and reporting the results, it is possible to both make use of the research results already generated and identify the research gaps that still exist [25].

Search Strategy

From May to July 2022, a comprehensive search was conducted in 2 specialist databases (MEDLINE via PubMed and CINAHL via EBSCO) according to predefined inclusion and exclusion criteria, which are presented in Textbox 1. The search strings for the literature search in the databases are presented in Multimedia Appendix 1. In addition to publications identified in the reference lists that appeared to be suitable according to the keywords and were available as full texts, grey literature from other databases and websites available online was also taken into account and included in the screening of abstracts and full texts.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

Publications

- All publication types.
- Publications until May 5, 2022.
- Available full text (author request if applicable).
- German or English publications.
- International studies.

Population

- Trainees, students, and teachers in nursing care.
- Working nursing professionals participating in continuing education programs.

Concept

- Virtual reality (VR) applications in the professional qualification of nursing staff:
  - VR
  - Immersive applications
  - Use of a head-mounted display
- Outcome:
  - Effectiveness
  - Acceptance
  - Trust
  - Usability
  - Design features

Context

- Basic training as a nursing professional.
- Basic studies to become a nursing professional.
- Continuing education and training for nursing professionals.
- Interprofessional teaching-learning settings in which future nursing professionals also participate.

Exclusion criteria

Publications

- Publications after May 5, 2022.
- Full text subject to a fee.
- Non-German or non-English publications.

Population

- Exclusively students of human, dental, or veterinary medicine.
- Exclusively practitioners of human, dental, or veterinary medicine.
- Exclusively trainees of other health professions.
- Exclusively practitioners of other health professions.

Concept

- Other simulation-based teaching-learning forms without VR or immersive approaches.
- Programming aspects of VR applications only.

Context

- VR applications in medical-therapeutic settings without an educational purpose.
- VR applications in other educational or recreational contexts.

Study Selection

The online tool "Rayyan" [26] was used for the consolidation and further processing of internationally published German or English articles, which were initially selected on the basis of the title and abstract. With the help of this tool, the research team was able to process the data set independently and in a blinded manner on the one hand but still cooperatively on the other. In this way, articles published up to May 1, 2022, were checked for their suitability with regard to the research question, and relevant hits were identified and extracted in a structured manner. No selection was made with regard to the study design, but both the population involved and the technologies used were taken into consideration. Therefore, we included studies in which nursing trainees, students, and teachers or nursing professionals in further education or training tested the use of VR in the form of head-mounted displays as a medium in targeted teaching-learning arrangements or helped to shape the development process. Of relevance here were, above all, statements on questions of effectiveness; information on increases in knowledge, technology acceptance, and usability; and concrete information on the didactic design of scenarios.

For this reason, publications that focused on other forms of VR representation (eg, nonimmersive 2D representation on a screen or Cardboard app–based models) or use as an assistive technology in nursing or medicine, or focused exclusively on other health professions were not considered. Furthermore, studies that focused on technological details and programming issues, but did not address the educational context, were also excluded.

Data Extraction and Synthesis

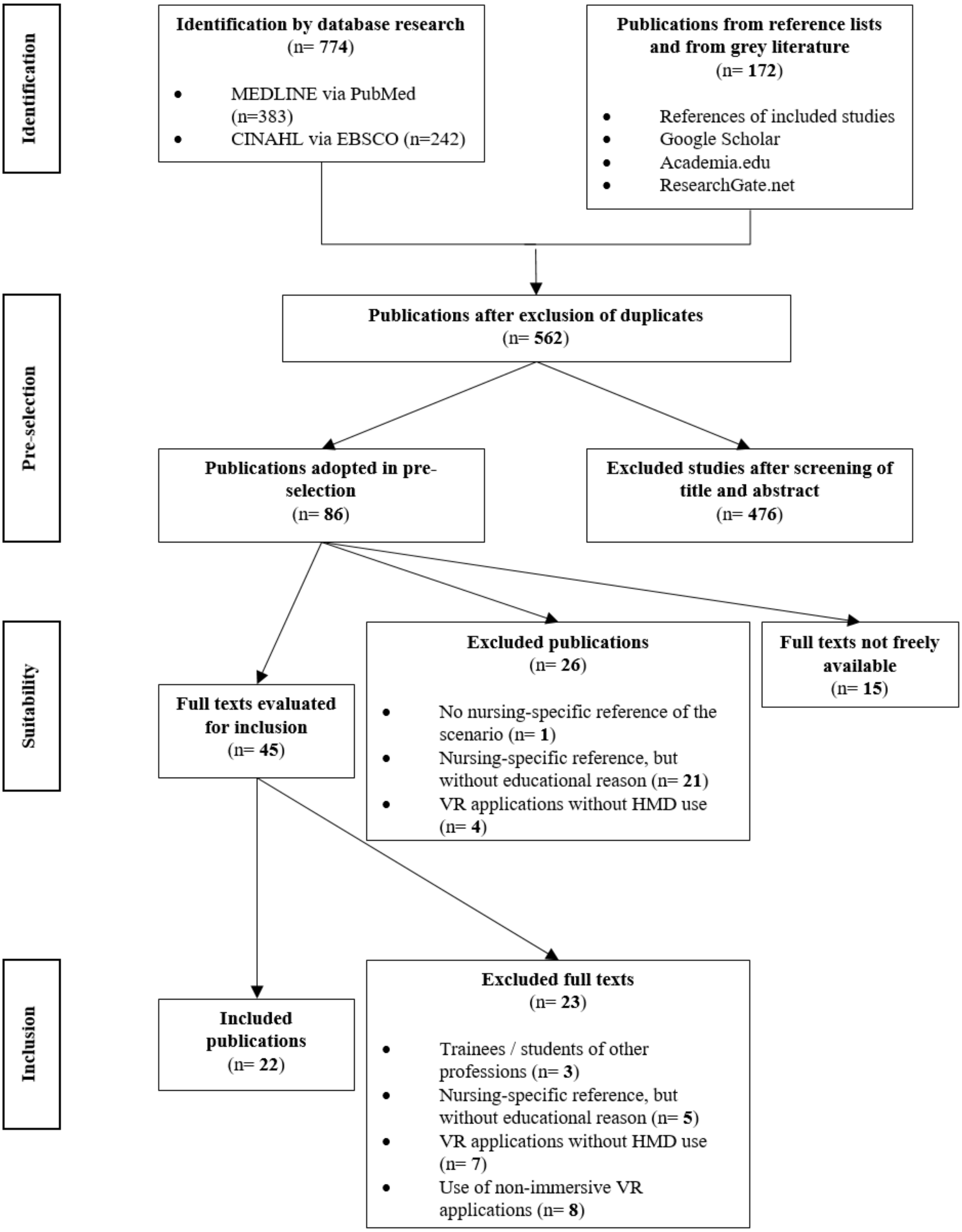
To systematically extract, summarize, and present the information on the current state of science that is relevant to answering the research question, the included studies were first processed narratively in a data table. The analysis and structuring of the data were carried out in terms of the study characteristics and the categories deduced in advance. Accordingly, the upper categories "creative design aspects," "methodological-didactic indications for use," and "research recommendations" served as a tabular and thematic structural basis for the present evidence synthesis (Multimedia Appendix 2 [8,27-47]). The category "general conclusions" included further relevant statements that did not fit into these categories.

Results

Research and Selection of Studies

Through a comprehensive database search, 774 potentially relevant studies were initially found. These were supplemented by 172 publications from a hand search. Studies automatically identified as duplicates by the program were only excluded after an additional manual cross-check, and 562 studies initially remained for the review process. The preselection of 45 articles, which was carried out by a double-blinded examination of the titles and abstracts according to previously defined criteria, led to the evaluation of the full text according to the inclusion criteria. Eventually, the data synthesis included 22 articles. There were no conflicts between the independent reviewers during this process. Figure 1 depicts this procedure graphically in the form of a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart [48].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for literature selection. HMD: head-mounted display; VR: virtual reality.



Characteristics of the Studies Included in the Assessment

The 22 studies included in the assessment were from 14 different countries. Of the 22 studies, 8 were from the United States [31,33,35,36,40,42,46], with 1 co-authored by researchers from Australia [39]; 2 from Ireland [27,47]; 2 from Germany [29,45]; and 1 each from Switzerland [37], Belgium [43], Scotland [28], Norway [8], Canada [43], Brazil [30], South Africa [34], Singapore [8], South Korea [38], and Taiwan [32]. The publication language was mainly English, apart from 1 study, which was published in German. The date of release of more than 77% of the studies was between 2020 and 2022, and only 5 had been published between 2014 and 2019. This is probably due to the fact that VR technologies have become significantly more affordable in recent years through several manufacturers and have thus found their way into private households as well as practical application contexts, with accompanying research. With regard to the study design, the publications were very heterogeneous as is to be expected in a scoping review. Most of the articles involved mixed methods studies [28,30,34-38,40,41,43,45-47]. Moreover, there were 2 qualitative studies [27,32], 2 experimental studies [33,41], and 3 theoretical papers [31,39,42]. Some of these studies were also partially cited in 3 included systematic reviews [8,29,44]. Therefore, this review attempted to reveal aspects that provided hints and recommendations to the chosen categories for the development of virtually supported teaching-learning scenarios for nursing trainees from among different types of publications with regard to their divergent study objectives, settings, and populations. On the basis of this, the research team scanned and divided the publications into groups according to their contribution to one or more of the 3 predefined categories. However, this scoping review aimed to provide a summary of the results of interest and not an all-encompassing presentation of the results. Table 1 illustrates the contribution of the included studies to the deductive categories.

Table 1. Contribution of the included studies to the deductive categories.

Study (author, year)	Creative design aspects	Methodological-didactic indications for use	Research recommendations
Weiß et al [29], 2018	Yes	Yes	Yes
Hara et al [30], 2021	Yes	Yes	Yes
Wells-Beede et al [31], 2022	Yes	Yes	No
Chang et al [32], 2020	Yes	Yes	Yes
Adhikari et al [28], 2021	Yes	Yes	No
Ma et al [33], 2021	Yes	Yes	No
Botha et al [34], 2021	Yes	Yes	No
Butt et al [35], 2018	Yes	Yes	Yes
Shah et al [36], 2021	No	Yes	Yes
Saab et al [27], 2021	Yes	Yes	Yes
Schlegel et al [37], 2019	Yes	Yes	Yes
Lee et al [38], 2020	Yes	Yes	No
Dean et al [39], 2020	Yes	Yes	No
Dorozhkin et al [40], 2017	Yes	Yes	No
Paquay et al [41], 2022	Yes	Yes	Yes
INACSL[a] Standards Committee [42], 2021	Yes	Yes	Yes
Thompson et al [43], 2020	No	Yes	No
Plotzky et al [44], 2021	Yes	Yes	Yes
Kleven et al [45], 2014	Yes	Yes	Yes
Breitkreutz et al [46], 2021	Yes	Yes	No
Shorey et al [8], 2021	Yes	Yes	No
Hardie et al [47], 2020	Yes	Yes	Yes

[a]INACSL: International Nursing Association of Clinical and Simulation Learning.

Characteristics of the Study Participants

Of the 22 studies, 19 described the methodical approach in the empirical surveys (for the cumulative data from the 3 theoretical papers [31,39,42], reference is made here to the respective publication). Accordingly, these included a total of 1193 participants, consisting of 14 teachers, 1018 nursing students or trainees at different stages of their studies or training, and 112 learners from other study programs. The students were studying midwifery (52 participants) [47], emergency medical services (24 participants) [41], and medicine (24 participants) [41], and were participating in interprofessional courses with nursing students in which VR was used for study purposes. A total of 12 students from other nonmedical programs completed the virtual learning program as a control group [45]. One study included 49 participants from a conference, but their professions and levels of education were not explicitly stated [40]. Nevertheless, the study was included because it tested an application that is also explicitly aimed at nurses in training and practice. The sociodemographic data, which were not given in detail in all publications, showed that the participants in the learner group were predominantly female and in an age range of 18 to 36 years but were mostly younger than 25 years. Most of the studies were conducted at a single institution, while only

4 publications presented their results from multicenter studies [30,33,41,46]. Nevertheless, almost all researchers reported that the participants had heterogeneous experiences with VR at the time of the first surveys.

Potential of Implementing VR Into Nursing Education

The included studies depicted a variety of potentials and didactic contexts in which virtually supported teaching-learning scenarios can efficiently supplement conventional teaching methods.

Due to the almost unlimited scope for design, virtually supported teaching-learning scenarios offer a wide range of content-related practice areas for future nursing professionals. This can range from free practice and reflection of communication occasions to technology-assisted patient assessments and nursing actions [36].

The acquisition of knowledge with VR scenarios is based, on the one hand, on the theory of situated learning [8] in order to promote an active connection of didactic principles with clinical competencies [27,28,30,35,36,43,46]. The learning process in VR takes place in the context of specific action goals, competencies, structures, and rules of the simulated nursing action [49]. It also offers a way in which interactions can be experienced and practiced in the social context of a "Community of Practice (CoP)" [50]. On the other hand, VR offers a way for experiential and constructivist learning [8,27,33,37,43,47] by allowing learners to gain meaningful and realistic experiences, even in stressful, rare, and dangerous situations [28,30,37,38,40,41,47]. This comes into play especially when conventional teaching methods can only deficiently depict those situations or if it is important to control them more intensively than in the reality of care [39]. In this way, safe; low-risk; contactless; and shame-, stress-, and fear-free learning is possible [8,27,28,30,32,36,44] when the paradigm of experimental knowledge acquisition and the associated trial-and-error strategy [27,28,32] can be considered its basis. Learners thus perceive less direct pressure that can be exerted by teachers during exercise [32] and can acquire a better understanding of the relevance and effects of individual action steps through directly experienced and concrete consequences [38,46,47]. Combined with gamification elements, trainees enter into a playful learning experience with a positive culture of error [30,37], which, when used sensibly, is able to increase interest and engagement in learning as well as motivation and willingness to actively acquire and discuss the learning content [8,28,30,34,35,37,43,45-47]. While this indicates a positive added value for a better and more satisfying learning experience [27,34,38], the use of VR-supported scenarios can emphasize increased self-confidence and self-awareness from a didactic perspective [27,28,32,43]. The pride and perception of the enhanced competence of learners as well as the playfully conveyed pleasure in a challenge or a competition with fellow students can promote the attention and memorability of the content [27,28,47].

While traditional teaching methods (eg, teaching in the skills lab) require observation and subsequent assessment of performance by a teacher, VR allows learners to gather experiences and impressions unobserved but still in a kind of safe space [46] and to discuss and analyze them later with teachers. As an assistive teaching-learning tool that usefully expands and supplements existing methods [27,28,32,44], VR could replace up to 50% of the clinical hours in the conventional teaching of nursing students [36]. Thus, both novices and experts [30,44] can benefit from it during training (eg, improvement of soft skills such as empathy, interprofessional communication, and collaboration) [29,36,44,45]. Here, the function of the change of perspective or the location-independent multiplayer game option is suitable, which places learners in a realistically depicted setting of care in a targeted situation (eg, communicatively challenging situation of case discussion or family counseling) and thus extends conventional role play [44,45]. Similarly, in this environment, it is possible to expand competencies, such as observation and reflection, on a case- and action-related basis or to look at doubtful situations from different perspectives and then discuss them together with the teacher or in a class group [30,47]. Accordingly, the deliberate use of virtual scenarios in generalist nursing education is well suited to acquiring new knowledge for the first time in a multisensory way, consolidating it through repetition or different action requirements, and forming abstract concepts in an experience-based and feedback-supported way [27,28,32,46,47].

Above all, the option of applying theoretically taught content with the accompanying required practical skills and abilities of nursing in a situation-oriented manner and connected with virtual persons who require care enables learners to gather and reflect on practical experiences even before their first clinical assignments. This, in turn, can result in them being more courageous, motivated, and committed in their active engagement with real patients. On the other hand, they might also experience a feeling of greater competence and self-efficacy, which can reduce the theory-practice transfer that is often perceived as difficult [27,30,36,38,42-47]. Although the respective studies had some limitations with regard to design and generalizability, some research teams reported objectively ascertainable cognitive, procedural, and psychomotor gains, in addition to the rather subjectively assessed personal and affective added values.

However, VR is not suitable as a stand-alone teaching approach that can or should replace teaching without specific instructions and guidance [27]. Rather, it is a matter of meaningfully integrating the possibilities offered by immersive virtual simulations into the teaching of prospective nurses or into the continuing education and training of nurses who are already working.

Didactic-Methodological Recommendations for the Use of Virtually Supported Scenarios in Nursing Education

The possible values of virtually supported teaching-learning scenarios in generalist nursing education are numerous. Nevertheless, there is a need for some didactic-methodological considerations and measures to be able to use them.

The prerequisite for this, however, is the economic, intentional, systematic, flexible, and learner-centered concept that facilitates the use of this medium, which has been adapted to the respective groups of learners, their levels of skills and knowledge, their learning experiences, and previous methods [30,32,35,37,42]. According to the deliberate practice theory, exercises should be selected on the basis of clearly defined, specific, appropriate, and measurable learning objectives that correspond to the real requirements of nursing practice [29,30,35,37,38,42]. Thus, teachers face the task of didactically reducing the available virtual possibilities by focusing on single aspects and significant content [44,47]. Due to this and several other factors, VR is not an adequate substitute for experienced instructors to teach professional nursing [27].

Saab et al [27] emphasized that the core of the nursing values of care and compassion is still human interaction. Trainees cannot acquire these exclusively through simulations. Rather, the personal and professional experience of teachers should convey these values. Additionally, they have to stimulate an empathic curiosity to generate a greater willingness to put oneself in the situation of the person receiving care and to support an accompanying in-depth understanding of the respective situation [39]. Moreover, VR cannot replace the deepening practice with the person receiving care or those involved in care for the hermeneutic case understanding of learners, in which individualized or at least partially individualized decisions about interventions in particular cases should be made with direct communicative reflection on needs and requirements. Furthermore, the abovementioned group of authors stated that there is no adequate substitute for personal and continuous feedback from the teacher for the preparation, support, and reflection of the learning situation [27]. In addition to this reflected use in general, it requires a well thought-out concept to leverage the potential of virtually supported scenarios in nursing education.

Considerations on Implementing VR Into Teaching

The naive use of VR for self-purposes or entertainment should only be found in the leisure sector. However, in order to make a purposeful and targeted contribution to the acquisition of skills and abilities by nursing trainees, it is important to proceed in a planned and systematic manner. Thus, Dean et al [39] called for users to not become passive VR consumers but to continue to maintain a critical, analytical, and thoughtful attitude for transferability to the reality of care. This also reflects the basic attitude of caregivers. Since future nursing professionals should always adopt a critical and reflexive attitude in the course of the increasing use of technologies in nursing practice and should also be sensitized to this in their training, this applies equally to not only teachers but also learners with regard to the use of virtual scenarios.

This also presupposes that teachers organize optimal framework conditions. The International Nursing Association of Clinical and Simulation Learning (INACSL) Standards Committee [42] and Hardie et al [47] therefore recommended a detailed prebriefing for preparation and introduction to the handling of the technology. This includes concrete preinstructions [32,38,41,44], which involve the correct use of technical devices, such as the controller [30,37], and getting used to the glasses and the changing perception [30,32]. Instructions on the associated teaching material and learning content and the requirements of the scenario should also be part of the introduction. This can be done either face-to-face with tutors or instructors or via a video [28] or interactive tutorial. Good instructions and ease of use open up the potential for learners to use VR independently of teachers and thus of location [44,46], and possibly even use a multiplayer version [35,40,44]. A final debriefing in the form of feedback sessions or accompanying reflection tasks supplements meaningful usage. This can support teachers and trainees to identify learners' current strengths and weaknesses or to analyze and discuss discrepancies between the learning experience provided and the understanding of the nursing concepts presented or even the reality of care.

Accordingly, the use of virtual scenarios is recommended especially for smaller classes [27], so that trainers can handle the organization of the set-up of the simulations as well as the assignment and creation of rotation plans [36] in a manageable and efficient way. Under certain circumstances, the information or involvement of additional teachers should be considered [36]. Moreover, it should be considered whether specifically trained instructors or fellow students should provide support for the learners, for example, to secure the environment [43]. This also becomes relevant if trainees are given the opportunity to use or borrow VR headsets for voluntary practice in their free time in a separate room [27].

This, in turn, would not only enable the self-organized learning demanded by learners and the curricula [38] but also invalidate the argument that VR isolates users and show that it promotes social interaction [44], which is a highly relevant component in nursing.

In addition to the meaningful intention to implement VR in educational contexts, design aspects are crucial elements for using the various potentials of VR.

Overview of the Design Considerations of Virtually Supported Scenarios

With the exception of 2 studies [36,43], all included publications contained mostly experience-based hints and recommendations regarding the design of virtually supported teaching-learning scenarios for generalist nursing education.

To empower trainees to handle care situations, specific circumstances, and various settings, virtually supported scenarios should provide a realistic, plausible, and immersive learning environment [28,38,41,45,47], which should have consistent [31], clinically correct [34], and narrative story structures [47]. Due to this, multiprofessionally composed development and research groups [32,42] have to predefine concrete learning objectives [42] adapted to the current ability and knowledge level of learners [30], whereby learners come across the subjective relevance of the scenario they have experienced [51]. This forms the basis for the deduction of the most profitable specific means, details, and features. Authentic, motivating, and moderately challenging experiences should always be the goal. Thus, focusing on typical visual and auditory details relevant to the nursing process [44,46] is a major aspect. A basic prerequisite is high image and sound quality [27,34,38,41]. By observing the correct lens focus [28] and a high refresh rate, users can read text insertions [31] or recognize facial expressions and gestures [30] more easily. Moreover, this can increase the sense of immersion and perceived spatial presence within the chosen scenario [33,38,45] and prevent motion sickness [30]. Considering the cognitive load, the INACSL Standards Committee [42] recommends the selection of the type and degree of fidelity (eg, with regard to physical, conceptual, or psychological parameters). Therefore, trainees are able to draw their attention to the respective action demands [27] and challenges of different stress levels [29,37,38]. This could be supported by the targeted use of visual cues, including color markings, highlights, or animations [30,31]. The function of pausing during an exercise [31] in order to reflect on the next steps or to record the entire exercise [8] for later discussion can help to create a critical reflective attitude toward one's performance. In addition to realistic visual and auditory details, the integration of tactile stimuli in the sense of a mixed reality experience [46] and 360° views, which enable observation of one's performance from different angles [31], could be useful. The change of perspective [51] or modality of experience provides a basis for the reflection and discussion of actions and reactions demanded in specific situations. Since VR allows for slipping into other roles, it can promote essential nursing skills, such as empathy [52-54], which is particularly important for recreating other life perspectives and situations. The understanding of the needs and requirements of virtual people who require care on a physiological level can be supported by interactive models visualizing anatomical structures as well as regular or pathological processes [27,45,47].

Another essential recommendation for VR learning scenarios is the use of gamification elements [55]. The implementation of game-based details in the applications enables the strengthening of memory pathways [46,47], which in turn can positively influence learning outcomes [28]. This includes, for example, scores or rewards in the form of medals, congratulatory banners, or colored lights [30]. The given feedback can additionally motivate learners to perform practical nursing activities in VR [28] and support them in the development of problem-solving [27] and procedural skills [30]. Likewise, this can be supported by time limits for the execution of individual nursing actions.

Furthermore, if the application enables the collection of game-played data [41] and, for example, allows their visualization to both learners and teachers in the form of error rates [40], it can not only document but also promote intended learning outcomes, particularly when learners use this feature to analyze the learning gains according to performance [8]. Consequently, the scenarios, which rise in complexity in more challenging difficulty levels [35], should offer the possibility of the repetition of exercises [35,44] in order to achieve an increase in competence individually and in terms of one's responsibility. If a trial-and-error strategy [27] forms the basis for this procedure, the learning process can be positively reinforced.

However, virtually supported scenarios can only unfold their potential if the handling scenarios allow. On the one hand, uncomplicated and trouble-free handling and experience of technical possibilities can increase the learners' sense of presence in the situation [41]. On the other hand, it is an essential factor in the prevention of motion sickness [30]. Therefore, designers should pay attention to a high degree of correspondence between the image and the respective head movement, and the use of high-resolution graphics and the mitigation of technical overreactions, for example, can be useful when reaching for objects in virtual space [8,46].

Thus, design aspects should address the questions of handling and acceptance and the associated benefits for learning.

Research Recommendations

Owing to the greater availability of and interest in the use of VR as a teaching-learning medium, the corpus of studies in the field of nursing education has grown immensely in recent years. Nevertheless, the studies included in this review have stated the inconsistent use of VR terminology, indicating the need for an unambiguous definition in publications [29,44], and have mentioned the requirement for further research with larger samples and associated statistical analysis [27,29] with regard to various aspects.

On the one hand, this involves the investigation of technical parameters and interactive possibilities, such as stereognosis [32], motion tracking, and the integration of haptic devices, enhancing VR interface elements or social media and other mobile technologies to enable collaborative learning and effective distribution of educational content [45]. On the other hand, there is a demand for further investigation of the learning process itself by means of virtual simulations and the transferability of learned nursing-relevant content to real clinical practice [30]. Shah et al [36] recommended quantitative ethnography as a possible research method to take a closer look at associated emotions; ways of thinking and acting during immersion; and how, why, or when learning groups differ in this respect. If researchers use such comparative studies, for example, to analyze several sessions with the same and different instructors and assess learners' perceptions during the instructions in prebriefings and debriefings or with regard to different content [36], they should take care to pilot the study [42] and to provide comparable test conditions for participants in the control group so that they can, for example, walk through a real patient room in search of faulty aspects of patient and workplace safety [37]. They should ensure almost the same conditions when surveying individual learning experiences [42], learning gains [36], and long-term knowledge retention or improved patient outcomes [35]. Studies for examining and evaluating the use of VR scenarios in education in more detail should also survey possible previous VR experiences of users [41] to be able to consider possible influencing factors or risks of bias.

For the use and design of virtually supported teaching-learning scenarios for generalist nursing education, the integration of a best-practice simulation framework [47] (eg, INACSL criteria [42] and Jeffries' Simulation Theory [32]) for the consideration of not only microdidactic but also meso- and macrodidactic influencing factors is recommended. Thus, in addition to design and application aspects, questions about financial effects or the return on investment [35] also come into focus, and interprofessional cooperation [32] should take these into account, especially for continuous modification and optimization of scenarios. Targeted needs assessment [42] and continuous learner and teacher involvement in development [30] are critical factors for the appropriate and economic development of an effective teaching-learning medium.

Discussion

Implications and Aspects of the Use of VR Teaching-Learning Scenarios

Within the framework of the literature research and the results presented, it must be stated that there are various ways to define VR, and it can encompass different devices, degrees of immersion, and interactions. Uniform definitions of the terms used would therefore be desirable [35,50]. Nevertheless, this medium in its various manifestations is generating successively more interest not only within the private leisure sector but also as a supplementary teaching-learning instrument in both general education and medical and nursing education contexts, as VR can meaningfully expand the number of methods with regard to various teaching-learning outcomes [36,38].

On the one hand, a virtual change of perspectives, role plays, or teamwork tasks in authentically depicted nursing scenarios could support the learning, practice, and repetition of personal and social competencies, such as empathy, heuristic case understanding, and targeted observation, which are relevant in the relational profession of nursing [2,39]. On the other hand, learners can consolidate procedural skills and abilities in virtually supported care situations by means of demonstrations, step-by-step instructions, and various feedback mechanisms [8]. In this way, they can safely apply theoretical content before, during, or even after a practical assignment in a concrete action situation and thus consolidate or assess their knowledge. This has the potential to soften limiting framework conditions and facilitate theory-practice transfer [2,36]. Teachers can benefit from the targeted use of VR in that they can give trainees learning tasks that are not bound to time and place, and these trainees are in turn more motivated and committed to partly self-directed teaching [41,50]. In addition, teachers and trainees command content illustrated more practically for appropriate discussion and reflection together [48,53].

Beyond the possibility of enhancing practical skills continuously in a standardized way, the use of VR supports trainees' sensitization to digital nursing technologies and helps expand their digital skills in a practical setting. Even sustainability issues can be addressed in this way [38].

Nevertheless, it is important to note that almost all studies unanimously emphasized that virtually supported teaching-learning scenarios are still not an omnipotent substitute in teaching and that their use is rather critically reflected and well-considered at those points where conventional teaching methods reach their limits [40,46] to comprehensively prepare learners for the future role of a professionally acting nurse [38]. This scoping review offers an overview of the implications, considerations, and recommendations to develop and implement virtually supported scenarios reasonably and purposefully for educational demands in nursing education. An excerpt of the results is shown in Textbox 2.

This includes not handing out VR glasses to learners in an uncontrolled manner, but rather embedding the application methodically and didactically in the lessons in a meaningful way to ensure pre- and postdiscussions as well as parallel professional and technical support. Only then can the presented content effectively support individual learning [2,42,48].

VR is consequently highly recommended to complement the third location of learning, that is, the skills lab [50]. The complexity of the practice is only approximately representable owing to current restrictions, such as limitations in haptics or olfaction, which are of great relevance in the care sector, and the combination of these can lead to the high resemblance of daily nursing practice and the broad preparation of trainees [16,38]. In addition, technology in the field of VR will continue to develop in the future, and possibilities for realization may arise for those constraints. In the best case, this will happen based on the needs and requirements of respective target groups in multiprofessional teams and with co-creative participatory procedures [38]. Thus, further prospective research fields are emerging in addition to the current cost-benefit analyses, large randomized controlled studies in various teaching-learning settings, and surveys on improved patient outcomes [35,36,38,41], and these will offer further potential and provide focal points for investigation that need to be critically considered.

Textbox 2. Recommendations for the development and implementation of virtual reality scenarios.

Design recommendations

- Realism and plausibility
- Attractive playful design with high image and sound quality
- Dialog-based narration
- Adoption of perspective
- Direct feedback and tangible consequences of action
- Hierarchical structure
- Data collection and reproduction
- Clear handling, navigation, and instructions
- Pause, repeat, and record functions
- Location-independent multiplayer option

Didactic considerations

- Assistive, activating, and motivating
- Multimodal, learner-centered, and experience-based teaching concept
- Specifically formulated learning goals
- Secure standardized environment
- Consolidation of theoretical, procedural, and application knowledge
- One-to-one support including feedback
- Situational testing
- Independent and flexible in terms of time, and repeatable as often as required
- Heuristically reflexive decision-making and problem-solving processes
- Self-confidence in processes, expertise, and communication skills

Research recommendations

- Clear definition of terminology
- Cooperative and co-creative development processes
- Larger samples and statistical analysis
- Varying settings and conditions
- Evaluation of improved patient outcomes
- Longitudinal studies on knowledge
- Cost-benefit analyses
- Inclusion of additional interactive functions
- Consideration of theoretical frameworks
- Integration of best-practice simulation frameworks (eg, International Nursing Association of Clinical and Simulation Learning criteria)

Limitations

A methodological strength of this scoping review is the comprehensive and supplementary hand search conducted in parallel with the database search and the citation tracking to counteract the risk of excluding relevant hits. Furthermore, the research team used a tool for blinded analysis to avoid selection bias in the selection of studies as far as possible. Nevertheless, the initial decision for a sensitive search principle was changed in favor of a specific procedure, as there has been an enormous growth of extended reality (XR) applications in educational and medical contexts in recent years. Accordingly, there is a growing amount of research papers on a wide variety of focal points. However, these often only correspond to the previously defined inclusion criteria in individual points, and thus, they do not answer or inadequately answer the concrete underlying research questions for the selected target group or the corresponding application. This is also the reason for another limitation of the study. As nursing education in its generalist application in Germany is unique in a worldwide comparison, the largely international research results are only partly transferable to local framework conditions, teaching methods, and content, as well as the requirements in the initial training of future nursing professionals. Furthermore, the data protection regulations applicable in Germany should be taken into account. These can influence not only the choice of devices but also the processing and use of the generated data. Therefore, critical considerations are relevant in the reception and the attempt to generalize the results in other contexts, especially since the focus was on a selective collection of data and not on a dedicated analysis or detailed comparison of the studies with each other.

Comparison With Prior Work

This scoping review reveals the results of selected publications according to a specific search principle. Although the aim was not to compare the studies with regard to the respective design or the reported results, the latter could be summarized under the deductively created paragraphs that address recommendations and considerations.

Compared with previous studies, which were partly considered in this study, it was possible to generate a general overview of relevant aspects that fundamentally characterize virtually supported teaching-learning scenarios in initial nursing education. On the other hand, the basis for the identification was the very specific context of nursing education in Germany. However, a large number of studies published thus far have focused on other study populations from the medical and general education sectors or other definitions of VR in their surveys and explanations.

The inclusion and exclusion criteria used served primarily to provide those actors involved in German nursing education and training with information on the use and development of virtually supported teaching-learning scenarios, which corresponded to the international consensus and met the needs of the German context. Consequently, this publication can serve as a point of reference for both national and international recipients, provided that they critically evaluate it and, if necessary, supplement relevant aspects, which are related to the respective country, action, teaching, or study population background.

Conclusions

Flexible use, a positive error culture, and learning that can be individually controlled and adapted to the knowledge levels of trainees by means of virtually supported teaching-learning scenarios can increase learning motivation and satisfaction. Simultaneously and compared to other common teaching methods, VR can reduce time, personnel, and material resources, and future nursing professionals can specifically train, deepen, and consolidate the procedural, personal, and social competencies of professional nursing knowledge and actions in both theoretical and practical teaching sessions.

Nevertheless, VR cannot and should not replace experienced nursing teachers, especially to convey elementary nursing values such as care and compassion. Therefore, learners and teachers should be actively involved in the co-creative design and evaluation process of virtually supported teaching-learning scenarios for the acquisition of skills and competencies in a practical yet safe setting. This will help to reveal the needs of the target group from the beginning and to incorporate them directly into the development on an iterative basis. In addition, future users can identify weak points or errors in content or applications more quickly than nonspecialist developers who may focus on different aspects. This could also launch the systematic implementation of this medium in the curriculum. Moreover, trainees and teachers will be sensitized to apply it critically and reflectively owing to the deeper insights that accompany the process.

Acknowledgments

The study conducted as part of the DigiCare project was financed from funds of the Federal Ministry of Education and Research and within the implementation phase of the WIR! Programme – Change through innovation in the region (project duration: August 01, 2021, to July 31, 2024; funding code: 03WIR3106B).

Authors' Contributions

All involved authors equally contributed to all manuscript components.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings for the literature search in 2 separate databases.
[DOCX File , 30 KB - games_v12i1e53356_app1.docx]

Multimedia Appendix 2
Overview of the essential characteristics of the studies included in this scoping review.
[DOCX File , 54 KB - games_v12i1e53356_app2.docx]

Multimedia Appendix 3
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.

[PDF File (Adobe PDF File), 498 KB - games_v12i1e53356_app3.pdf]

References

1. Prognose zu Bedarf und Angebot an Pflegekräften in Deutschland nach Szenario in den Jahren 2024 bis 2049. Statista. URL: https://de.statista.com/statistik/daten/studie/172651/umfrage/bedarf-an-pflegekraeften-2025/ [accessed 2024-06-13]
2. Arbeitsbedingungen in der Pflege. Hans Böckler Stiftung. URL: https://www.boeckler.de/de/auf-einen-blick-17945-zahlen-und-studien-zum-pflegenotstand-und-wege-hinaus-17962.htm [accessed 2024-06-13]
3. Können Innovation und Digitalisierung den Pflegenotstand lindern? Pflegenot Deutschland. URL: https://www.pflegenot-deutschland.de/pflegepolitik/innovation-digitalisierung-pflegenotstand/ [accessed 2024-06-13]
4. Weiß T, Meißner T, Kempa S. Pflegeberufereformgesetz (PflBRefG). Wiesbaden: Springer Gabler; 2018.
5. Evans M. Soziale Dienstleistungsarbeit im Spiegel der Digitalisierung: Ein Impuls zur Analyse der Arrangements von Organisation, Profession und Klient. In: Ernst G, Zühlke-Robinet K, Finking G, Bach U, editors. Digitale Transformation: Arbeit in Dienstleistungssystemen. Baden-Baden: Nomos Publishing; 2020:275-288.
6. Bowman DA, McMahan RP. Virtual reality: how much immersion is enough? Computer 2007 Jul;40(7):36-43. [doi: 10.1109/mc.2007.257]
7. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inform Syst 1994;12(12):1321-1329 [FREE Full text]
8. Shorey S, Ng ED. The use of virtual reality simulation among nursing students and registered nurses: A systematic review. Nurse Educ Today 2021 Mar;98:104662. [doi: 10.1016/j.nedt.2020.104662] [Medline: 33203545]
9. Ghanbarzadeh R, Ghapanchi AH, Blumenstein M, Talaei-Khoei A. A decade of research on the use of three-dimensional virtual worlds in health care: a systematic literature review. J Med Internet Res 2014 Feb 18;16(2):e47 [FREE Full text] [doi: 10.2196/jmir.3097] [Medline: 24550130]
10. Strangman N, Hall T, National Center on Accessing the General Curriculum. Virtual reality/computer simulations. CiteSeer. 2003. URL: https://tinyurl.com/y54bfp4z [accessed 2024-06-13]
11. Choi DH, Dailey-Hebert A, Estes JS. Emerging Tools and Applications of Virtual Reality in Education. Hershey, PA: IGI Global; 2016.
12. Kilmon C, Brown L, Ghosh S, Mikitiuk A. Immersive virtual reality simulations in nursing education. Nurs Educ Perspect 2010;31(5):314-317. [Medline: 21086871]
13. Einsatz von Virtual Reality gestützten Absaugtraining für professionell Pflegende in Ausbildung und Praxis. pflegekongress. URL: https://www.pflegekongress.at/html/publicpages/Vortraege2019/Lindewedel-Reime_pflegekongress19.pdf [accessed 2024-06-13]
14. Chen F, Leng Y, Ge J, Wang D, Li C, Chen B, et al. Effectiveness of virtual reality in nursing education: meta-analysis. J Med Internet Res 2020 Sep 15;22(9):e18290 [FREE Full text] [doi: 10.2196/18290] [Medline: 32930664]
15. Wernbacher S. Die Förderung der kommunikativen Kompetenz in der Gesundheits- und Krankenpflegeausbildung. Unipub. URL: https://unipub.uni-graz.at/obvugrhs/content/titleinfo/3466363/full.pdf [accessed 2024-06-13]
16. Virtuelle Realität und E-Learning. e-teaching. URL: https://www.e-teaching.org/didaktik/gestaltung/vr/vr.pdf [accessed 2024-06-13]
17. Schmal J. Unterrichten und Präsentieren in Gesundheitsfachberufen. Methodik und Didaktik für Praktiker. Heidelberg: Springer Berlin; 2023.
18. Pfeifer L, Nauerth A, Raschper P, Freese C, Bräkling S. Virtual Reality basierte Digital Reusable Learning Objects in der Pflegeausbildung - Rahmenbedingungen, Anforderungen und Bedarfe aus pflegepädagogischer Sicht. In: Innovative Lehr-/Lernszenarien in den Pflege- und Gesundheitsberufen. Working Paper-Reihe der Projekte DiViFaG und ViRDiPA. Vol 2 aktual. Version. Bielefeld: Fakultät für Erziehungswissenschaft; 2021.
19. Kleinknecht M, Schneider J. What do teachers think and feel when analyzing videos of themselves and other teachers teaching? Teaching and Teacher Education 2013 Jul;33:13-23. [doi: 10.1016/j.tate.2013.02.002]
20. Landesprojekt: Theorie-Praxis-Transfer in der Ausbildung in den Pflegeberufen. Branchenmonitoring Gesundheitsfachberufe Rheinland-Pfalz. URL: https://www.branchenmonitoring-gesundheitsfachberufe-rlp.de/wp-content/uploads/2021/11/BadP_28_Endfassung_0.pdf [accessed 2024-06-13]
21. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
22. Kaufhold M, Weyland U. Auswirkungen der Digitalisierung in den Gesundheitsberufen – Ansätze im Projekt HumanTec am Beispiel Pflege. Forum der Arbeitsgemeinschaft Berufsbildungsforschungsnetz. URL: https://www.agbfn.de/dokumente/pdf/AGBFN_betrieblichesLernen_Abstract_Kaufhold,Weylandt_(FachhochschuleBielefeld).pdf [accessed 2024-06-13]
23. von Elm E, Schreiber G, Haupt CC. Methodische Anleitung für Scoping Reviews (JBI-Methodologie). Z Evid Fortbild Qual Gesundhwes 2019 Jun;143:1-7. [doi: 10.1016/j.zefq.2019.05.004] [Medline: 31296451]
24. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009 Jun 27;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]

25. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
26. Hirt J, Nordhausen T. Digitale Anwendungen für die Studienauswahl im Rahmen von systematischen Evidenzsynthesen. Pflege 2019;32(5):277-278 [FREE Full text] [doi: 10.1024/1012-5302/a000690] [Medline: 31538877]
27. Saab MM, Hegarty J, Murphy D, Landers M. Incorporating virtual reality in nurse education: A qualitative study of nursing students' perspectives. Nurse Educ Today 2021 Oct;105:105045 [FREE Full text] [doi: 10.1016/j.nedt.2021.105045] [Medline: 34245956]
28. Adhikari R, Kydonaki C, Lawrie J, O'Reilly M, Ballantyne B, Whitehorn J, et al. A mixed-methods feasibility study to assess the acceptability and applicability of immersive virtual reality sepsis game as an adjunct to nursing education. Nurse Educ Today 2021 Aug;103:104944. [doi: 10.1016/j.nedt.2021.104944] [Medline: 34015677]
29. Weiß S, Bongartz H, Boll S, Heuten W. Applications of immersive VR in nursing education. Zukunft der Pflege - Innovative Technologien für die Pflege 2018:174-179 [FREE Full text]
30. Hara CYN, Goes FDSN, Camargo RAA, Fonseca LMM, Aredes NDA. Design and evaluation of a 3D serious game for communication learning in nursing education. Nurse Educ Today 2021 May;100:104846. [doi: 10.1016/j.nedt.2021.104846] [Medline: 33751998]
31. Wells-Beede E, Garcia B, Chun SW, Kicklighter C, Seo JH. Creative solutions for complex circumstances: the utilization of virtual reality in a specialty course. Clinical Simulation in Nursing 2022 Apr;65:82-85. [doi: 10.1016/j.ecns.2022.01.004]
32. Chang YM, Lai CL. Exploring the experiences of nursing students in using immersive virtual reality to learn nursing skills. Nurse Educ Today 2021 Feb;97:104670. [doi: 10.1016/j.nedt.2020.104670] [Medline: 33264739]
33. Ma Z, Huang K, Yao L. Feasibility of a computer role-playing game to promote empathy in nursing students: the role of immersiveness and perspective. Cyberpsychol Behav Soc Netw 2021 Nov;24(11):750-755. [doi: 10.1089/cyber.2020.0371] [Medline: 33989057]
34. Botha BS, de Wet L, Botma Y. Undergraduate Nursing Student Experiences in Using Immersive Virtual Reality to Manage a Patient With a Foreign Object in the Right Lung. Clinical Simulation in Nursing 2021 Jul;56:76-83. [doi: 10.1016/j.ecns.2020.10.008]
35. Butt AL, Kardong-Edgren S, Ellertson A. Using game-based virtual reality with haptics for skill acquisition. Clinical Simulation in Nursing 2018 Mar;16:25-32. [doi: 10.1016/j.ecns.2017.09.010]
36. Shah M, Siebert-Evenstone A, Eagan B, Holthaus R. Modeling educator use of virtual reality simulations in nursing education using epistemic network analysis. 2021 Presented at: 7th International Conference of the Immersive Learning Research Network (iLRN); May 17, 2021-June 10, 2021; Eureka, CA. [doi: 10.23919/iLRN52045.2021.9459408]
37. Schlegel C, Weber U. Lernen mit Virtual Reality: Ein Hype in der Pflegeausbildung? ResearchGate. URL: https://tinyurl.com/srt3kx3x [accessed 2024-06-13]
38. Lee Y, Kim SK, Eom M. Usability of mental illness simulation involving scenarios with patients with schizophrenia via immersive virtual reality: A mixed methods study. PLoS One 2020 Sep 16;15(9):e0238437 [FREE Full text] [doi: 10.1371/journal.pone.0238437] [Medline: 32936813]
39. Dean S, Halpern J, McAllister M, Lazenby M. Nursing education, virtual reality and empathy? Nurs Open 2020 Nov;7(6):2056-2059 [FREE Full text] [doi: 10.1002/nop2.551] [Medline: 33072391]
40. Dorozhkin D, Olasky J, Jones DB, Schwaitzberg SD, Jones SB, Cao CGL, SAGES FUSE Committee. OR fire virtual training simulator: design and face validity. Surg Endosc 2017 Sep 30;31(9):3527-3533 [FREE Full text] [doi: 10.1007/s00464-016-5379-7] [Medline: 28039649]
41. Paquay M, Goffoy J, Chevalier S, Servotte J, Ghuysen A. Relationships between internal factors, social factors and the sense of presence in virtual reality-based simulations. Clinical Simulation in Nursing 2022 Jan;62:1-11. [doi: 10.1016/j.ecns.2021.09.006]
42. INACSL Standards Committee. Healthcare Simulation Standards of Best PracticeTM Simulation Design. Clinical Simulation in Nursing 2021 Sep;58:14-21. [doi: 10.1016/j.ecns.2021.08.009]
43. Thompson D, Thompson A, McConnell K. Nursing students' engagement and experiences with virtual reality in an undergraduate bioscience course. Int J Nurs Educ Scholarsh 2020 Sep 16;17(1):17 [FREE Full text] [doi: 10.1515/ijnes-2019-0081] [Medline: 32941161]
44. Plotzky C, Lindwedel U, Sorber M, Loessl B, König P, Kunze C, et al. Virtual reality simulations in nurse education: A systematic mapping review. Nurse Educ Today 2021 Jun;101:104868. [doi: 10.1016/j.nedt.2021.104868] [Medline: 33798987]
45. Kleven N, Prasolova-Førland E, Fominykh M, Hansen A, Rasmussen G, Sagberg L, et al. Training nurses and educating the public using a virtual operating room with Oculus Rift. 2014 Presented at: 2014 International Conference on Virtual Systems & Multimedia (VSMM); December 09-12, 2014; Hong Kong, China. [doi: 10.1109/VSMM.2014.7136687]
46. Breitkreuz KR, Kardong-Edgren S, Gilbert GE, DeBlieck C, Maske M, Hallock C, et al. A multi-site study examining the usability of a virtual reality game designed to improve retention of sterile catheterization skills in nursing students. Simulation & Gaming 2020 Sep 25;52(2):169-184. [doi: 10.1177/1046878120954891]

47. Hardie P, Darley A, Carroll L, Redmond C, Campbell A, Jarvis S. Nursing and Midwifery students' experience of immersive virtual reality storytelling: an evaluative study. BMC Nurs 2020;19:78 [FREE Full text] [doi: 10.1186/s12912-020-00471-5] [Medline: 32821245]
48. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
49. Dehnbostel P, Gießler W. Arbeitsintegrierte Berufsqualifizierung in der Pflege. Lernen im Prozess der Arbeit in neuen Lernformen und erweiterten Lernorten. In: Marchwacka MA, editor. Handbuch Pflegeausbildung. Theorie – Empirie – Praxis. Bern: Hofgrefe; 2023:127-138.
50. Lave J, Wenger E. Situated Learning: Legitimate Peripheral Participation. Cambridge: Cambridge University Press; 1991.
51. Siebert H. Didaktik - mehr als die Kunst des Lehrens? REPORT Literatur- und Forschungsreport Weiterbildung: Didaktik 2005;3:9-17 [FREE Full text]
52. Seinfeld S, Hasler BS, Banakou D, Levy J. Editorial: Virtual reality and empathy. Front Psychol 2022 Nov 21;13:1089006 [FREE Full text] [doi: 10.3389/fpsyg.2022.1089006] [Medline: 36483728]
53. Bischoff-Wanner C. Empathie in der Pflege. Begriffserklärung und Entwicklung eines Rahmenmodells. Bern: Huber; 2002.
54. Klimasch G. Pflegerische Empathie (lernen) – Sichtweisen von Pflegelernenden. Pflege 2022 Oct;35(5):314-315. [doi: 10.1024/1012-5302/a000907] [Medline: 36117437]
55. Routledge H. Why Games Are Good For Business: How to Leverage the Power of Serious Games, Gamification and Simulations. London: Palgrave Macmillan; 2016.

Abbreviations

dVRS: desktop virtual reality simulation
INACSL: International Nursing Association of Clinical and Simulation Learning
iVRS: immersive virtual reality simulation
VR: virtual reality

Edited by N Ahmadpour; submitted 04.10.23; peer-reviewed by X Yi, W Li, S Matsuda; comments to author 25.02.24; revised version received 16.03.24; accepted 16.04.24; published 08.07.24.

Please cite as:
Steindorff JV, Redlich LM, Paulicke D, Jahn P
Use and Design of Virtual Reality–Supported Learning Scenarios in the Vocational Qualification of Nursing Professionals: Scoping Review
JMIR Serious Games 2024;12:e53356
URL: https://games.jmir.org/2024/1/e53356
doi:10.2196/53356
PMID:

Original Paper

Impact of Facilitation on Cognitive Flow in a Novel Diabetes Management Rehearsal Game for Health Professions Education: Mixed Methods, Open-Label, Superiority Randomized Controlled Trial

Jun Wen Tan[1*], BSc; Gabriel Tan[2*], MBBS; Xia Lian[3*], APN; Darren Kai Siang Chong[1*]; Preman Rajalingam[4*], PhD; Rinkoo Dalan[1,3*], MBBS, PhD; Sreenivasulu Reddy Mogali[1*], MSc, PhD

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]MOH Holdings Pte Ltd, Singapore, Singapore

[3]Department of Endocrinology, Tan Tock Seng Hospital, Singapore, Singapore

[4]Institute of Learning, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates

[*]all authors contributed equally

Corresponding Author:
Sreenivasulu Reddy Mogali, MSc, PhD
Lee Kong Chian School of Medicine
Nanyang Technological University
Office of Medical Education, Clinical Sciences Building
11 Mandalay Road
Singapore, 308232
Singapore
Phone: 65 65923114
Email: sreenivasulu.reddy@ntu.edu.sg

Abstract

Background: Though the prevalence of diabetes is set to increase, most serious game solutions typically target patient self-management and education. Few games target health care professions education, and even fewer consider the factors that may increase their efficacies. The impact of facilitation, a prominent feature of health professions education, is examined in the context of a rehearsal-based diabetes management serious game.

Objective: In this mixed methods, open-label, superiority randomized controlled trial, we compare student performance, attitudes, and perceptions of a rehearsal-based diabetes management game for health care professionals.

Methods: Student participants were randomized into 2 groups to play a diabetes management game. The control group played the game alone, and the intervention group played the same game alongside a facilitator tasked to moderate overall challenge levels and address queries. Both groups were administered the Flow Short Scale, a 13-item measure rated on a 7-point Likert scale ranging from 1 (“not at all”) to 7 (“very much”) immediately after the game. Students were then invited to voluntary focus group discussions to elicit their attitudes and perceptions of the game. Findings were subject to between-group comparisons and inductive thematic analysis respectively.

Results: A total of 48 (26 control, 22 intervention) clinical-year undergraduates from the Lee Kong Chian School of Medicine in Singapore participated in this study, with 18 continuing to the focus group discussions. Flow Short Scale results indicated the superiority of the intervention group for overall flow (t_{46}=–2.17, P=.04) and the absorption subdomain (t_{46}=–2.6, P=.01). Qualitative results indicated students viewed facilitation as helpful and appropriate, and were able to identify improvable elements of the game’s theoretical foundations and overall design.

Conclusions: While serious games are efficacious means of rehearsing previously learned knowledge, facilitation allows for their efficiency to be greatly increased. Such increases are likely crucial in the coming years with the increased digitization of health care professions education and the prevalence of diabetes.

Trial Registration: ClinicalTrials.gov NCT05637749; https://www.clinicaltrials.gov/study/NCT05637749

(JMIR Serious Games 2024;12:e54703) doi:10.2196/54703

KEYWORDS

serious game; randomized control trial; facilitation; diabetes; diabetes management; flow; education

Introduction

Background

Diabetes is a chronic disease characterized by sustained high blood glucose and, when left unmanaged, is associated with severe health consequences and premature mortality [1]. Globally, an upward trend in the incidence of the disease has been noted in most regions, driven primarily by type 2 diabetes (T2D) and the prevalence of its associated risk factors [2-4].

This coincides with the increasing complexity of medicine and its need for a skilled workforce capable of taking up new knowledge in constrained time frames [5,6], necessitating the development of new methods to facilitate continuous health care professions education [7].

Digital Interventions

Despite numerous digital interventions developed in the wake of these needs, notable gaps remain. Educational interventions remain focused on patients with comparatively little for caregivers and health care professionals. Despite favorable outcomes, such interventions may have difficulties sustaining user attention and engagement [8-10]. Reviews of digital diabetes educational interventions suggest their full potential may be stymied by the absence of a human expert to guide the user [11], though it remains unclear if this is a consequence of the rapid uptake of digital technologies and the greater push toward self-reliance.

Jeon and Park [12] noted that self-care apps improve social motivation and behaviors, but not knowledge, behavioral skills, or personal motivation, which instead benefited from in-person interventions, while caregivers actively sought peer-to-peer support to alleviate concerns [13]. In continuous education, learners paradoxically reported great acceptance of distance learning methodologies, while also desiring face-to-face teaching [14,15]. Additional evidence suggests improving motivation and subject attitudes may be more important than enhancing knowledge [16], as observed in interactions with diabetes care services [16]. As a result, despite the student and patient-centered approaches used by modern developments [17-20], evidence of the exact benefits afforded by including a human expert, trainer, or facilitator in these interventions remains insubstantial.

Serious Games

Serious games may increase patient motivation [21-23], and the management skill of patients and caregivers [12,16,24], in the context of chronic diseases such as diabetes. They are usually defined as games designed for nonrecreational purposes such as education and therapy [25-27], and have enjoyed increasing uptake in both educational and clinical settings due to the ease by which they enhance motivation in their users [25-27]. They are distinct from gamification, which incorporates game-like elements into nongame interventions [28].

Such games are unable to replace qualified professionals, and instead support the promotion of exercise [29,30], deliver diabetes care education [31,32], and facilitate self-management [33,34]. Like their game-free counterparts, they are heavily focused on patients, with few health care professionals or caregivers. The majority appear to be gamified interventions as opposed to serious games, with literature reviews of the past decade returning only a nondigital escape room to improve diabetes management knowledge in pharmacy students and teaching insulin therapy to primary care physicians [35,36].

This dearth likely stems from early attempts to gamify existing methods, which failed due to poor game and instructional design [37,38]. Modern developments are understandably subjected to rigorous validation studies before implementation, and recent reviews of the literature suggest this remains the focus of a vast majority of game-based research—newly developed games are trialed against an established game-free control group, and efficacy is determined by the degree to which the game fulfills its intended purpose [25,39-42]. This user-centric focus, while meritorious, leaves unaddressed the key mechanisms of action responsible for a game's success, much less to what extent such mechanisms may be controlled to influence how players enjoy or learn from them.

Roles of Human Educators

Human educators provide emotional intelligence, empathy, and context awareness [43], key drivers of learner engagement and overarching educational outcomes despite advancements in intervention design [44]. With games, prior investigations suggest educators may facilitate learners transiting from passive knowledge retention and learning to encourage self-directed inquiry and active learning [43], as well as provide customized, empathetic, and learner-specific feedback that digital systems may not fully emulate [45].

Facilitation and Flow

Facilitation is considered by the Promoting Action on Research Implementation in Health Services framework as a process that supports and enables others' self-improvement and goal attainment [46,47]. In medical education, facilitators are credited for the success of collaborative [48], guided, yet autonomous learning experiences such as team-based learning [49,50]. Success may be attributed to facilitators exhibiting prosocial traits such as empathy, flexibility, authenticity, pragmatism, and credibility [51,52], as well as easing difficulties and keeping students invested in the activity [53-55].

Despite this, facilitation remains a broadly defined concept and the benefits of prosocial traits may not wholly translate to serious game-based interventions. Human experts introduced to games may assume multiple roles such as facilitators, instructors, and mentors [56], among others, with no role being universal due to the myriad roles games may play. Nonetheless, facilitation and moderation are the most likely drivers of success in serious games due to their means of adjusting a game activity to better meet the needs of individual learners.

The meeting of these needs is often a precursor of a flow state, a crucial yet often overlooked feature of serious games. Flow is a cognitive state characterized by absolute attention toward an optimally challenging task and the fluency of one's actions, seemingly without conscious thought [57-59]. Notably, high rates of flow are associated with a willingness to return and repeat an activity [25,59-61], critical in education where rehearsal facilitates the committing of new knowledge into long-term memory.

Understanding how much influence facilitation may exert on flow generation in serious games is thus key to increasing the efficacy of such games in medical education and further enabling the continuous education of health care professionals in the future.

Study Aims

To this end, this mixed methods study aims to identify and, where possible, quantify the benefits arising from human-assisted facilitation in digital game-based interventions. This will be accomplished via an open-label superiority randomized controlled trial, then a focus group discussion to elicit greater insight and provide additional context into participant perceptions and attitudes.

A rehearsal-based diabetes management game has been developed for this purpose and includes a special role for a human facilitator tasked with ensuring an optimal game environment for the player.

The quantitative aspect of this study hypothesizes that subjects assigned to a facilitated game group will report statistically significantly higher flow scores than subjects of the unfacilitated group.

Methods

Participants

Subjects were recruited from third-, fourth-, and fifth-year medical students undertaking their Bachelor of Medicine and Bachelor of Surgery degree at the Lee Kong Chian School of Medicine, Nanyang Technological University in Singapore, where this study also takes place. These students were selected due to their completion of the endocrinology segments of their internal medicine clinical postings and thus had basic familiarity with diabetes management in both clinical and community health care settings. Recruitment was performed by email advertisements and snowball sampling via word-of-mouth, and the completely voluntary and benefits-free nature of this study was repeatedly stressed. Exclusions included students who had not completed the endocrinology segments of their internal medicine postings, diseases of the eye not including myopia, noticeable psychosocial difficulties, and any other characteristics that may put them at risk while playing the game. All interested participants were instructed to read this study's information sheet, had the same sheet read to them before consent taking, and were repeatedly informed that they could ask questions and that participation in both the qualitative and quantitative aspects of this study was voluntary.

Ethical Considerations

Ethical approval was obtained from the Nanyang Technological University Institutional Review Board (IRB-2022-739).

Theoretical Bases

Before the formal study, informal focus group discussions with clinical-year medical students were conducted to gauge interest, elicit suggestions, and identify key features of the diabetes management game. Following this and subsequent literature reviews [62], it was determined the game intervention would best be developed based on self-determination theory (SDT), flow theory, and experiential learning theory.

SDT posits that an activity becomes intrinsically motivating when the needs of competence, autonomy, and relatedness are met [63]. Competence was addressed by mirroring the behaviors of both nonplayer characters (NPCs) and their ambulatory glucose profiles (AGPs) as closely as possible to case studies students would encounter as part of their education. To address the need for autonomy, players were permitted to manage NPCs in any manner. Relatedness was expected to be established by both the presence of the facilitator and the role the facilitator plays when checking in on the player's progress.

The facilitator's role also overlaps with flow theory, as they may adjust game difficulty based on real-time player feedback. Flow refers to a deep cognitive state wherein an individual directs absolute attention toward an optimally challenging task, simultaneously experiencing near-complete control over the activity and a total loss of awareness of the self [57-59]. Such states are associated with increased accomplishment across the breadth of the human developmental life span and, in the context of education and rehearsal, the willingness to return and repeat an activity [25,59-61].

Experiential learning theory stipulates that learning occurs when an individual partakes in the activity or task to be learned as opposed to receiving knowledge through instruction [64], and further overlaps with the aforementioned fulfillment of competence as defined by SDT.

Game Intervention

The digital diabetes management game is comprised of a single-player management game centered on a 2D community populated with NPCs who all have type 1 diabetes (T1D), T2D, or gestational diabetes (GD). NPCs work, consume meals, and partake in recreational activities within the game environment of their own accord and may not be directly controlled by the player. The player interacts with the game using the mouse and controls the administration of insulin, snacks, and oral medication. Upon clicking the respective buttons, players are presented with a dosage and may adjust it with further mouse clicks before confirming the action. NPCs do not partake in these activities of their own accord.

Each NPC possesses individually tracked blood glucose, visible to the player via an AGP, with changes simulated in response to stimuli, such as the physical intensity of current activities, insulin and oral medication dosages, consumption of meals or snacks, and phenotypic characteristics such as insulin resistance. Should extended or severe hyper- or hypoglycemia occur, NPCs will faint, be removed from the game, and the player will be informed that said NPC has been evacuated to an off-site hospital. Upon selecting an NPC, players may view their relevant clinical history, present symptoms, and all past actions they were administered (see Figure 1; a short technical demonstration is included in Multimedia Appendix 1).

Figure 1. Overview of the game world as viewed by the player, including sidebar display of relevant NPC details. NPC: nonplayer character.

Following a tutorial with actual gameplay, the player is given 12 minutes to play the game with 1 real-time minute corresponding to 1 in-game hour. They are tasked to keep the blood glucose levels of all NPCs in an ideal target range (subject to phenotype) as much as possible. With every real-time minute, each NPC's AGP is updated based on their in-game activities and the player's inputs to them thus far.

By default, the game begins with 10 NPCs comprising 7 T2D, 2 T1D, and 1 NPC with GD, reflective of the incidence of each phenotype. This number may change based on the actions of a human facilitator, who may access a game in progress from a separate machine. The facilitator is provided the same information as the player and may additionally create new NPCs for the player to manage, remove existing NPCs from play, or as an alternative to removal "freeze" the AGP of existing NPCs such that they need not be managed by the player until "unfrozen" (see Multimedia Appendix 2).

Due to the short duration of gameplay, the role of the facilitator was governed by a strict set of rules that they were not permitted to deviate from during this study. They were not permitted to offer knowledge a player has clearly forgotten (ie, administration of metformin to an NPC with compromised renal function) unless explicitly asked. They were to remind players of the function of game controls if asked, or if the player repeatedly made control-based mistakes (ie, trying to use the right mouse button or keyboard, which have no function). They were to succinctly explain to the player the reason behind an NPC fainting and being conveyed to the hospital should an instance occur (ie, too much bolus insulin 2 hours ago and NPC became hypoglycemic) to avoid disrupting the flow of the game. To minimize the influence of extraneous factors that may result from personal communication styles, the same facilitator was used for all games.

In the first 3 minutes of a facilitated game, the player will play the game solo with the facilitator remaining out of sight and taking no actions. On the third minute, and every 2 minutes after, the facilitator will approach the player and ask, "How are you faring?" The player would then indicate how well they are handling the present difficulty and if they desire a change in the number of active NPCs.

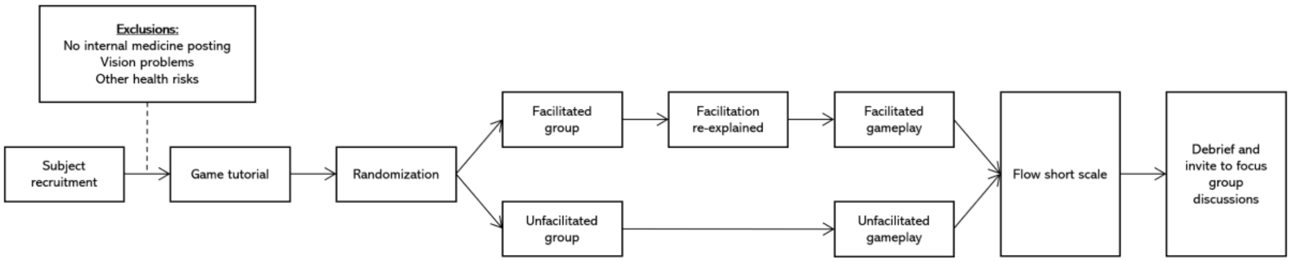
Study Design

This study used the CONSORT (Consolidated Standards of Reporting Trials) as guidelines and was conceptualized as an open-label, superiority randomized controlled trial (Multimedia Appendix 3). Following consent by the on-site member of this study's team, subjects were briefed on their objectives, given the tutorial, and allowed to familiarize themselves with playing the game until they had no further questions. Subjects were then randomized into either the facilitated intervention or facilitator-free control via simple randomization using Sealed Envelope, a secure web-based randomization service based in London, the United Kingdom [65], that allowed for allocations to be concealed from all parties until after a subject was enrolled and ready to partake in the intervention. Aside from the secure password to enable each randomization, subjects were permitted to view the result of their randomization.

Should the subject be randomized into the facilitated intervention group, the role of the facilitator would be repeated to them, and any last-minute questions answered. Otherwise, the facilitator would ensure the start of the game and then exit this study's site until the control's game had elapsed. The game itself was app-based and played on an internet-enabled university laptop belonging to this study's team, with the facilitator remotely joining from a separate laptop on the same network.

Upon conclusion of the gameplay, subjects were immediately administered the Flow Short Scale (FSS), issued an e-voucher as an inconvenience fee, and invited to the focus group discussion (see Figure 2). Subjects who attended the focus group discussion were reminded that the focus group discussions would be recorded for transcription by a third-party transcription company and any subject unwilling to consent again was allowed to leave. Subjects were then shown and allowed to refamiliarize themselves with the game through play. The guiding questions of the focus group comprised: (1) Do you recall becoming really immersed in the game? What were you doing just before? (2) What did the game do to capture and retain your attention for extended time periods? (3) If you could improve the game to make it clearer and more balanced, what would you do?

Figure 2. Overview of the study design.



Flow Short Scale

Following gameplay, subjects of both groups were administered the FSS. The FSS consists of 13 items on a 7-point Likert scale ranging from 1 ("not at all") to 7 ("very much") [66]. The FSS demonstrates good construct validity, psychometric properties, and a stable 3-factor structure comprising fluency of performance, absorption by activity, and perceived importance or outcome importance of said activity [67]. Flow itself comprises the first 10 items and the domains of fluency and absorption [66]. The scale is typically administered immediately after an activity as a retrospective measure of flow in said activity, and was, for this study, hosted on a university-secured Google Forms and transmitted to subjects via a QR code.

Power

Power calculations were performed via Sealed Envelope sample size calculations [68], with an α level of 5%, 90% power, and the anticipated control group means of 4 (0.5 above the mean of 3.5 due to games innately being conduits of flow) and anticipated intervention means of 5.05 (15% higher than the control), and an SD of 1. The increase of 15% was based on the results of prior studies comparing game-based interventions and established nongame controls on the results of the FSS [69,70]. An estimated 22 subjects per group for an overall 44 was expected.

Data Analysis

Quantitative analysis will comprise group comparisons of either discrete or continuous data drawn from FSS and in-game scores. If normally distributed, data will be analyzed via independent samples *t* tests (2-tailed across the board), with Welch correction performed should equality of variances not be observed. Data that are not normally distributed will be compared with Mann-Whitney *U* tests instead. Effect sizes will be calculated for FSS data to better visualize the degree of impact facilitation has on flow generation. Exploratory examinations of all data collected automatically by the diabetes management game will be performed to identify any notable differences between groups. The threshold for statistical significance is set for <.05 per convention, and will be performed using R (version 4.3.1; R Foundation for Statistical Computing).

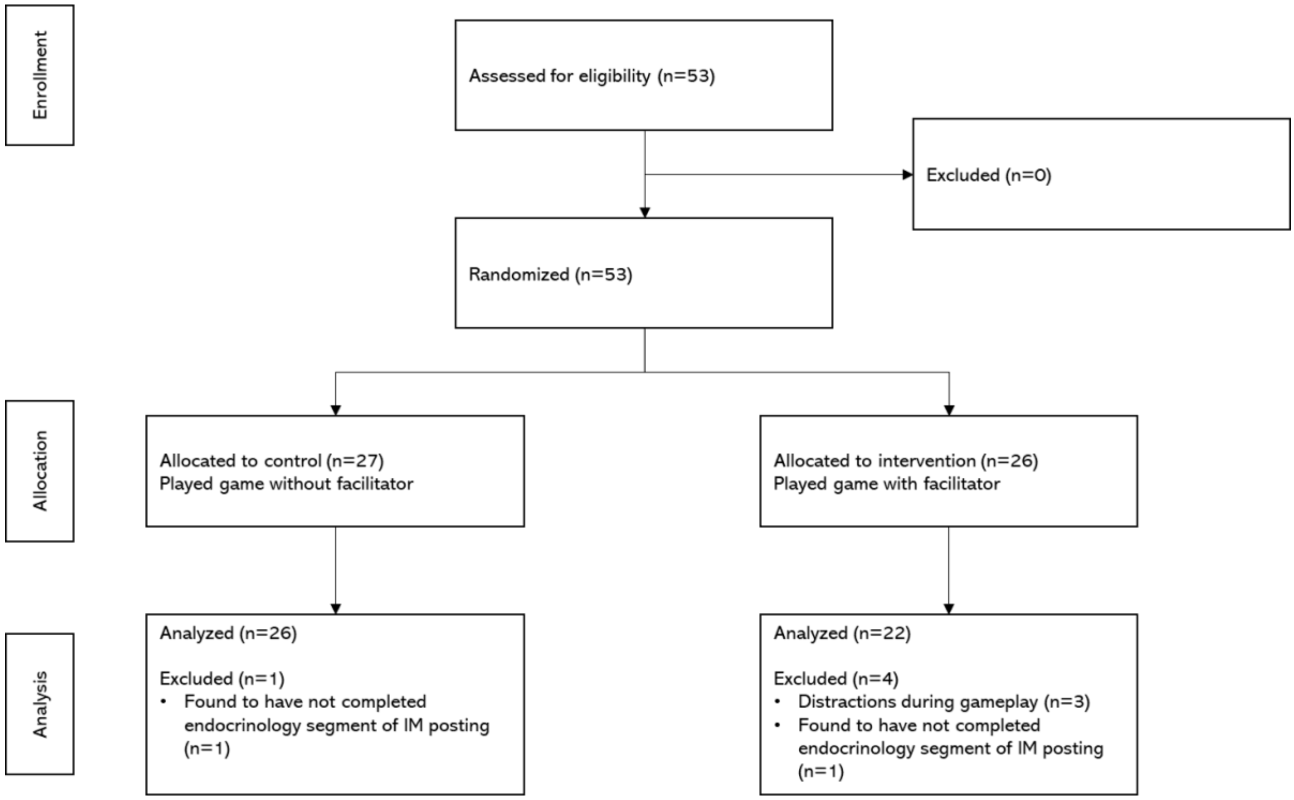
Due to the exploratory aspect of this study and the novelty of the proposed intervention design, transcribed focus group discussions were subject to inductive thematic analysis per the guidelines of Braun and Clarke [71,72]. A reflexive approach was adopted [72], with transcriptions performed by 2 authors (JWT and DKSC) with a third (SRM) acting as a referee. Of the 3, JWT and DKSC have prior histories of playing video games; only JWT plays recreational video games regularly, DKSC no longer plays video games, while SRM has a very limited history of playing video games. To minimize transference of bias, no communication or input was permitted during the initial stages of analyses. Both JWT and DKSC would first read the transcripts until familiarized, then begin a preliminary coding process and generate a series of initial themes. At this stage, prospective codes were deemed to be anything that appeared to be related to student perceptions of advantages afforded by facilitation, disadvantages resulting from its absence, and any other factor not accounted for by the research question but deemed serendipitous by the coders. Upon completion, these initial themes were then individually reviewed against previously identified codes and refined as necessary. Each analysis was then compared and discussed, with all differences highlighted for discussion and resolution with the referee. Themes were then cross-checked to ensure they represented clear patterns, and iteratively reviewed until each possessed a distinct scope with minimal overlap.

Results

Participants

Of the 53 subjects recruited into this study, 5 were eventually excluded due to being visibly distracted (ie, mobile phones) while playing the game or were found to have not actually completed the endocrinology segments of their internal medicine postings. A total of 48 subjects were thus included; comprising 26 control and 22 intervention randomizations (see Figure 3 for CONSORT flowchart).

Figure 3. CONSORT flow diagram for participant allocation. CONSORT: Consolidated Standards of Reporting Trials; IM: internal medicine.



Of the 48 analyzed subjects, there were 25 (52.1%) male and 23 (47.9%) female students aged between 21 and 25 (mean 22.44, SD 1.17) years. Control group subjects comprised 13 male and female students, while intervention group subjects comprised 12 male and 10 female students. The mean ages for both groups were 22.5 (SD 1.36) and 22.4 (SD 0.91) years respectively.

A total of 18 students continued on to the focus group discussions. Of these, 5 male and 5 female students were from the control group, and 5 male and 3 female students were from the intervention. Three sessions were conducted comprising 4 subjects (3 male, 1 female) with 1 control and 3 intervention assignments, 8 subjects (4 male and 4 female) with 6 control and 2 intervention assignments, and 6 subjects (3 male and 3 female), with 3 control and 3 intervention assignments. Table 1 presents an overview of participant demographics for each group in this study.

Table 1. Overview of participant demographics for both the randomized controlled trial and the focus group discussions.

Characteristics	Control	Intervention
Game randomized controlled trial		
Total participants (n)	26	22
Male, n (%)	13 (50)	12 (54.5)
Female, n (%)	13 (50)	10 (45.5)
Age (years), mean (SD)	22.5 (1.36)	22.4 (0.91)
Focus group discussions		
Total participants (n)	10	8
Male, n (%)	5 (50)	5 (62.5)
Female, n (%)	5 (50)	3 (37.5)
Age (years), mean (SD)	21.7 (0.48)	22 (0.76)

Between-Group Comparisons

Table 2 presents a summary of the performance of between-group comparisons. Normal distribution of data and homogeneity of variances were observed. Independent samples *t* tests were performed across FSS data, and results indicated superiority for overall flow (t_{46}=–2.17, *P*=.04), weighted primarily on the absorption subdomain (t_{46}=–2.6, *P*=.01) of the intervention group. No significant differences were observed between the fluency subdomain F and importance (t_{46}=–0.2, *P*=.84). These results suggest a moderate to high degree of flow for both the intervention and control groups, with notably high absorption for the intervention group, and are supported by the moderate to relatively high effect sizes of 0.63 for overall flow and 0.75 for absorption respectively.

Table 2. Summary of *t* tests and effect size calculations between the control (unfacilitated) and intervention (facilitated) conditions.

Variables	Control, mean (SD)	Intervention, mean (SD)	*t* test[a] (*df*)	*P* value	Effect size, Cohen *d*
Overall flow	4.4 (0.89)	4.95 (0.85)	–2.17 (46)	.04	0.63
Fluency	4.03 (1.16)	4.52 (1.06)	–1.5 (46)	.14	0.44
Absorption	4.96 (0.83)	5.6 (0.87)	–2.6 (46)	.01	0.75
Importance	4.26 (1.39)	4.35 (1.78)	–0.2 (46)	.84	0.06
Mean hours of ideal glucose	2.66 (1.36)	2.35 (1.09)	0.86 (46)	.4	N/A[b]
Metformin errors	2.77 (1.88)	2.73 (1.67)	0.08 (46)	.94	N/A
Number of evacuations	1.58 (1.94)	0.64 (0.95)	2.07 (46)	.04	N/A

[a]Two-tailed.

[b]N/A: Not applicable.

Gameplay analytics indicated no significant differences between hours NPCs spent at an ideal blood glucose level (t_{46}=–0.86, *P*=.4) and the number of NPCs administered metformin when contraindicated (t_{46}=0.08, *P*=.94). Ideal blood glucose was defined as 5 to 9 mmol/L for T1D and T2D, and 5 to 7 mmol/L for GD.

A significant difference was noted for the number of NPCs requiring medical evacuation, which occurs when they have extended hyper- or hypoglycemia and faint, in favor of the intervention group (t_{46}=2.07, *P*=.04).

Thematic Analysis

The focus group discussions elicited a broad range of student insights and perceptions, evidence in support of the theoretical foundations that afford such games their efficacy, perceived advantages of human-directed moderation of game difficulty, and suggestions on how to improve the intervention. Focus group discussion transcripts were anonymized to preserve the confidentiality of the research data and all mention of names were removed. Although the sessions were conducted in English, all participants were Singaporean and frequently communicated in Singlish—an English-based creole that, while comprising almost entirely of English words, uses a grammatical structure that deviates heavily from standard English. Sessions were transcribed verbatim to avoid accidental changes in meaning and retained Singlish terms such as “ya,” usually an analogue of “yes” but sometimes occurs as general affirmation, and the ubiquitous particle “lah,” typically found at the end of sentences that, when spoken with an appropriate tone, may modify an utterance akin to the use of adverbs in standard English.

Perceptions on Facilitation

Thematic analysis of the focus group discussions indicated facilitation was mostly helpful, and that students felt a sensation of safety and were more likely to undertake greater challenges as a result.

> *Yeah, I think it was quite… It was sort of like a safety blanket, you know?* [Student 1A]
>
> *I think I just felt like if anything happened I can go to the facilitator, like, hi, can you help? Can you take out one person? That would be like the guy over there.*
>
> *Yeah, same. [facilitator’s name] actually was basically my lifeline when I think about it.* [Student 1B]

Conversely, unfacilitated students experienced increased challenge and performance went down when this was too much for their skill levels, and desired facilitation when this occurred.

> *Okay, so it was challenging, but it was very frustrating, and I didn’t know what I could do to resolve it.* [Student 1C]
>
> *Yeah. I think a facilitator would have been good or at least there would be, like, instructions on the screen, lah.* [Student 1C]

Perceptions of Support for SDT

Support for SDT was deemed as features already present or features that if added would support the theory in the context of meeting the needs of competence, autonomy, and relatedness. Students who perceived themselves struggling with underperformance, actual or otherwise, requested additional modifications to the game beyond what the facilitator was capable of.

> *But yeah, it will be better if there’s, like, a tutorial or something from the easy levels to high levels, like that, yeah.* [Student 1C]

> *And then, after that, it's like, you're frustrated 'cause, like, your course is not really going well.* [Student 1B]

Feelings of autonomy were noted to already be present due to the numerous means of resolving problems and that actions were free of true consequences.

> *But because I didn't feel like there was any serious consequence, because it was a game, so I thought it was quite fun.* [Student 2D]

Perceptions of relatedness were most prevalent during attempts to involve peers as fellow participants and included comparisons to popular cooperative recreational games.

> *So, instead of it being confined to just the cafeteria and the outdoor exercise area, we could have the opportunity to explore more places…*
>
> *I'm thinking like an Overcooked kind of thing, like, different islands.* [Students 1A and 1B]

Perceptions of Support for Flow Theory

Discussions of the game activity suggested students who were facilitated were more likely to experience an altered perception of time despite there being a clock in the game.

> *I think for me, I didn't really care too much about the time. So, like, when [facilitator's name] stopped me, I eh 12 minutes already?* [Student 3B]
>
> *But it was a fun experience. I felt engaged, because every minute I would check everyone's [blood glucose]. So, I did not realise, like, that time had passed.* [Student 2D]

When queried, students were retrospectively aware of becoming completely absorbed in the activity to the point of forgetting about the facilitator's presence, despite the regular check-ins.

> *It's like I don't have the mental capacity to focus on anything else.* [Student 3C]
>
> *I think I completely forgot that I can ask the facilitator questions.* [Student 3A]
>
> *I just kept clicking around each patient to see where it was going, and the threads, and whatnot. I think that that's what really kept my attention most of the time…[Student 2C]*

Perceptions of Game Design Elements

Students generally perceived the game as fun, enjoyable, and an appropriate means of revising diabetes management knowledge. The intervention was perceived as both challenging and a safe space in which to commit mistakes harmlessly.

> *Especially fun cause there's the whole threat of them possibly dying in the hospital makes it, like, more exciting and more fun to play.* [Student 2C]
>
> *But because I didn't feel like there was any serious consequence, because it was a game, so I thought it was quite fun.* [Student 2D]

Student discussions frequently resulted in feedback and disagreements on the merits of said feedback were likely evidence of the specific needs of students playing the game.

> *I think they should stop moving. Like, moving doesn't help anything and it doesn't add anything.*
>
> *I like the moving though.*
>
> *The moving was fun lah you just keep chasing the guy around.* [Students 1B, 1A, and 1D]

Despite not being prompted, students were able to raise requests for changes to better align the game with SDT and flow theory. Changes in line with SDT from the game design perspective primarily focused on being able to play the game with other people and meet the need for relatedness.

> *Like, you can play with a friend… Unless, I don't know, there's some multi-player function introduced.* [Student 1A]

Changes in line with flow theory focused on how the game should have better-presented information to students, ranging from succinct to full and detailed explanations.

> *Maybe at the start, before you start playing, that there's a screen that shows everybody with all their conditions.* [Student 3E]
>
> *So, either working on a different way of showing they were thirsty, like maybe an icon that shows that they're thirsty instead…* [Student 2G]

Additional Findings

Although not the focus of this study, it was noted that certain student characteristics may exert some influence over the degree they engage with the game activity and facilitator. Further, 1 student indicated altruistic motives as a driver of engagement.

> *I was pretty immersed in the game, and especially with the fact when the people started dying and getting hospitalised. I think, like, when… Once that's happening, then, yeah, like, oh no, and then you feel more immersed in the game, because you want to keep everyone else alive.* [Student 3E]

Students who appeared more forgetful than their peers were also likely to express frustration that inhibits engagement.

> *…apparently the endocrinology emphasised that during multiple tutorials, but I don't have any recollection of that at all.* [Student 3B]

Discussion

Principal Findings

Results from the between-comparisons indicate support for the hypothesis; the facilitated group is superior, based on the moderate to fairly large effect sizes, to the control in terms of overall flow and the subdomain of absorption, but not fluency. Analyses of the focus group discussions suggest that, beyond flow, ideal conditions for flow were supported by the perception of safety and its related willingness to push oneself toward greater challenge. These findings were unlikely results of differing competencies between groups, evidenced by the nonsignificant differences in time NPCs spent at healthy blood glucose and the number of inappropriate administrations of metformin. Additionally, students of the intervention group almost universally forgot they were allowed to clarify the effects of medication and refrained from doing so as a result. This forgetting to ask for help renders the lower rate of medical evacuation to be most likely the result of the intervention group's difficulty adjustment as opposed to the facilitator reminding students of the effects of medication. This is likely the result of intense concentration on the game activity resulting from a high flow state as indicated by student reports of total attention being given to the activity. The higher flow scores of the intervention group also offer support to the notion that facilitation confers tangible benefits that result in increased engagement [56,73]. Due to substantial correlations between flow and intrinsic motivation [74], it is likely that students would be more willing to engage in a facilitated serious game due to interest and its enjoyability, and thus be more likely to re-engage in the activity without the need for an incentive [63].

Analyses of the focus group discussions have also indicated substantial support for flow theory and SDT as theoretical foundations of serious games in this design. Even when unprompted, students frequently requested the modification of game features that would circumvent a specific difficulty they experienced, only for other students to disagree with the merits of said requests. This both highlights the facilitator's role in helping students circumvent specific difficulties, and flow theory's need for a balance between the challenge of the activity and the learner's perceived skill [58]. Similarly, reports of feeling safe when paired with a facilitator likely stem from the need for relatedness as defined by SDT. Due to students' unfamiliarity with the facilitator, this is likely in the context of a mentor-mentee relationship as opposed to friendships [73,75] and is likely not observable in serious games featuring dynamic game balancing as the sole option for difficulty modulation [76]. While it could not be readily determined from the focus group discussions, there is a possibility that perceptions of safety stemmed from fulfilling the need for autonomy, due to a facilitated game affording students a means of exercising greater control over their learning environment [56,77].

While the use of games in medical education and training is not new, an embedded role for a facilitator remains uncommon even in nonmedical literature, with self-selected or dynamic game balancing remaining the common form of difficulty modulation [76]. As a result, the design of the present intervention appears unique to the best of the authors' knowledge in examinations of both gray and peer-reviewed literature. The inclusion of a facilitator remains beneficial particularly to beginners and players exhibiting low confidence, as evidenced by student perceptions of a "safety blanket." Comparable studies include a diabetes education and self-management study that paired young patients with mentors to alleviate the emotional stresses of adolescence and reduced the socioeconomic costs of the disease [75], a game-based learning tool for children with content that could easily be modified by an educator resulted in greatly increased student engagement and willingness to participate [78], and a qualitative study that suggested the facilitators required a mix of managerial and technical skills to blend away difficulties faced by students such that they may fully engage in a game-based activity [56]. This study's design nonetheless aligns with facilitation's role to support, give, and encourage learners as opposed to teaching the content in question, itself key to simulation and game-based interventions to which the diabetes management game belongs [79]. In addition, the results suggest some relation to the sociocultural theory of cognitive development as defined by Vygotsky, which posits that learning is a social process occurring primarily through interactions between a learner and an expert mentor [80]. Though the theory was not central to the development of the intervention, it may imply that learners used to learning with facilitators may be more receptive to the intervention than those used to learning on their own.

Serious games for diabetes management are almost universally directed at those affected by the disease and understandably target behavioral change as the ultimate goal of education [81]. This study presents one of the few serious games for use in health care education that includes an option for facilitation, itself understudied and uncommon to games even beyond the health care setting [56,79], but is limited by a lack of suitable comparators in the medical literature. In engineering education, it was noted that no 1-model decided the best means of facilitating a serious game, but that learning tended toward being experiential in nature [56], limiting its comparable applicability with the content of the diabetes management game, which falls primarily between the "Knows" and "Knows How" tiers of Miller's pyramid [82,83]. However, due to the clear distinction between the role of the facilitator and the game intervention itself, it should be possible for the facilitator's role to be generalized to other topics within health professions education. This may be further supported by the game's focus on the lower tiers of Miller's pyramid, and that the standalone game may be played relatively effectively without facilitation.

Finally, though the intervention was intended as a supplementary rehearsal tool, it appears to be a suitable means of formative assessment when played without facilitation [84], and for the rapid detection of learning gaps and their prevalence in a cohort.

Limitations

This study's inclusion of a dedicated role for a facilitator in a rehearsal-focused game-based intervention appears to be unique. This role, and its ability to moderate a learner's gameplay in real-time, does not feature in the recent literature and limits comparisons. This study's focus on ascertaining the benefits of facilitation in the context of a rehearsal-based game for diabetes management knowledge is itself a limitation, for the exact long-term effects on the topic cannot be determined without a follow-up long-term study involving a version of the game refined after player feedback. Though sufficiently powered and based on a strong theoretical foundation, this study nonetheless presents a sample drawn from a single institution and cohort of students.

Similarly, triangulating the findings via studies involving the same methods of facilitation, but with different topics within health professions education, will help determine the long-term effects of facilitation and its benefits to medicine. The semiexploratory nature of the intervention, and the use of a one-to-one facilitator-student ratio means the intervention cannot be sustainably upscaled without first determining an optional means of increasing the ratio, limiting its deployment to smaller scales. Finally, an element of self-selection bias favoring students with preexisting interests in video games may have been present due to the voluntary nature of this study.

Conclusions

The inclusion of a facilitator in a rehearsal-based medical serious game can increase the degree to which a student may engage in an activity and elicit sensations of safety with the corresponding willingness to embrace greater challenge. The benefits appear particularly notable for participants who are beginners or unconfident in their abilities and are likely to be the result of facilitators easing difficulties to greater align the participant with the conditions of flow and SDT.

Acknowledgments

This project was supported by the Games for Health Innovations Centre (ALIVE) Serious Games Grant. We would like to thank Dr Han Siew Ping and Dr Olivia Ng Xie Wei for their assistance with the focus group discussions.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

JWT, RD, and SRM were involved in the conceptualization of this study, its methodology, and funding acquisition. JWT, GT, XL, RD, and SRM were involved in the design of the software. JWT and DKSC were involved in the data curation and analysis with supervision from SRM and PR. All authors were involved in writing the original draft, and reviewing and editing this paper.

Conflicts of Interest

SRM is the Associate Editor for *JMIR Medical Education*. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Technical demonstration of the diabetes management game's control system. All participants will interact with the game in this manner.
[MP4 File (MP4 Video), 32725 KB - games_v12i1e54703_app1.mp4]

Multimedia Appendix 2
A sample video demonstrating the freezing, unfreezing, removal, and addition of a non-player character (NPC) by a facilitator. Participants assigned to the facilitated gameplay will experience these.
[MP4 File (MP4 Video), 12660 KB - games_v12i1e54703_app2.mp4]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 9478 KB - games_v12i1e54703_app3.pdf]

References

1. Heald AH, Stedman M, Davies M, Livingston M, Alshames R, Lunt M, et al. Estimating life years lost to diabetes: outcomes from analysis of national diabetes audit and office of national statistics data. Cardiovasc Endocrinol Metab 2020;9(4):183-185 [FREE Full text] [doi: 10.1097/XCE.0000000000000210] [Medline: 33225235]
2. Chan JCN, Lim LL, Wareham NJ, Shaw JE, Orchard TJ, Zhang P, et al. The lancet commission on diabetes: using data to transform diabetes care and patient lives. Lancet 2020;396(10267):2019-2082. [doi: 10.1016/S0140-6736(20)32374-6] [Medline: 33189186]
3. Magliano DJ, Islam RM, Barr ELM, Gregg EW, Pavkov ME, Harding JL, et al. Trends in incidence of total or type 2 diabetes: systematic review. BMJ 2019;366:l5003 [FREE Full text] [doi: 10.1136/bmj.l5003] [Medline: 31511236]
4. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022;183:109119. [doi: 10.1016/j.diabres.2021.109119] [Medline: 34879977]
5. Smolle J. Complexity and medical education. Saf Health 2016;2(1):16. [doi: 10.1186/s40886-016-0051-4]
6. Woodruff JN. Accounting for complexity in medical education: a model of adaptive behaviour in medicine. Med Educ 2019;53(9):861-873. [doi: 10.1111/medu.13905] [Medline: 31106901]
7. Baddeley AD. Human Memory: Theory and Practice, Revised Edition. Hove, East Sussex: Psychology Press; Apr 01, 1997.
8. Adu MD, Malabu UH, Malau-Aduli AEO, Malau-Aduli BS. Users' preferences and design recommendations to promote engagements with mobile apps for diabetes self-management: multi-national perspectives. PLoS One 2018;13(12):e0208942 [FREE Full text] [doi: 10.1371/journal.pone.0208942] [Medline: 30532235]

9. He Q, Zhao X, Wang Y, Xie Q, Cheng L. Effectiveness of smartphone application-based self-management interventions in patients with type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. J Adv Nurs 2022;78(2):348-362. [doi: 10.1111/jan.14993] [Medline: 34324218]
10. Kayyali R, Peletidi A, Ismail M, Hashim Z, Bandeira P, Bonnah J. Awareness and use of mHealth apps: a study from England. Pharmacy (Basel) 2017;5(2):33 [FREE Full text] [doi: 10.3390/pharmacy5020033] [Medline: 28970445]
11. Reuter J, Ferreira DM, Amorim M, Madaleno M, Veloso C. Educators as facilitators of game-based learning. In: Handbook of Research on Acquiring 21st Century Literacy Skills Through Game-Based Learning. Pennsylvania, United States: IGI Global; 2022.
12. Jeon E, Park HA. Experiences of patients with a diabetes self-care app developed based on the information-motivation-behavioral skills model: before-and-after study. JMIR Diabetes 2019;4(2):e11590 [FREE Full text] [doi: 10.2196/11590] [Medline: 30998218]
13. Oser TK, Oser SM, McGinley EL, Stuckey HL. A novel approach to identifying barriers and facilitators in raising a child with type 1 diabetes: qualitative analysis of caregiver blogs. JMIR Diabetes 2017;2(2):e27 [FREE Full text] [doi: 10.2196/diabetes.8966] [Medline: 30291073]
14. Bin Mubayrik HF. Exploring adult learners' viewpoints and motivation regarding distance learning in medical education. Adv Med Educ Pract 2020;11:139-146 [FREE Full text] [doi: 10.2147/AMEP.S231651] [Medline: 32110134]
15. Setia S, Tay JC, Chia YC, Subramaniam K. Massive open online courses (MOOCs) for continuing medical education—why and how? Adv Med Educ Pract 2019;10:805-812 [FREE Full text] [doi: 10.2147/AMEP.S219104] [Medline: 31572042]
16. Captieux M, Pearce G, Parke HL, Epiphaniou E, Wild S, Taylor SJC, et al. Supported self-management for people with type 2 diabetes: a meta-review of quantitative systematic reviews. BMJ Open 2018;8(12):e024262 [FREE Full text] [doi: 10.1136/bmjopen-2018-024262] [Medline: 30552277]
17. Ang ET, Chan JM, Gopal V, Li Shia N. Gamifying anatomy education. Clin Anat 2018;31(7):997-1005. [doi: 10.1002/ca.23249] [Medline: 30168609]
18. Ramnanan CJ, Pound LD. Advances in medical education and practice: student perceptions of the flipped classroom. Adv Med Educ Pract 2017;8:63-73. [doi: 10.2147/amep.s109037]
19. Chen F, Lui AM, Martinelli SM. A systematic review of the effectiveness of flipped classrooms in medical education. Med Educ 2017;51(6):585-597. [doi: 10.1111/medu.13272] [Medline: 28488303]
20. Poeppelman RS, Liebert CA, Vegas DB, Germann CA, Volerman A. A narrative review and novel framework for application of team-based learning in graduate medical education. J Grad Med Educ 2016;8(4):510-517 [FREE Full text] [doi: 10.4300/JGME-D-15-00516.1] [Medline: 27777660]
21. Von Bargen T, Zientz C, Haux R. Gamification for mHealth—a review of playful mobile healthcare. Stud Health Technol Inform 2014;202:225-228. [Medline: 25000057]
22. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
23. Miller AS, Cafazzo JA, Seto E. A game plan: gamification design principles in mHealth applications for chronic disease management. Health Inf J 2016;22(2):184-193 [FREE Full text] [doi: 10.1177/1460458214537511] [Medline: 24986104]
24. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed e-Health 2012;18(3):175-184 [FREE Full text] [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
25. Gorbanev I, Agudelo-Londoño S, González RA, Cortes A, Pomares A, Delgadillo V, et al. A systematic review of serious games in medical education: quality of evidence and pedagogical strategy. Med Educ Online 2018;23(1):1438718 [FREE Full text] [doi: 10.1080/10872981.2018.1438718] [Medline: 29457760]
26. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. Comput Educ 2012;59(2):661-686. [doi: 10.1016/j.compedu.2012.03.004]
27. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012;99(10):1322-1330. [doi: 10.1002/bjs.8819] [Medline: 22961509]
28. Stieglitz S. In: Stieglitz S, Lattemann C, Robra-Bissantz S, Zarnekow R, Brockmann T, editors. Gamification: using game elements in serious contexts. Cham: Springer International Publishing; 2016.
29. Anderson-Hanley C, Arciero PJ, Westen SC, Nimon J, Zimmerman E. Neuropsychological benefits of stationary bike exercise and a cybercycle exergame for older adults with diabetes: an exploratory analysis. J Diabetes Sci Technol 2012;6(4):849-857 [FREE Full text] [doi: 10.1177/193229681200600416] [Medline: 22920811]
30. Kempf K, Martin S. Autonomous exercise game use improves metabolic control and quality of life in type 2 diabetes patients - a randomized controlled trial. BMC Endocr Disord 2013;13:57 [FREE Full text] [doi: 10.1186/1472-6823-13-57] [Medline: 24321337]
31. DeShazo J, Harris L, Turner A, Pratt W. Designing and remotely testing mobile diabetes video games. J Telemed Telecare 2010;16(7):378-382. [doi: 10.1258/jtt.2010.091012] [Medline: 20679406]
32. Fuchslocher A, Niesenhaus J, Krämer N. Serious games for health: an empirical study of the game “Balance” for teenagers with diabetes mellitus. Entertainment Comput 2011;2(2):97-101. [doi: 10.1016/j.entcom.2010.12.001]

33. Johnson C, Feinglos M, Pereira K, Hassell N, Blascovich J, Nicollerat J, et al. Feasibility and preliminary effects of a virtual environment for adults with type 2 diabetes: pilot study. JMIR Res Protoc 2014;3(2):e23 [FREE Full text] [doi: 10.2196/resprot.3045] [Medline: 24713420]
34. Rosal MC, Heyden R, Mejilla R, Capelson R, Chalmers KA, Rizzo DePaoli M, et al. A virtual world versus face-to-face intervention format to promote diabetes self-management among african American women: a pilot randomized clinical trial. JMIR Res Protoc 2014;3(4):e54 [FREE Full text] [doi: 10.2196/resprot.3412] [Medline: 25344620]
35. Eukel HN, Frenzel JE, Cernusca D. Educational gaming for pharmacy students—design and evaluation of a diabetes-themed escape room. Am J Pharm Educ 2017;81(7):6265 [FREE Full text] [doi: 10.5688/ajpe8176265] [Medline: 29109566]
36. Diehl LA, Souza RM, Alves JB, Gordan PA, Esteves RZ, Jorge MLSG, et al. InsuOnline, a serious game to teach insulin therapy to primary care physicians: design of the game and a randomized controlled trial for educational validation. JMIR Res Protoc 2013;2(1):e5 [FREE Full text] [doi: 10.2196/resprot.2431] [Medline: 23612462]
37. Becker K, Parker JR. The guide to computer simulations and games. British Columbia: Wiley Publishing; 2011.
38. Cindy I. Instructional-design theories and models, volume III: building a common knowledge base. Palmerston North: International Forum of Educational Technology & Society; 2010:219-221.
39. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. J Med Internet Res 2019;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
40. Ravyse WS, Seugnet Blignaut A, Leendertz V, Woolner A. Success factors for serious games to enhance learning: a systematic review. Virtual Reality 2016;21(1):31-58. [doi: 10.1007/s10055-016-0298-4]
41. Hays RT. The effectiveness of instructional games: a literature review and discussion. 2005. URL: https://apps.dtic.mil/sti/tr/pdf/ADA441935.pdf [accessed 2024-06-25]
42. Haoran G, Bazakidi E, Zary N. Serious games in health professions education: review of trends and learning efficacy. Yearb Med Inform 2019;28(1):240-248 [FREE Full text] [doi: 10.1055/s-0039-1677904] [Medline: 31022747]
43. Wouters P, van Nimwegen C, van Oostendorp H, van der Spek ED. A meta-analysis of the cognitive and motivational effects of serious games. J Educ Psychol 2013;105(2):249-265. [doi: 10.1037/a0031311]
44. Passarelli M, Dagnino F, Earp J, Manganello F, Persico D, Pozzi F, et al. Educational games as a motivational tool: considerations on their potential and limitations. 2019 Presented at: 11th International Conference on Computer Supported Education; January 1, 2018; The Netherland. [doi: 10.5220/0007586503300337]
45. Erhel S, Jamet E. Digital game-based learning: impact of instructions and feedback on motivation and learning effectiveness. Comput Educ 2013;67:156-167. [doi: 10.1016/j.compedu.2013.02.019]
46. Helfrich CD, Damschroder LJ, Hagedorn HJ, Daggett GS, Sahay A, Ritchie M, et al. A critical synthesis of literature on the promoting action on research implementation in health services (PARIHS) framework. Implement Sci 2010;5:82 [FREE Full text] [doi: 10.1186/1748-5908-5-82] [Medline: 20973988]
47. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the expert recommendations for implementing change (ERIC) project. Implement Sci 2015;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
48. Hong JM, Rajalingam P. Geographic trends in team-based learning (TBL) research and implementation in medical schools. Health Professions Educ 2020;6(1):47-60. [doi: 10.1016/j.hpe.2019.11.005]
49. Yang L, Rajalingam P. Are two teachers better than one? team teaching in TBL. Med Sci Educ 2020;30(1):643-647 [FREE Full text] [doi: 10.1007/s40670-019-00828-9] [Medline: 34457717]
50. Gullo C, Ha TC, Cook S. Twelve tips for facilitating team-based learning. Med Teach 2015;37(9):819-824 [FREE Full text] [doi: 10.3109/0142159X.2014.1001729] [Medline: 25665624]
51. Dogherty EJ, Harrison MB, Graham ID, Vandyk AD, Keeping-Burke L. Turning knowledge into action at the point-of-care: the collective experience of nurses facilitating the implementation of evidence-based practice. Worldviews Evid Based Nurs 2013;10(3):129-139 [FREE Full text] [doi: 10.1111/wvn.12009] [Medline: 23796066]
52. Harvey G, Loftus-Hills A, Rycroft-Malone J, Titchen A, Kitson A, McCormack B, et al. Getting evidence into practice: the role and function of facilitation. J Adv Nurs 2002;37(6):577-588. [doi: 10.1046/j.1365-2648.2002.02126.x] [Medline: 11879422]
53. Wallin L, Målqvist M, Nga NT, Eriksson L, Persson LA, Hoa DP, et al. Implementing knowledge into practice for improved neonatal survival; a cluster-randomised, community-based trial in Quang Ninh province, Vietnam. BMC Health Serv Res 2011;11:239 [FREE Full text] [doi: 10.1186/1472-6963-11-239] [Medline: 21951770]
54. Berta W, Cranley L, Dearing JW, Dogherty EJ, Squires JE, Estabrooks CA. Why (we think) facilitation works: insights from organizational learning theory. Implement Sci 2015;10:141 [FREE Full text] [doi: 10.1186/s13012-015-0323-0] [Medline: 26443999]
55. Cranley LA, Cummings GG, Profetto-McGrath J, Toth F, Estabrooks CA. Facilitation roles and characteristics associated with research use by healthcare professionals: a scoping review. BMJ Open 2017;7(8):e014384 [FREE Full text] [doi: 10.1136/bmjopen-2016-014384] [Medline: 28801388]

56. Hauge JB, Söbke H, Bröker T, Lim T, Luccini AM, Kornevs M, et al. Current competencies of game facilitators and their potential optimization in higher education: multimethod study. JMIR Serious Games 2021;9(2):e25481 [FREE Full text] [doi: 10.2196/25481] [Medline: 33949956]
57. Csikszentmihalyi M. Beyond boredom and anxiety. San Francisco: Jossey-Bass Publishers; 1975:231.
58. Csikszentmihalyi M. Flow : the psychology of optimal experience. New York: Harper & Row; 1990:303.
59. Engeser S. Advances in flow research. New York, NY: Springer-Verlag; 2012.
60. Csikszentmihalyi M. Applications of flow in human development and education: the collected works of Mihaly Csikszentmihalyi. Dordrecht: Springer Netherlands; 2014.
61. Kiili K, Perttula A, Lindstedt A, Arnab S, Suominen M. Flow experience as a quality measure in evaluating physically activating collaborative serious games. Int J Serious Games 2014;1(3):35-49. [doi: 10.17083/ijsg.v1i3.23]
62. Krath J, Schürmann L, von Korflesch HFO. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comput Human Behav 2021;125:106963. [doi: 10.1016/j.chb.2021.106963]
63. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
64. Kolb DA. Experiential learning: experience as the source of learning and development. US: FT press; 2014.
65. sealed envelope. Simple randomisation service. sealed envelope. URL: https://www.sealedenvelope.com/simple-randomiser/v1/ [accessed 2024-06-25]
66. Rheinberg F, Vollmeyer R, Engeser S. Assessment of flow experiences. In: Stiensmeier-Pelster J, Rheinberg F, editors. Diagnosis of Motivation and Self-Concept (Tests and Trends NF2). Göttingen: Hogrefe; 2003:261-279.
67. Kyriazos TA, Stalikas A, Prassa K, Galanakis M, Flora K, Chatzilia V. The flow short scale (FSS) dimensionality and what MIMIC shows on heterogeneity and invariance. Psychology 2018;09(06):1357-1382. [doi: 10.4236/psych.2018.96083]
68. Power (sample size) calculators London, United Kingdom: sealed envelope ltd. Sealed Envelope Ltd. 2022. URL: https://www.sealedenvelope.com/ [accessed 2024-06-22]
69. Tan JW, Ng KB, Mogali SR. An exploratory digital board game approach to the review and reinforcement of complex medical subjects like anatomical education: cross-sectional and mixed methods study. JMIR Serious Games 2022;10(1):e33282 [FREE Full text] [doi: 10.2196/33282] [Medline: 35006080]
70. Tan JW, Chong DKS, Ng KB, Car LT, Mogali SR. Rehearsal-based digital serious boardgame versus a game-free e-learning tool for anatomical education: quasi-randomized controlled trial. Anat Sci Educ 2023;16(5):830-842. [doi: 10.1002/ase.2286] [Medline: 37166085]
71. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa] [Medline: 36656762]
72. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2020;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
73. Taylor ASA. Facilitation matters: a framework for instructor-led serious gaming. University of Skövde. Facilitation matters: A framework for instructor-led serious gaming: University of Skövde; 2014. URL: https://www.diva-portal.org/smash/get/diva2:745056/FULLTEXT01.pdf [accessed 2024-06-25]
74. Keller J, Ringelhan S, Blomann F. Does skills–demands compatibility result in intrinsic motivation? Experimental test of a basic notion proposed in the theory of flow-experiences. J Positive Psychol 2011;6(5):408-417. [doi: 10.1080/17439760.2011.604041]
75. Kim KH, Lee DY. The challenges of diabetes education for young diabetes patients: a 1:1 mentoring program teaching method. J Korean Diabetes 2017;18(1):49-52. [doi: 10.4093/jkd.2017.18.1.49]
76. Andrade G, Ramalho G, Gomes A, Corruble V. Dynamic game balancing: an evaluation of user satisfaction. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment 2021;2(1):3-8. [doi: 10.1609/aiide.v2i1.18739]
77. Kortmann R, Peters V. Demystifying the unseen helmsman: towards a competency model for game facilitators. 2017. URL: https://research.tudelft.nl/en/publications/demystifying-the-unseen-helmsman-towards-a-competency-model-for-g [accessed 2024-06-22]
78. Scarlatos LL, Scarlatos T. Teacher directed active learning games. J Educ Technol Syst 2008;37(1):3-18. [doi: 10.2190/et.37.1.b]
79. Leigh E, Spindler L. Congruent facilitation of simulations and games. In: Shiratori R, Arai K, Kato F, editors. Gaming, Simulations, and Society: Research Scope and Perspective. Tokyo: Springer Tokyo; 2005:189-198.
80. Vygotsky L. Sociocultural theory of cognitive development. Hillsdale, NJ: Erlbaun; 1987.
81. Martos-Cabrera MB, Membrive-Jiménez MJ, Suleiman-Martos N, Mota-Romero E, Cañadas-De la Fuente GA, Gómez-Urquiza JL, et al. Games and health education for diabetes control: a systematic review with meta-analysis. Healthcare (Basel) 2020;8(4):399 [FREE Full text] [doi: 10.3390/healthcare8040399] [Medline: 33066372]
82. Miller GE. The assessment of clinical skills/competence/performance. Acad Med 1990;65(9 Suppl):S63-S67. [doi: 10.1097/00001888-199009000-00045] [Medline: 2400509]

83. Witheridge A, Ferns G, Scott-Smith W. Revisiting Miller's pyramid in medical education: the gap between traditional assessment and diagnostic reasoning. Int J Med Educ 2019;10:191-192 [FREE Full text] [doi: 10.5116/ijme.5d9b.0c37] [Medline: 31655795]
84. Caballero-Hernández JA, Palomo-Duarte M, Dodero JM. Skill assessment in learning experiences based on serious games: a systematic mapping study. Comput Educ 2017;113:42-60. [doi: 10.1016/j.compedu.2017.05.008]

Abbreviations

AGP: ambulatory glucose profile
CONSORT: Consolidated Standards of Reporting Trials
FSS: Flow Short Scale
GD: gestational diabetes
NPC: nonplayer character
SDT: self-determination theory
T1D: type 1 diabetes
T2D: type 2 diabetes

Edited by G Eysenbach, T Leung; submitted 20.11.23; peer-reviewed by K Erenli; comments to author 18.12.23; revised version received 02.01.24; accepted 31.05.24; published 17.07.24.

Please cite as:
Tan JW, Tan G, Lian X, Chong DKS, Rajalingam P, Dalan R, Mogali SR
Impact of Facilitation on Cognitive Flow in a Novel Diabetes Management Rehearsal Game for Health Professions Education: Mixed Methods, Open-Label, Superiority Randomized Controlled Trial
JMIR Serious Games 2024;12:e54703
URL: https://games.jmir.org/2024/1/e54703
doi:10.2196/54703
PMID:38900700

Original Paper

Comparing Learning Outcomes of Machine-Guided Virtual Reality–Based Training With Educator-Guided Training in a Metaverse Environment: Randomized Controlled Trial

Dilek Kitapcioglu[1], MD, PhD; Mehmet Emin Aksoy[1,2*], MSc, MD, PhD; Arun Ekin Ozkan[1*], MSci; Tuba Usseli[3*], MSSc

[1]Center of Advanced Simulation and Education (CASE), Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

[2]Department of Biomedical Device Technology, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

[3]Vocational School for Anaesthesiology Technicians, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

[*]these authors contributed equally

Corresponding Author:
Mehmet Emin Aksoy, MSc, MD, PhD
Department of Biomedical Device Technology
Acibadem Mehmet Ali Aydinlar University
Kayisdagi cad. No:32 Atasehir
Istanbul, 34752
Turkey
Phone: 90 05052685158
Email: dreminaksoy@gmail.com

Abstract

Background: Virtual reality (VR) modules are commonly used for health care training, such as adult advanced cardiac life support (ACLS), due to immersion and engagement. The metaverse differs from current VR serious gaming by enabling shared social connections, while current VR modules focus on computer-based content without social interaction. Educators in the metaverse can foster communication and collaboration during training sessions.

Objective: This study aimed to compare learning outcomes of VR-based, machine-guided training with educator-guided, VR-based training in the metaverse environment.

Methods: A total of 62 volunteered students from Acibadem Mehmet Ali Aydinlar University Vocational School for Anesthesiology were randomly divided into 2 groups of 31 participants each: one group received VR-based training with machine guidance (MG), and the other received VR-based training with educator guidance (EG) in the metaverse. The members of both groups undertook VR-based basic training for ACLS. Afterward, the MG group was trained with a VR-based advanced training module, which provides training with full MG, whereas the EG group attended the VR-based, educator-guided training in the metaverse. The primary outcome of the study was determined by the exam score of the VR-based training module. Descriptive statistics defined continuous variables such as VR exam scores and time spent on machine- or educator-guided training. The correlation between training time and VR exam scores was assessed with the Spearman rank correlation, and nonnormally distributed variables were compared using the Mann-Whitney *U* test. Statistical significance was set at $P<.05$, with analyses executed by MedCalc Statistical Software (version 12.7.7).

Results: Comparing the VR test scores between the MG and EG groups revealed no statistically significant difference. The VR test scores for the EG group had a median of 86 (range 11-100). In contrast, the MG group scores had a median of 66 (range 13-100; $P=.08$). Regarding the correlation between the duration of machine-guided or educator-guided training and VR-based exam scores, for the MG group, =0.569 and $P=.005$ were obtained. For the EG group, this correlation was found to be =0.298 and $P=.10$. While this correlation is statistically significant for the MG group, it is not significant for the EG group. The post hoc power analysis (80%), considering the correlation between the time spent on training and exam scores, supported this finding.

Conclusions: The results of this study suggest that a well-designed, VR-based serious gaming module with MG could provide comparable learning outcomes to VR training in the metaverse with EG for adult ACLS training. Future research with a larger sample size could explore whether social interaction with educators in a metaverse environment offers added benefits for learners.

Trial Registration: ClinicalTrials.gov NCT06288087; https://clinicaltrials.gov/study/NCT06288087

(JMIR Serious Games 2024;12:e58654) doi:10.2196/58654

KEYWORDS

metaverse; serious gaming; virtual reality; educator guidance; educator; learning; machine guided; VR; guided training; randomized controlled trial; mixed reality; training; training module; module; correlation; gaming; gaming module; serious game; game; games

Introduction

Serious games are now widely used for training as an additional modality to simulation-based education by various industries [1-5]. Training using serious gaming modules can be conducted through various platforms, such as personal computers; tablet personal computers; virtual reality (VR); or mixed reality systems, such as augmented reality or augmented virtuality. Due to the immersive effect and enhanced student engagement, VR and mixed reality modules are currently favored for health care training, such as advanced cardiac life support (ACLS) training [6-8]. The COVID-19 pandemic has also accelerated the process of digital learning [9]. The shift from conventional in-person classrooms to online education has been expedited by the global pandemic, underscoring the importance of maintaining physical distancing during these training sessions [10-12].

The portability and increased affordability of VR and mixed reality systems in recent years have contributed to a significant growth in VR-based learning. A notable advantage of a VR-based serious gaming module is its compatibility with wireless VR headsets, eliminating the need for a personal computer and an expensive cable-based VR headset [13,14].

Due to advancements in the internet connection speed, the organization of VR-based multiplayer training sessions has become feasible. This facilitates learners' collaborative interaction within a group setting to solve medical cases [10,15,16].

Online education stands as one of the potential application areas of the metaverse. The metaverse, which is an immersive, 3D, computer-based, and multiuser online environment, serves as an optimal platform for orchestrating multiplayer training sessions [17,18]. The metaverse merges the physical and computer-based environments by integrating the internet, the web, mixed reality, augmented reality, and VR [10].

The metaverse offers shared social connections, thus differentiating it from existing VR-based serious gaming modules, which primarily focus on presenting computer-based content and environments without fostering communal interaction [19]. This facilitates interaction between learners and educators during metaverse-based training sessions, allowing educators to guide the learners and provide feedback [19,20]. The presence of the educators fosters communication and collaboration during training sessions, thereby enhancing the motivation and engagement of the learners [16,21]. While the metaverse has the potential to provide collaborative and individualized learning experiences, the presence of an educator is crucial in guiding and enhancing the learning process [20,22]. Educator-guided learning in the metaverse has the potential to be superior compared with machine-guided learning, as educators can provide personalized feedback, facilitate discussions, and adapt the learning experience to the unique needs and preferences of each student, whereas machine-guided learning can be particularly beneficial for self-learning, where a fixed learning content is available [10,22].

In our university, a VR-based multiplayer training module within the metaverse environment specifically designed for advanced cardiac life support (ACLS) training has been developed [23]. This module serves as an additional training tool to simulation-based training, enabling learners to receive guidance either from machine algorithms or educators within the metaverse. In light of the limitations of machine-guided learning in VR for collaborative learning and the time constraints associated with educator-guided training in a metaverse environment, a comparative analysis of these methods has been deemed necessary [10,22,24]. The hypothesis of this research is to evaluate and contrast the learning outcomes achieved through VR-based, machine-guided training against educator-guided, VR-based training within the metaverse setting.

Methods

Recruitment

The study was advertised to fourth-semester (2023-2024 Spring Semester) students at the Vocational School for Anesthesiology at Acibadem Mehmet Ali Aydinlar University, and volunteers were accepted for participation. In total, 62 participants volunteered for the study, including 36 female participants and 26 male participants between 20 and 22 years of age. Participants were randomized into 2 groups based on their university ID numbers, with students being assigned to one group if their ID number was odd and to the other group if their ID number was even. By using this randomization process, the participants were divided into 2 groups, each consisting of 31 individuals: one group received VR-based training with machine guidance (MG), and the other group received VR-based training with educator guidance (EG) in the metaverse. All participants read and signed an informed consent document outlining the study flow. The study took place at the Center of Advanced Simulation and Education of Acibadem Mehmet Ali Aydinlar University. Exclusion criteria included previously experienced VR-induced motion sickness; previous ACLS training; and other medical conditions, such as episodes of vertigo or being on medication that causes vertigo-like symptoms.

Study Flow

From this point forward, the group trained with VR-based training with MG will be referred to as the MG group and the group trained with VR-based training with EG in metaverse will be referred to as the EG group. A total of 2 educators with more than 10-year experience in medical simulation guided the participants in metaverse during the study.

The members of both groups undertook VR-based basic training for ACLS. Afterward, participants of the MG group were trained with a VR-based advanced training module, which provides training with full MG, whereas the EG group attended VR-based, educator-guided training in the metaverse.

Serious Gaming Module

The VR-based serious game used in this study, named "3DMedsim ACLS VR," was developed in line with the ACLS standards of the European Resuscitation Council (ERC) and American Heart Association (AHA) [25,26]. The development process was reviewed by clinicians to ensure adherence to crisis resource management and the criteria of the latest AHA and ERC guidelines [27-29].

The serious game is combined with a learning management system along with a learning record store (LRS), allowing the storage of credentials of the users within a shared database [30-32]. In addition, it incorporates a 3D visualization engine integrated with the LRS, enabling the tracking of user interactions and the generation of experience application programming interface (xAPI) calls for each action taken [31]. When users engage in predefined interactions with significant objects within the computer-based environment, corresponding xAPI events are automatically generated [30,33]. This functionality is made possible through the development of a Unity extension software library making necessary xAPI web service calls to the LRS servers through the HTTP protocol [34]. User actions are defined in terms of actor, verb, and object parameters, with the library automatically creating xAPI calls for each action, subsequently recorded by the LRS. The library is also capable of automatically creating xAPI calls for actions not performed within specified time limits or in the correct sequence. Security measures and user authentication features essential for record-keeping are also included in this library, using the basic access authentication method inherent in the http protocol standard.

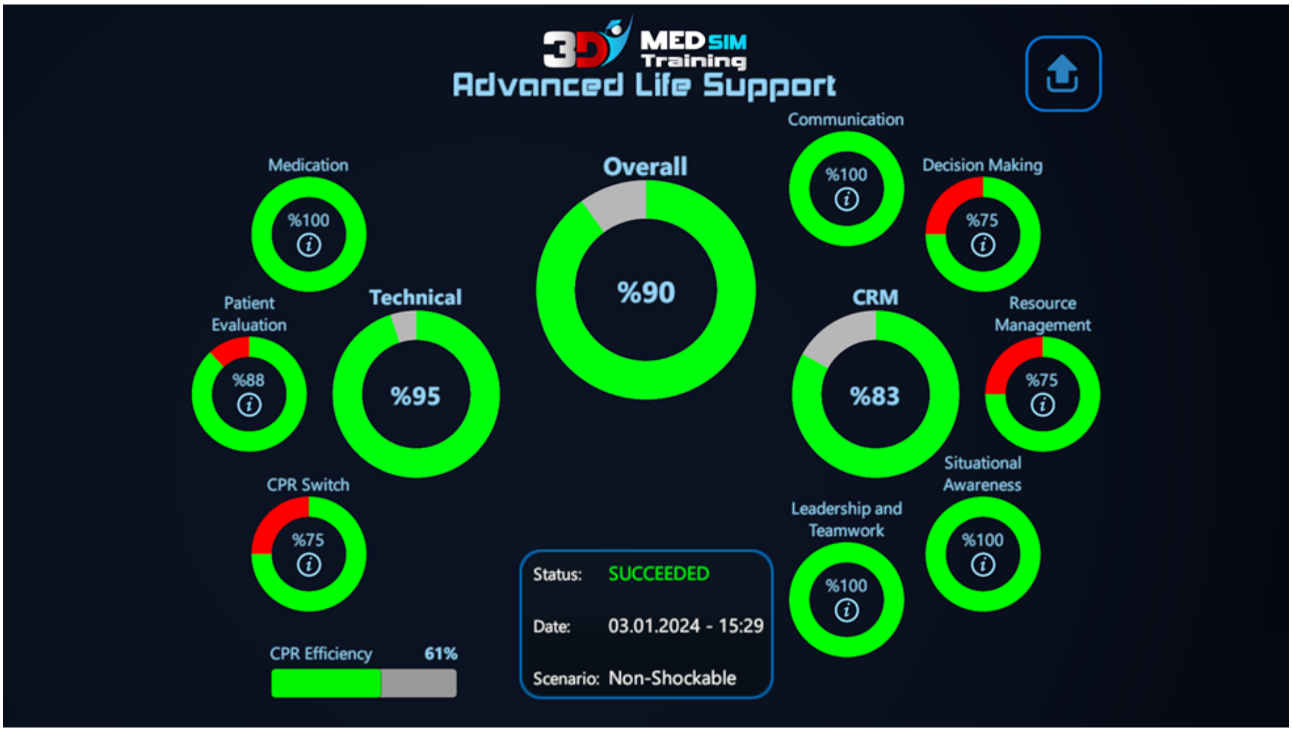
The ACLS serious game includes 4 different stages: basic training, advanced training with machine support, advanced training with educator support in the metaverse environment, and test mode. All of these stages take place in a computer-based hospital room setting to increase immersion, as shown in Figure 1. This serious game also allows users to create a multiplayer lobby and invite others as observers to join the training. These multiplayer lobbies include a voice chat for communication between the observers and the player [13].

In the first stage, basic training is developed with complete beginners in mind. This stage aims to walk the player through the ACLS algorithm step by step with visual and audio cues for each interaction. The user is only able to interact with correct objects in this stage, with no time limits, until they felt ready for advanced-level training mode. The removal of timings and wrong choices removes the stress element and allows the user to familiarize themselves with the ACLS algorithm [13].

In the second stage, advanced training builds on top of basic training. In this stage, visual and audio help for each step is removed. The user is also able to interact with any object or button, which means making mistakes is a possibility at this stage. However, if the user makes a mistake, they are warned by MG. The system monitors each action performed by the player and generates verbal warnings when errors are detected or if time limits are exceeded. In the study, the MG group played this mode independently with MG, while the EG group played the mode in the metaverse environment, accompanied by an educator, as depicted in Figure 2. In educator-guided mode, the system disables machine-guided warnings, enabling educators to provide real-time guidance during training in the metaverse [13].

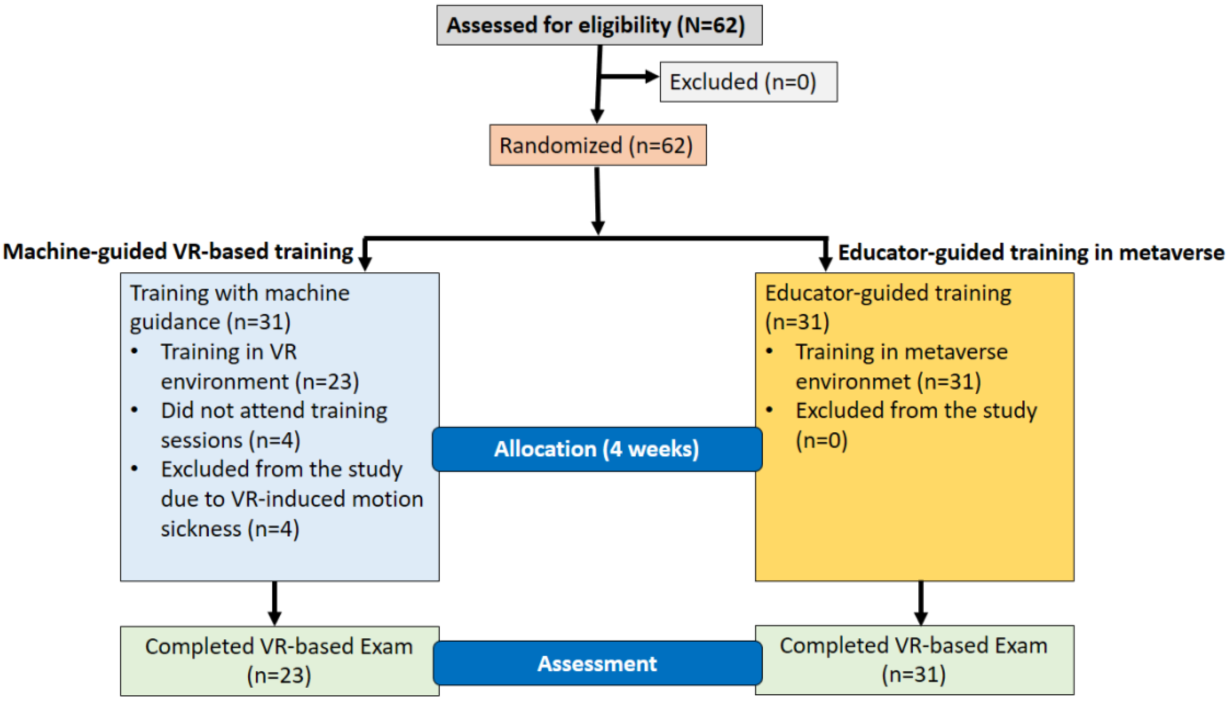
Figure 1. A computer-based hospital room setting in the serious gaming module, including selection options for the next step in the advanced cardiac algorithm. CPR: cardiopulmonary resuscitation.

Figure 2. The score screen of the advanced cardiac life support serious game exam mode, including separate scores for technical skills, nontechnical skills, and the total score. CPR: cardiopulmonary resuscitation.



The final stage is the VR-based exam mode, which was completed by participants immediately following the MG and EG training modes on the same day. In this mode, each action taken by the user is graded according to its timing, order, and accuracy. Contrary to the advanced training mode, there was no guidance for the users. The final score consisted of 2 categories. In total, 70% (70/100) of the total performance score is derived from ACLS assessments, with the remaining percentage (30/100, 30%) calculated from crisis resource management performances of the participants. The primary outcome of the study was determined by the total score obtained through the exam mode of the VR-based training module [13]. A sample test score report can be seen in Figure 3.

Figure 3. Randomized controlled trial flow diagram of the study. VR: virtual reality.



Data Collection

The mean test scores obtained by participants in the VR-based training module exam, along with the duration of machine-guided or educator-guided training (in minutes), were chosen as key indicators for the outcomes if the study. The correlation between exam scores and time spent on training was investigated.

Statistical Analysis

Descriptive statistics were used to define continuous variables, that is, VR exam scores and time spent for machine- or educator-guided training (minutes; mean, SD, median, and range). The correlation between time spent for machine- or educator-guided training and VR exam scores that did not show normal distribution was evaluated using the Spearman rank correlation. The comparison of MG and EG groups that did not show normal distribution in terms of VR exam score and time spent for machine- or educator-guided training was performed using the Mann-Whitney *U* test. The statistical significance level was set at *P*=.05. Analyses were conducted using MedCalc Statistical Software (version 12.7.7) [35].

Ethical Considerations

This study was approved by the Scientific Ethical Committee of Acibadem Mehmet Ali Aydinlar University (approval 2023-20/672). All participants read and signed an informed consent document outlining the study flow. No secondary analysis was performed using these data (Multimedia Appendices 1 and 2).

The data used in the research do not contain any information that can identify individuals. Each participant was assigned a unique code that replaced personal identifiers. The key linking codes to personal information were securely stored separately from the research data. Only essential research team members have access to the deidentified data. Participants did not receive any financial compensation. The participants were informed with the consent form that the results of this research may be used in scientific and professional publications or for educational purposes, but the identities of the participants will remain confidential. The relevant consent form (Turkish or English) can be seen in Multimedia Appendices 1 and 2.

Results

Overview

There were 4 dropouts from the MG group because they did not meet the inclusion criteria. In addition to this, 4 participants from the MG group did not attend the training sessions. Therefore, the EG consisted of 31 participants, whereas the MG group consisted of 23 participants (Figure 3 and Multimedia Appendix 3).

After undertaking a VR-based basic training for ACLS, the participants were divided into 2 groups, MG (n=23) and EG (n=31). The participants of the MG group were trained with a VR-based advanced training module with full MG. The participants of the EG group attended a VR-based, educator-guided training in the metaverse environment. In order to evaluate the knowledge level of the participants, the exam mode of the VR-based training module was used.

As shown in Table 1, the average score was 53 (SD 33) out of 100 for the MG group participants and 68 (SD 35) out 100 for the EG group participants. When comparing the VR test scores, no significant difference was found between the MG and EG groups. The average VR test scores for the EG group were 68 (SD 35), with a median of 86 (range 11-100), whereas MG group scores were 53 (SD 33), with a median of 66 (range 13-100; *P*=.08), as seen in Table 1.

Table 1. Mean and median scores of educator guidance and machine guidance group participants.

Virtual reality exam score (total 100 points)	Machine guidance group	Educator guidance group	*P* value[a]
Mean (SD)	53 (33)	68 (35)	.08
Median (range)	66 (13-100)	86 (11-100)	_[b]

[a]*P*<.005 for all exam scores.

[b]Not applicable.

Mann-Whitney U Test

After comparing the time spent (in minutes) for machine-guided training and educator-guided training for each participant, a significant difference was found between the MG and EG groups. The time spent in the EG group had a median of 21 (range 14-62) minutes. In contrast, the MG group had a median of 15 (range 5-30) minutes (*P*=.002).

Correlation Between VR Test Scores and Duration of the Training

Regarding the correlation between the duration of machine- or educator-guided training and VR test scores, for the MG group, =0.569 and *P*=.005 were obtained. For the EG group, this correlation was found to be =0.298 and *P*=.10. While this correlation is statistically significant for the MG group, it is not significant for the EG group.

Spearman Rank Correlation Test

If the power analysis is based on the exam scores of the 2 groups, the type I error rate was found to be 40% at a 5% significance level. However, when a post hoc power analysis was conducted, considering the correlation between the time spent on training and exam scores, the type I error rate increased to 80% at the same significance level [36,37].

Discussion

Principal Findings

This study focuses on comparing the learning outcomes of machine- versus educator-guided, VR-based training in the metaverse environment. Participants of the EG and MG groups were evaluated using the exam mode of the VR-based training module, which was defined as the main outcome indicator of this study. The findings of the study revealed that there was no statistical difference between the VR-based exam scores of the MG and EG groups. Although these values are statistically significant, the post hoc power analysis based on the exam scores of the 2 groups indicated a power of only 40%. Since this is a preliminary study, a power analysis was not conducted beforehand. While the results are statistically valid, they might be influenced by insufficient power. From an educational perspective, however, the observed difference appears relevant. In addition to this finding, the participants of the EG group spent more time during the advanced training compared with the participants of the MG group (*P*=.002). The correlation between the duration of machine- or educator-guided training and VR-based exam scores is statistically significant for the MG group (=0.569, *P*=.005), but not significant for the EG group (=0.298, *P*=.10). These findings indicate that dedicating additional time to VR-based, machine-guided training holds promise for boosting exam scores of the participants. The post hoc power analysis (80%), considering the correlation between the time spent on training and exam scores, supported this finding.

Using well-designed, VR-based serious gaming modules for specific tasks can enhance learners' motivation, resulting in improved learning outcomes [9,10,16,19,22,38]. The serious gaming module used in this study, specifically designed for adult ACLS training, has proven its capability of improving the knowledge level during our previous study [13].

Numerous studies have demonstrated that VR-based training within the metaverse environment enhances learning outcomes, attributed to its immersive effects and interactive features [15,20,38-41]. Interactivity enables users to engage and communicate through avatars, enhancing embodiment, immersion, and user engagement within the metaverse [9,10,20,39]. The training module used in this study provides different real-life hospital environments to enhance the immersive effect and multiplayer mode to enable users to interact with each other.

A gamified interaction may experience insufficient supervision. The presence of an educator in the metaverse has the potential to enhance learning outcomes, particularly for learners who struggle with self-regulation [9,21]. Introducing multiplayer features within the metaverse can address this limitation, fostering social interaction and potentially enhancing learning outcomes further [19,20]. However, performance data gathered during this study indicated that there was no statistical difference in the exam scores between the MG and EG groups, suggesting that EG did not provide any additional advantage in terms of achieving better learning outcomes.

According to previous studies, an important limitation of metaverse-based learning is the lack of sufficient metaverse literacy among both learners and educators, revealing the importance of familiarization with effective metaverse usage [19,24]. As VR-based serious gaming modules have been embedded into the curricula of training programs of our university for the last 6 years, the students and educators, who have participated in this study, did not encounter any issues with the usage of VR hardware.

Widespread adoption of VR hardware is still restricted among a notable portion of the population. This problem is exacerbated by the lack of high-speed internet connectivity in certain regions of the world. [11,42]. As the same problem also exists in our country, this study was performed on the university campus with high-speed internet access and the VR headsets were provided by the simulation center of our university.

Limitations

The first limitation of the study was the small sample size, which was attributed to the challenges faced in recruiting volunteers. The second limitation of the study was the limited processing power and graphics capability of the headsets used in this study. Therefore, only the heads of the avatars were displayed during the game, leading to a decreased level of immersion. The third limitation of this study was the potential for selection bias. Participants were recruited from a specific population, which may not represent the broader demographic. This can limit the generalizability of our findings. In addition, the voluntary nature of participation may attract individuals with particular characteristics or interests, further skewing the results. Future studies should aim for a more diverse and random sample to mitigate this bias.

Conclusions

Based on the findings of this study, a properly designed VR-based serious gaming module with MG could potentially offer similar learning outcomes to VR-based training in the metaverse with EG. Future research with a larger sample size could explore whether the effect of social interaction with educators in the metaverse environment may provide additional benefits for learners.

Acknowledgments

The authors express their gratitude to the biomedical support team at the Center of Advanced Simulation and Education of Acibadem Mehmet Ali Aydinlar University for their invaluable support during this study. This research did not receive any specific grants from public, commercial, or not-for-profit funding agencies.

Data Availability

All data generated or analyzed during this study are included in this published paper and Multimedia Appendix 4.

Authors' Contributions

DK was responsible for study conceptualization, methodology, and reviewing and editing the manuscript. MEA was responsible for study conceptualization, methodology, and resources and drafting, reviewing, and correspondence. AEO was responsible for the software and methodology. TU was responsible for the recruitment of the participants and data collection.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consent form for the study (Turkish Version).
[PDF File (Adobe PDF File), 547 KB - games_v12i1e58654_app1.pdf]

Multimedia Appendix 2
Consent form for the study (English Version).
[PDF File (Adobe PDF File), 460 KB - games_v12i1e58654_app2.pdf]

Multimedia Appendix 3
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 89 KB - games_v12i1e58654_app3.pdf]

Multimedia Appendix 4
Raw data of the study.
[PDF File (Adobe PDF File), 224 KB - games_v12i1e58654_app4.pdf]

References

1. Barteit S, Lanfermann L, Bärnighausen T, Neuhann F, Beiersmann C. Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. JMIR Serious Games 2021 Jul 08;9(3):e29080 [FREE Full text] [doi: 10.2196/29080] [Medline: 34255668]
2. Hu H, Xiao Y, Li H. The effectiveness of a serious game versus online lectures for improving medical students' coronavirus disease 2019 knowledge. Games Health J 2021 Apr;10(2):139-144. [doi: 10.1089/g4h.2020.0140] [Medline: 33434099]
3. Flavián C, Ibáñez-Sánchez S, Orús C. The impact of virtual, augmented and mixed reality technologies on the customer experience. J Bus Res 2019 Jul;100:547-560. [doi: 10.1016/j.jbusres.2018.10.050]
4. Verkuyl M, Hughes M, Tsui J, Betts L, St-Amant O, Lapum JL. Virtual gaming simulation in nursing education: a focus group study. J Nurs Educ 2017 May 01;56(5):274-280. [doi: 10.3928/01484834-20170421-04] [Medline: 28467555]
5. Choi C, Seok M, Choi S, Kim T, Kim S. Military serious game federation development and execution process based on interoperation between game application and constructive simulators. International Journal of Simulation and Process Modelling 2015;10(2):103 [FREE Full text] [doi: 10.1504/IJSPM.2015.070462]
6. Dewolf P, Clarebout G, Wauters L, Van Kerkhoven J, Verelst S. The effect of teaching nontechnical skills in advanced life support: a systematic review. AEM Educ Train 2021 Jul;5(3):e10522 [FREE Full text] [doi: 10.1002/aet2.10522] [Medline: 34041431]
7. Khanal P, Vankipuram A, Ashby A, Vankipuram M, Gupta A, Drumm-Gurnee D, et al. Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles. J Biomed Inform 2014 Oct;51:49-59 [FREE Full text] [doi: 10.1016/j.jbi.2014.04.005] [Medline: 24732098]
8. Peek JJ, Max SA, Bakhuis W, Huig IC, Rosalia RA, Sadeghi AH, et al. Virtual reality simulator versus conventional advanced life support training for cardiopulmonary resuscitation post-cardiac surgery: a randomized controlled trial. J Cardiovasc Dev Dis 2023;10(2):67 [FREE Full text] [doi: 10.3390/jcdd10020067] [Medline: 36826563]
9. Li M, Yu Z. A systematic review on the metaverse-based blended english learning. Front Psychol 2022;13:1087508 [FREE Full text] [doi: 10.3389/fpsyg.2022.1087508] [Medline: 36687938]
10. Mitra S. Metaverse: a potential virtual-physical ecosystem for innovative blended education and training. Journal of Metaverse 2023;3(1):66 [FREE Full text] [doi: 10.57019/jmv.1168056]
11. Wei X. Online gamification learning during the COVID-19 pandemic. Journal of Education, Humanities and Social Sciences 2023;10:126-132. [doi: 10.54097/ehss.v10i.6902]

12. Montalbano L, Gallo L, Ferrante G, Malizia V, Cilluffo G, Fasola S, et al. Serious games: a new approach to foster information and practices about COVID-19? Front Robot AI 2022;9:830950 [FREE Full text] [doi: 10.3389/frobt.2022.830950] [Medline: 35677083]
13. Aksoy ME, Özkan AE, Kitapcioglu D, Usseli T. Comparing the outcomes of virtual reality-based serious gaming and lecture-based training for advanced life support training: randomized controlled trial. JMIR Serious Games 2023;11:e46964 [FREE Full text] [doi: 10.2196/46964] [Medline: 37768719]
14. Meta Quest 2 Headset. Meta. 2023. URL: https://www.meta.com/quest/products/quest-2/ [accessed 2024-07-10]
15. Kaddoura S, Al Husseiny F. The rising trend of metaverse in education: challenges, opportunities, and ethical considerations. PeerJ Comput Sci 2023;9:e1252 [FREE Full text] [doi: 10.7717/peerj-cs.1252] [Medline: 37346578]
16. Zhang X, Chen Y, Hu L, Wang Y. The metaverse in education: definition, framework, features, potential applications, challenges, and future research topics. Front Psychol 2022;13:1016300 [FREE Full text] [doi: 10.3389/fpsyg.2022.1016300] [Medline: 36304866]
17. Yue K. Breaking down the barrier between teachers and students by using metaverse technology in education. 2022 Presented at: 13th International Conference on E-Education, E-Business, E-Management, and E-Learning; January 14-17, 2022; Tokyo, Japan p. 40-44. [doi: 10.1145/3514262.3514345]
18. Díaz JEM, Saldaña CAD, Ávila CAR. Virtual world as a resource for hybrid education. Int J Emerg Technol Learn 2020;15(15):94-109. [doi: 10.3991/ijet.v15i15.13025]
19. Ng DTK. What is the metaverse? definitions, technologies and the community of inquiry. Australasian Journal of Educational Technology 2022;38(4):190-205. [doi: 10.14742/ajet.7945]
20. Onu P, Pradhan A, Mbohwa C. Potential to use metaverse for future teaching and learning. Educ Inf Technol 2023 Sep 02;29(7):8893-8924. [doi: 10.1007/s10639-023-12167-9]
21. Huh S. Application of computer-based testing in the Korean Medical Licensing Examination, the emergence of the metaverse in medical education, journal metrics and statistics, and appreciation to reviewers and volunteers. J Educ Eval Health Prof 2022;19:2 [FREE Full text] [doi: 10.3352/jeehp.2022.19.2] [Medline: 35021317]
22. Han J, Liu G, Gao Y. Learners in the metaverse: a systematic review on the use of Roblox in learning. Education Sciences 2023 Mar 12;13(3):296. [doi: 10.3390/educsci13030296]
23. Berg KM, Soar J, Andersen LW, Böttiger BW, Cacciola S, Callaway CW, et al. Adult advanced life support: 2020 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. Circulation 2020 Oct 20;142(16_suppl_1):S92-S139 [FREE Full text] [doi: 10.1161/CIR.0000000000000893] [Medline: 33084390]
24. Chang C, Kuo H, Du Z. The role of digital literacy in augmented, virtual, and mixed reality in popular science education: a review study and an educational framework development. Virtual Reality 2023 Jun 20;27(3):2461-2479. [doi: 10.1007/s10055-023-00817-9]
25. Panchal AR, Bartos JA, Cabañas JG, Donnino MW, Drennan IR, Hirsch KG, et al. Part 3: adult basic and advanced life support: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020 Oct 20;142(16_suppl_2):S366-S468 [FREE Full text] [doi: 10.1161/CIR.0000000000000916] [Medline: 33081529]
26. Kane AD, Nolan JP. Changes to the European Resuscitation Council guidelines for adult resuscitation. BJA Educ 2022 Jul;22(7):265-272 [FREE Full text] [doi: 10.1016/j.bjae.2022.02.004] [Medline: 35754855]
27. Krage R, Zwaan L, Tjon Soei Len L, Kolenbrander MW, van Groeningen D, Loer SA, et al. Relationship between non-technical skills and technical performance during cardiopulmonary resuscitation: does stress have an influence? Emerg Med J 2017 Nov;34(11):728-733 [FREE Full text] [doi: 10.1136/emermed-2016-205754] [Medline: 28844039]
28. Saravana-Bawan BB, Fulton C, Riley B, Katulka J, King S, Paton-Gay D, et al. Evaluating best methods for crisis resource management education: didactic teaching or noncontextual active learning. Simul Healthc 2019 Dec;14(6):366-371. [doi: 10.1097/SIH.0000000000000388] [Medline: 31490864]
29. Peltonen V, Peltonen L, Salanterä S, Hoppu S, Elomaa J, Pappila T, et al. An observational study of technical and non-technical skills in advanced life support in the clinical setting. Resuscitation 2020 Aug;153:162-168. [doi: 10.1016/j.resuscitation.2020.06.010] [Medline: 32561474]
30. Clemente FJG, Torre LDL, Dormido S, Salzmann S, Gillet D. Collecting Experience Data from Remotely Hosted Learning Applications. Cham: Springer; 2018:170-181.
31. Learning record store (LRS). Rustici Software. 2023. URL: https://xapi.com/learning-record-store/ [accessed 2023-02-28]
32. Muñoz MLA, Orfali CH. The role of Learning Management Systems (LMS) in educational environments: didactic and technical use. 2018 Presented at: 11th annual International Conference of Education, Research and Innovation; 2018 November 12-14; Seville, Spain p. 10473-10479. [doi: 10.21125/iceri.2018.0976]
33. Akbari F, Nasiri M, Rashidi N, Zonoori S, Amirmohseni L, Eslami J, et al. Comparison of the effects of virtual training by serious game and lecture on operating room novices' knowledge and performance about surgical instruments setup: a multi-center, two-arm study. BMC Med Educ 2022 Apr 11;22(1):268 [FREE Full text] [doi: 10.1186/s12909-022-03351-5] [Medline: 35410279]
34. Unity game engine. URL: https://unity3d.com/ [accessed 2024-07-10]

35. MedCalc Statistical Software. URL: http://www.medcalc.org [accessed 2024-07-10]
36. Rosner B. Fundamentals of Biostatistics. Boston, MA: Cengage Learning; 2016:927.
37. Levine M, Ensom MH. Post hoc power analysis: an idea whose time has passed? Pharmacotherapy 2001 Apr;21(4):405-409. [doi: 10.1592/phco.21.5.405.34503] [Medline: 11310512]
38. Dahan NA, Al-Razgan M, Al-Laith A, Alsoufi MA, Al-Asaly MS, Alfakih T. Metaverse framework: a case study on e-learning environment (ELEM). Electronics 2022 May 19;11(10):1616. [doi: 10.3390/electronics11101616]
39. Agrati LS. Tutoring in the metaverse. study on student-teachers' and tutors' perceptions about NPC tutor. Front. Educ 2023 Oct 13;8. [doi: 10.3389/feduc.2023.1202442]
40. Almarzouqi A, Aburayya A, Salloum SA. Prediction of user's intention to use metaverse system in medical education: a hybrid SEM-ML learning approach. IEEE Access 2022;10:43421-43434. [doi: 10.1109/access.2022.3169285]
41. Lee H, Hwang Y. Technology-enhanced education through VR-making and metaverse-linking to foster teacher readiness and sustainable learning. Sustainability 2022 Apr 16;14(8):4786. [doi: 10.3390/su14084786]
42. Ball C, Huang K, Francis J. Virtual reality adoption during the COVID-19 pandemic: a uses and gratifications perspective. Telemat Inform 2021 Dec;65:101728 [FREE Full text] [doi: 10.1016/j.tele.2021.101728] [Medline: 34887619]

Abbreviations

ACLS: advanced cardiac life support
AHA: American Heart Association
EG: educator guidance
ERC: European Resuscitation Council
LRS: learning record store
MG: machine guidance
VR: virtual reality
xAPI: experience application programming interface

Edited by A Mavragani; submitted 22.03.24; peer-reviewed by O Tanriover, O Cakmakkaya; comments to author 23.04.24; revised version received 23.05.24; accepted 06.07.24; published 07.08.24.

Please cite as:
Kitapcioglu D, Aksoy ME, Ozkan AE, Usseli T
Comparing Learning Outcomes of Machine-Guided Virtual Reality–Based Training With Educator-Guided Training in a Metaverse Environment: Randomized Controlled Trial
JMIR Serious Games 2024;12:e58654
URL: https://games.jmir.org/2024/1/e58654
doi:10.2196/58654
PMID:39110497

Original Paper

Exploring Immersive Multimodal Virtual Reality Training, Affective States, and Ecological Validity in Healthy Firefighters: Quasi-Experimental Study

Joana Oliveira[1,2*], MSc; Joana Aires Dias[1*], MSc; Rita Correia[1], MSc; Raquel Pinheiro[3], MSc; Vítor Reis[3], PhD; Daniela Sousa[1,2], MSc; Daniel Agostinho[1,4], MSc; Marco Simões[1,2], PhD; Miguel Castelo-Branco[1,2], MD, PhD

[1]Coimbra Institute for Biomedical Imaging and Translational Research, Institute for Nuclear Sciences Applied to Health, University of Coimbra, Coimbra, Portugal

[2]Faculty of Medicine, University of Coimbra, Coimbra, Portugal

[3]National Fire Service School, Sintra, Portugal

[4]Center for Informatics and Systems of University of Coimbra, Faculty of Science and Technology, University of Coimbra, Coimbra, Portugal

[*]these authors contributed equally

Corresponding Author:
Miguel Castelo-Branco, MD, PhD
Coimbra Institute for Biomedical Imaging and Translational Research, Institute for Nuclear Sciences Applied to Health, University of Coimbra
Pólo das Ciências da Saúde
Azinhaga de Santa Comba
Coimbra, 3000-548
Portugal
Phone: 351 239488514
Email: mcbranco@fmed.uc.pt

Abstract

Background: Firefighters face stressful life-threatening events requiring fast decision-making. To better prepare for those situations, training is paramount, but errors in real-life training can be harmful. Virtual reality (VR) simulations provide the desired realism while enabling practice in a secure and controlled environment. Firefighters' affective states are also crucial as they are a higher-risk group.

Objective: To assess the impact on affective states of 2 simulated immersive experiences in a sample of healthy firefighters (before, during, and after the simulation), we pursued a multivariate approach comprising cognitive performance, situational awareness, depression, anxiety, stress, number of previous adverse events experienced, posttraumatic stress disorder (PTSD) severity, and emotions. The efficacy and ecological validity of an innovative VR haptic system were also tested, exploring its impact on performance.

Methods: In collaboration with the Portuguese National Fire Service School, we exposed 22 healthy firefighters to 2 immersive scenarios using the FLAIM Trainer VR system (neutral and arousing scenarios) while recording physiological data in a quasi-experimental study. Baseline cognitive performance, depression, anxiety, stress, number of adverse events, and severity of PTSD symptoms were evaluated. Positive and negative affective states were measured before, between, and after each scenario. Situational awareness, sense of presence, ecological validity, engagement, and negative effects resulting from VR immersion were tested.

Results: Baseline positive affect score was high (mean 32.4, SD 7.2) and increased after the VR tasks (partial η^2=0.52; Greenhouse-Geisser $F_{1.82,32.78}$=19.73; P<.001). Contrarily, mean negative affect score remained low (range 11.0-11.9) throughout the study (partial η^2=0.02; Greenhouse-Geisser $F_{2.13,38.4}$=0.39; P=.69). Participants' feedback on the VR sense of presence was also positive, reporting a high sense of physical space (mean score 3.9, SD 0.8), ecological validity (mean score 3.8, SD 0.6), and engagement (mean score 3.8, SD 0.6). Engagement was related to the number of previously experienced adverse events (r=0.49; P=.02) and positive affect (after the last VR task; r=0.55; P=.02). Conversely, participants reported few negative effects (mean score 1.7, SD 0.6). The negative effects correlated positively with negative affect (after the last VR task; r=0.53; P=.03); and avoidance (r=0.73; P<.001), a PTSD symptom, controlling for relevant baseline variables. Performance related to situational awareness was positive (mean 46.4, SD 34.5), although no relation was found to metacognitively perceived situational awareness (r=–0.12; P=.59).

Conclusions: We show that VR is an effective alternative to in-person training as it was considered ecologically valid and engaging while promoting positive emotions, with few negative repercussions. This corroborates the use of VR to test firefighters' performance and situational awareness. Further research is needed to ascertain that firefighters with PTSD symptomatology are not negatively affected by VR. This study favors the use of VR training and provides new insights on its emotional and cognitive impact on the trainee.

(JMIR Serious Games 2024;12:e53683) doi:10.2196/53683

KEYWORDS

virtual reality; firefighter; training; posttraumatic stress disorder; PTSD; emotion; situational awareness; engagement; ecological validity; multivariate approach

Introduction

Background

Over the last few years, gamification and serious games applied to nongame settings have become a visible theme, promoting knowledge acquisition, user engagement, and strategic skills [1]. More precisely, virtual reality (VR) serious game training systems have been showing high potential for many purposes, especially in the professional instruction field [2,3]. Simulated situations under high stress levels have been gaining greater interest for some time. As an example, Keitel et al [4] examined endocrine and psychological stress responses in medical students during 2 simulated stress scenarios (a stress situation induced at a laboratory and a simulated emergency situation) compared to a rest condition.

For firefighters, this assumes special importance in relation to the idiosyncrasies of their strenuous everyday duties [2]. Their functions require a repetitive and continuous preparation to acquire and preserve skills and refresh expertise, which can be potentially facilitated through VR tools. These systems present direct advantages compared to traditional in-person training as a cost-effective and safer realistic alternative to perform dangerous or complex tasks and train decision-making and related abilities in secure environments, replicating similar conditions to those in real life with less time and consumption of resources or other constraints.

Furthermore, VR applications make available a plurality of training scenarios with high ecological validity [2]. Ecological validity can be defined as "the ability to generalize experimental results to different populations, situations, and variables" [5] and is related to the cues contained in an experiment that permit the generalization from the laboratory to real life. VR settings have been successfully used to increase the ecological validity of several experiments [6] as well as enhance participants' daily activities, be it rehabilitation for older adults [7], improving cognitive function of children with attention-deficit/hyperactivity disorder [8], or simply shopping in a virtual supermarket [9].

The high ecological validity of VR scenarios allows for customized realistic immersive scenarios fit to the training purpose [10] without putting the workers' lives or properties in danger [11]. These systems present several other advantages, namely, being an observer [12]; repeated practice; and promotion of the users' engagement and motivation regarding the training tasks and, consequently, the retention of competencies and transference of knowledge [2,13]. VR experiences also encourage post–training session discussions and member collaboration [14]. This is critical because coordination and cooperation are essential between firefighters. Furthermore, VR technology provides user insights into simulated experiences and gains into real-life situations and can be applied to a wide range of personnel categories (eg, trainees and commanders).

As presented by Narciso et al [15], the ideal VR training system should reproduce with the best accuracy possible the real-world situations such that the user can be immersed in the virtual environment and perceive it as close to reality. Thus, the effectiveness of the VR system and its ability to mobilize the user are crucial points to consider regarding firefighting training. Moreover, these workers have to manage critical and dangerous incidents, operate in intense and stressful unpredictable conditions, face hazardous scenarios, deal with human suffering, and handle one's own and colleagues' emotions.

In the last years, some studies have applied simulated tasks [16] and VR systems using several devices and tools (eg, game engines using 3D graphics, software, and VR equipment) for the purpose of firefighter instruction [10,13,17-22]. Reis and Neves [23,24] presented a list of guidelines for simulation in training contexts using VR tools, namely, physical and psychological fidelity, interaction, immersive features, and realism. In addition, Saghafian et al [25] highlighted the importance of realism to increase the acceptance level of the users, effectiveness, and transfer of knowledge, as well as the role of emotions in VR contexts, for example, when using a VR fire extinguisher training tool for several companies from the industrial sector.

From our perspective, developments have been made in this area, with studies focused on VR for firefighting training purposes [15,23,26-28]. In a specific study by Narciso et al [29], virtual environments were used to investigate the effects of multisensory stimuli in 6 different methods—audiovisual or audiovisual combined with smell, or heat, or weight, or uniform, or mask, or a combination of these—on stress (using heart rate and self-report measures), fatigue, cybersickness, presence, and knowledge transfer. This study was conducted in the context of firefighter training and involved a sample of 91 individuals who were not firefighters. Another pilot study revealed the positive effects of resorting to VR tasks for firefighter training purposes, namely, in decision-making abilities during emergency actions [24].

Notwithstanding the existing literature, a clear gap remains in this area, including studies using multimodal experiments capable of integrating realistic immersive scenarios, neuropsychological measures, and biological parameters. The study by Andrews [16] called attention to the ecological and content validity issues of simulated emergency tasks for firefighters, aggregating findings from 60 protocols. The author also highlighted the lack of empirical studies concerning simulating tasks. Indeed, there is a clear need for in-depth ecological experiences, and VR systems might be a valid solution with high application potential.

Some explicit attempts have been made to increase realism in firefighter training [15,16,30]. For example, the study by Butler et al [30] has sought to integrate a multicomponent method including behavioral and biological parameters in a sample of incident commanders to investigate standard operating procedures and decision-making during simulated events. However, the technology used was scarcely immersive (the study used moving images and did not resort to VR technology). In addition, another study [31] investigated decision-making in a sample of Portuguese trainees who became certified for the job using a realistic VR system coupled to a haptic thermal device. The impact of VR was analyzed, but neither physiological measures nor the emotional and neuropsychological components were included.

Other studies have addressed firefighters' situational awareness [12,32-34], defined as the "knowledge or how well the individual discriminates true (signal) from false (noise) information" [32]. However, the literature is still scarce in studying the potential of including this variable applied to firefighter training from a broader perspective, as argued by Chiu [12], namely, combined with VR methodology.

Despite the aforementioned body of work, further research is still needed regarding the effectiveness of VR training applied to firefighter instruction. For example, recently, in their study, Narciso et al [35] applied a heart rate variability measure and self-reported questionnaires to a sample of Portuguese firefighters in this VR context. Although the authors found promising results related to stress, sense of presence, cybersickness, and transference of knowledge, training in real environments showed superior results compared to the VR setting. This is unsurprising and does not preclude the use VR in transfer situations in which learning in real settings is not practical.

To successfully perform their duties, firefighters need flexibility for diverse tasks and operations and have to act immediately, make critical decisions in adverse conditions, deal with fatigue, and control their emotions. After interventions, it is central for their mental health to return to a homeostatically appropriate state to carry on with their lives. In this regard, a systematic review [36] found some mental health issues within firefighter groups, namely, posttraumatic stress disorder (PTSD), anxiety, and depression. Oliveira et al [37] also investigated this issue and replicated these findings.

Emotions play a crucial role in the mental health of firefighters [38]. Findings from Godfrey et al [39] revealed that difficulties in emotion regulation mechanisms were positively associated with the severity of PTSD symptomatology in a sample of firefighters and might be a valuable target considering the vulnerability of this population. Stanley et al [40] highlighted greater levels of distress tolerance as a key trait for firefighting, defined as a capability to deal with negative emotional and physical states. Specifically, firefighting is emotionally challenging work, being cognitively demanding and physically overloading, with implications for the decision-making process [41]. In this regard, Robinson et al [42] found some impairments in cognitive functions among nonfirefighter participants (namely, in the visual declarative memory and working memory domains immediately after the exposure and 20 minutes after) in response to the threat during a simulated firefighting emergency. Furthermore, a study found that shift workers have a higher propensity to show cognitive impairment due to sleep deprivation [43].

This state might affect fundamental cognitive abilities that are crucial to guarantee the success, effectiveness, and safety of firefighting operations. Previous research has indicated that firefighters require higher cognitive function during their hazardous activities [44,45]. Some examples of this higher-order cognition are attention (as a broad construct)—the ability to maintain focus over the time; vigilance—the capacity to maintain continuous attention to specific stimuli, helping in information discrimination; processing speed—the ability to acquire and process information rapidly and, consequently, execute [46]; analytical thinking—the ability to solve problems by virtue of understanding logical principles and assessing the evidence [47]; accurate decision-making ability as a complex process that includes making difficult choices, weighing potential risks and rewards, taking actions, and analyzing the effects [48,49]; working memory—the capacity to maintain information accessible while performing other complex tasks [50], involving temporary storage and manipulation of information; cognitive flexibility—the ability to respond readily, changing selectively according to the environment's demands [51]; and awareness of the situation—the ability to continuously extract information from the surrounding dynamic environment [52], among other executive functions.

Considering the state of art, this study used an Australian VR realistic training system, the FLAIM Trainer, which uses immersive technology for firefighter training and learning. This simulator comprises several cutting-edge characteristics, including various immersive virtual fire scenarios, using a multisensory interface with several customized and haptic components (eg, heat suit, hose line, and a nozzle). For the purpose of this work, a multimodal approach was implemented using realistic training tasks combined with psychological and physiological measures, including evaluation of executive functions and firefighters' performance regarding situational awareness.

Objectives

The main goal of this work was to assess the impact of 2 simulated experiences on positive and negative emotions in a sample of healthy Portuguese firefighters pursuing a multivariate approach and, ultimately, investigate affective aspects of the immersive performance in the following two virtual environments: (1) a neutral scenario (exploration of the surrounding area of a rural residence, including preventive tasks) and (2) an arousing stressful scenario (extinction of a residential fire, including victim rescue actions). The relationship among cognitive performance; reported levels of anxiety, depression, and stress; PTSD severity; adverse events witnessed as a firefighter before the experiment; and affective states before, during, and after the VR task was also analyzed. Another aim was linked to this latter question and concerned evaluation of the impact of an innovative and haptic VR system (the FLAIM Trainer) on firefighter training and daily duties. The ecological effectiveness and the direct and indirect effects of these immersive scenarios were tested, including evaluation of firefighters' situational awareness fit to firefighter demands. These aspects have direct implications for firefighting work and performance in relation to firefighter training and the role of emotion regulation mechanisms in the daily duties of these workers.

Methods

Overview

The sample comprised 22 healthy firefighters from several fire stations in Portugal (n=17, 77% from Lisbon and the Tagus Valley area) recruited in collaboration with the Portuguese National Fire Service School at Sintra (Portugal). The exclusion criteria were the presence of a severe psychiatric or physical condition affecting the brain and behavior—including PTSD, drug or alcohol addiction, significant vision and hearing problems, and pharmacological treatment with effects on brain mechanisms—and never having participated in firefighting campaigns. The Results section presents a detailed description of the sample.

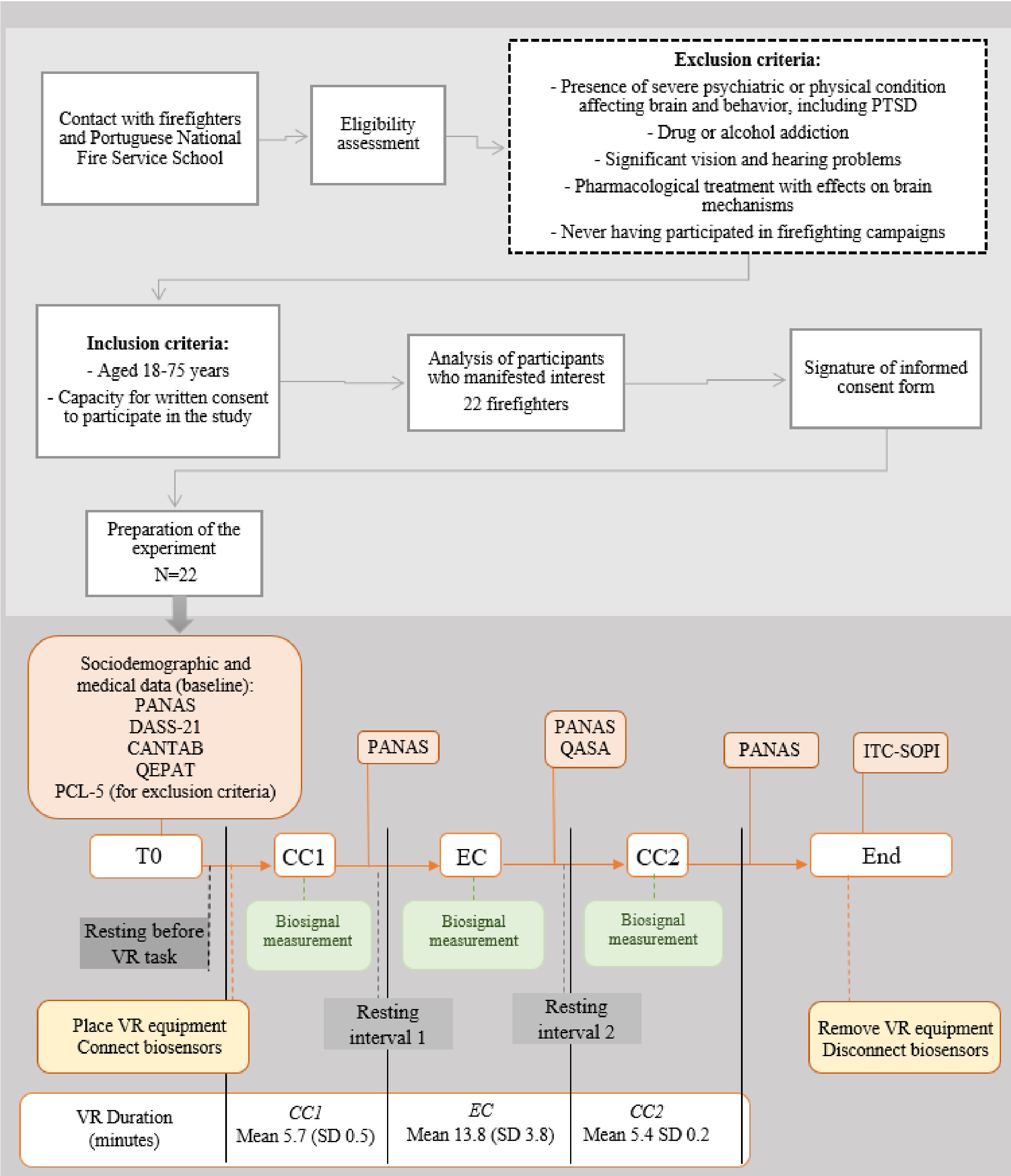
Protocol Procedure

The experiment was conducted at the Portuguese National Fire Service School facilities at Sintra (Portugal), except in the case of 9% (2/22) of the participants, who performed the tasks at the Institute for Nuclear Sciences Applied to Health, University of Coimbra (Portugal), using a quasi-experimental study design. Data collection at the Portuguese National Fire Service School occurred during daytime with participants who were at the school attending a course or called up from nearby fire brigades to participate in the study. Data were also collected during daytime at the Institute for Nuclear Sciences Applied to Health with 9% (2/22) of the firefighters, who purposely went to the institute to participate in the study on a day off from work. Collection at both sites took place during the participants' time off from work. All participants underwent an individual session, including a comprehensive semistructured interview regarding sociodemographic information and self-reported medical history followed by a psychological assessment comprising self-report measures and a battery of computerized tests to assess cognitive function. The total duration of the psychological assessment session was approximately 1 hour. Each participant was then conducted to a separate room to take part in the virtual experience session, including measurement of biosignals and filling out the self-report questionnaires immediately after undergoing each experimental scenario to evaluate their experience in the VR environment at various moments in time (Figure 1).

It was determined that the participants should complete the psychological assessment first and the VR experience afterward due to the consideration that performing the VR task first could elicit a particularly positive or negative mood and, thus, lead to mood-congruent memory recall bias [53,54] and inaccurate reports during the subsequent psychological assessment. To prevent the psychological assessment from having the opposite effect on the VR task performance, it was ensured that the participants had at least 25 minutes to rest between the psychological assessment and the VR task.

During all VR experiments (including during the filling out of the self-report measures), firefighters wore urban-standard fire combat equipment, comprising protective clothing and boots, to accurately simulate a real situation.

Figure 1. Diagram of the recruitment procedure and experimental sessions, including psychological assessment, virtual reality (VR) experiments, and biosignal recording. CANTAB: Cambridge Neuropsychological Test Automated Battery; CC1: control condition 1; CC2: control condition 2; DASS-21: 21-item Depression, Anxiety, and Stress Scales; EC: experimental condition; ITC-SOPI: ITC–Sense of Presence Inventory; PANAS: Positive and Negative Affect Schedule; PCL-5: Posttraumatic Stress Disorder Checklist for the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; PTSD: posttraumatic stress disorder; QASA: Quantitative Analysis of Situation Awareness; QEPAT: Questionário de Exposição e Perturbação dos Acontecimentos Traumáticos (a Portuguese questionnaire related to the exposure to and disturbance of traumatic events); T0: baseline.



Outcome measures

We collected sociodemographic and self-reported medical data, and the measurement instruments used were: (1) the Positive and Negative Affect Schedule (PANAS [55,56]); 2) 21-item Depression, Anxiety, and Stress Scales (DASS-21 [57,58]); 3) Questionário de Exposição e Perturbação dos Acontecimentos Traumáticos (QEPAT; a Portuguese questionnaire related to the exposure to and disturbance of traumatic events [59]); 4) Posttraumatic Stress Disorder Checklist for the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (PCL-5; [60]; [61]); and 5) Cambridge Neuropsychological Test Automated Battery (CANTAB) [62].

The following selected tests were administered in this sequence before the VR tasks: motor screening task (MOT), spatial working memory (SWM) task (extended version), rapid visual information processing (RVP) task (3-target version), intra/extradimensional set shift (IED; lines-first-repeated version), the 6) Quantitative Analysis of Situation Awareness (QASA; instrument created for this study following the methodology proposed by Edgar et al [63]), and the 7) ITC–Sense of Presence Inventory (ITC-SOPI [64,65]).

More detailed information about the outcome measures used in this study and the internal consistency of the scales can be found in Multimedia Appendix 1.

All scales' descriptive statistics, Cronbach α values, and detailed information about the duration of the experiment (conditions and resting intervals) can be found in Multimedia Appendix 2.

VR Session: Setup

The FLAIM Trainer VR technology was designed for firefighter training and learning, presenting a wide range of simulated fire environments (150 available scenarios in 28 languages) [66], which faithfully reproduce dangerous situations usually faced in fire duties (eg, highway, national park, and warehouse fires) or rarer events (eg, fire in a petrol station or engine fire on an aircraft). The system was developed in an attempt to fit the needs regarding fire duties and integrated professional feedback; instructors can monitor in real time and review and track trainees' performance. The simulator was acquired for research purposes within firefighter training.

Some case studies have been reported using the FLAIM Trainer VR system with positive results [67]. Higher levels of acceptance and usability of this VR system were found among 91 Brazilian military firefighters during a firefighting specialization course [68].

The simulator incorporates feedback and outcomes to assist firefighters in decision-making, along with a proprietary virtual fire behavior technology that mimics real fire. This technology aims to improve high-level skills such as risk assessment and dynamic thinking and practical abilities such as muscle memory, radio messaging, and hose and nozzle handling [69]. The hardware was designed considering user-friendliness and immersiveness. The user can explore and navigate the scenes freely, which is facilitated by a teleport mode, and interact with objects.

The hardware setup includes customized components, namely, a self-contained breathing apparatus kit where the VR computer is housed with a built-in HTC Vive Pro VR headset, a half facemask and lung demand valve to capture respiration and reproduce sounds of breathing, and a thermal imaging camera, as well as a fire proximity heat suit technology and a mobile hose line apparatus with haptic technology reproducing a realistic jet and nozzle force; the pump pressure and the suppressant type can be controlled and customized at any time by the instructor. Support equipment includes an instructor tablet, HTC Vive Tracker Puck, charging systems, cables, peripherical hardware, tripods, tracking hardware with base stations (Vive Lighthouses), and Vive Controller [69,70]. In this experiment, the half facemask, thermal imaging camera, and heat suit were not used. Figure 2 shows 2 firefighters performing the VR tasks.

During the VR sessions, biosignals were recorded simultaneously through noninvasive superficial sensors to study autonomic arousal related to firefighters' emotions and behavior and its impact on decision-making under stressful conditions. The parameters measured were electrodermal activity, photoplethysmography, electrocardiography, electroencephalography, and respiration. These data will be reported in a separate study.

The VR experience was divided into 3 moments with 2 different experimental scenarios. The experiment started with a brief tutorial for the user's familiarization with the interaction with the VR system and the use of the hose and the teleport mode. Then, the neutral scenario (*Property Emergency Prepare* scenario; Figure 3) was presented—control condition 1. The firefighters were instructed to explore the scenario for 5 minutes and clean the terrain and the nature around a rural property as a prevention and safeguard measure against potential bushfire risks. In this environment, it was not possible to enter the house, but there were several elements outside with which the participant should interact (eg, cleaning inflammable debris from the nearby zone, ensuring the safety of the animals, and preparing the area for a potential fire).

The experimental condition consisted of a residential fire in the first floor of a house inserted in an urban zone (Bedroom Fire scenario; Figure 4). The scenario ended when the firefighter extinguished the fire and rescued all human victims. The experiment ended with another immersion in the neutral scenario (Property Emergency Prepare scenario; Figure 3)—control condition 2—for 5 minutes, effectively acting as a postbaseline assessment. The firefighters were instructed to act as in a real situation during all conditions. A total of 14% (3/22) of the participants did not complete the last scenario (control condition 2) for health-related reasons (ie, shortness of breath caused by the physical demands of the experimental condition), due to technical issues (ie, system crashed due to lack of battery during the task), or because they did not wish to continue participating in the experiment.

Participants took a mean of 34.1 (SD 8.6) minutes to complete both the control and experimental conditions, also accounting for resting time between the conditions. While a 5-minute time limit was established for the control conditions, some participants needed more time to familiarize themselves with the VR headset and nozzle, so additional time was provided. This also had the aim of ensuring that the participants were engaged and motivated to perform the experimental condition. The first control condition had a mean completion time of 5.7 (SD 0.5) minutes, whereas the second control condition had a mean completion time of 5.4 (SD 0.2) minutes. No time limit was imposed for the experimental condition. Instead, it was determined that the participants had finished the experimental condition once they had completed the 2 objectives of putting out the fire and saving the victims or when they themselves decided to terminate the task. Even though participants were encouraged to complete the objectives, they could quit at any time. Only 5% (1/22) of the participants abandoned the experimental condition after not being able to complete it after 20.5 minutes. Furthermore, it was ensured that the participants had resting time between conditions to avoid fatigue, although the specific interval duration between conditions was determined by the participants. The mean resting times were 4.3 (SD 3.3) minutes between control condition 1 and the experimental condition and 6.5 (SD 1.6) minutes between the experimental condition and control condition 2.

Figure 2. Two firefighters perform the virtual reality tasks (images from the Portuguese National Fire Service School).

Figure 3. Scenario of control conditions 1 and 2 (provided by FLAIM).

Figure 4. Scenario of the experimental condition (provided by FLAIM).

Statistical Analysis

Statistical analysis was performed using SPSS (version 28; IBM Corp), and a significance level of 0.05 was established. Detailed information about the statistical analysis performed can be found in Multimedia Appendix 1.

Ethical Considerations

Participants provided written informed consent. This study was approved by the local research ethics committee of the Faculty of Medicine of the University of Coimbra (approval number: CE_Proc. CE-047/2020) and was guided by the Declaration of Helsinki of 1975 and its latest updates. All data were handled under rigorous protective measures in place to safeguard participant information; specifically, an alphanumeric code was assigned to all participants following a deidentification process to preserve participants' confidentiality and privacy. No compensation was provided for participation in the study.

Results

Participants

Most participants were male (15/22, 68%), with a mean age of 39.1 (SD 9.7; median 42.0, range 23-56) years, and had completed the mandatory education in Portugal (secondary education; 17/22, 77%). In addition, 91% (20/22) and 82% (18/22) of the sample reported consuming alcohol and coffee, respectively, and 36% (8/22) smoked regularly. Regarding their history of psychological help, 59% (13/22) of the firefighters reported having sought psychological assistance in the past. Concerning their activity as firefighters, the sample was equally distributed between professional (12/22, 55%) and volunteer (10/22, 45%) firefighters working primarily in shifts (19/22, 86%) for a mean of 36.8 (SD 17.5; median 42.0, range 7-60) hours per week. The participants had worked for a mean of 18.9 (SD 10.6; median 18.0, range 4-40) years as firefighters. Detailed sociodemographic data can be found in Multimedia Appendix 3.

Characterization of Psychological Measures

DASS-21 Results

Participants' scores in the 3 scales of the DASS-21 were low as they all had normal depression scores (0 to 4), 9% (2/22) had mild anxiety scores (4 to 5), and only 5% (1/22) had mild stress scores (8 to 9). The sample mean was 1.1 (SD 1.3; median 1.0, range 0-4) for the depression scale, 1.2 (SD 1.5; median 0.5, range 0-5) for the anxiety scale, and 3.7 (the highest; SD 2.6; median 3.5, range 0-8) for the stress scale.

QEPAT Results

Regarding the number of adverse events in the context of firefighting activities, participants had experienced a mean of 31.9 (SD 4.5; median 33.0, range 23-40) event types out of 42. No additional events were described by any participants.

The most frequently experienced adverse events were participating in firefighting activities in which goods and properties were at risk (mean 2.7, SD 0.8; median 3.0, range 1-4) and aiding or seeing injured and/or fragile older adults (mean 2.7, SD 0.7; median 3.0, range 2-4), with 59% (13/22) of the sample indicating experiencing each event many times or frequently. In addition, 50% (11/22) of the sample indicated having seen or aided gravely injured adults many times or frequently (mean 2.6, SD 1.1; median 2.5, range 1-4).

While analyzing the level of disturbance, we considered only the events the participants had experienced. The events that obtained the highest mean disturbance score were witnessing the death or grave injury of a firefighter colleague while in active duty (mean 3.4, SD 0.7; median 4.0, range 2-4) and having aided a firefighter colleague who died or was gravely injured while in active duty (mean 3.2, SD 0.7; median 3.0, range 2-4). Both of these events were experienced by 41% (9/22) of the participants. One other event type that obtained a high disturbance score and was experienced by all but 1 participant (21/22, 95%) was hearing through radio communications that firefighter colleagues were in danger, injured, or dead (mean 3.3, SD 0.7; median 3.0, range 2-4). Of the events experienced by all participants, having seen or helped injured children was the one with the highest level of disturbance (mean 2.2, SD 0.9; median 2.0, range 1-4).

PCL-5 Results

None of the participants surpassed the PCL-5 full-scale thresholds or fulfilled the *Diagnostic and Statistical Manual of Mental Disorders*, *Fifth Edition*, criteria for possible PTSD diagnosis, yielding a sample mean score of 3.9 (SD 4.5; median 2.0, range 0-12). The subscales had similar low mean scores. In the Intrusion Symptoms and Avoidance subscales, the sample held a mean of 1.0 (SD 1.5; median 0.0, range 0-5) and 1.0 (SD 1.9; median 0.0, range 0-8), respectively, similar to that of the Negative Alterations in Cognitions and Mood (NACM; mean 1.1, SD 1.7; median 0.0, range 0-6) subscale and slightly lower scores on the Alterations in Arousal and Reactivity (AAR) subscale (mean 0.9, SD 1.3; median 0.0, range 0-4).

PANAS Results

Participants had low negative affect scores for all 4 moments of administration of the PANAS, ranging from a mean of 11.0 after both control conditions (SD 1.3 for control condition 1 and SD 1.8 for control condition 2) to 11.9 (SD 3.2) after the experimental condition. All mean scores were <12 although higher than the minimum score of 10 ($P<.05$ for all comparisons). Conversely, the scores on the positive affect scale were moderate to high. At the baseline measure (T0), the sample had a mean of 32.4 (SD 7.2), close to the middle point of 30 (Cohen d=0.33; t_{21}=1.53; P=.14). The highest mean score was obtained for the measure after the second control condition (mean 40.7, SD 6.6), although it was lower than the maximum score of 50 on the scale (Cohen d=–1.42; t_{18}=–6.18; $P<.001$). Detailed results can be found in Multimedia Appendix 4.

CANTAB Results

All participants excelled in the MOT, concluding the assessment trials with 0 mistakes.

In the SWM task (more detailed information about the outcome measures used in this study can be found in Multimedia Appendix 1), the mean was –0.6, measured in normalized z scores (SD 1.3; median –1.0, range −1.9 to 2.3), which is only slightly lower than the norm mean of z=0 (Cohen d=–0.47; t_{21}=−2.22; P=.04). Only 18% (4/22) of the participants had values higher than the mean, with 14% (3/22) obtaining the highest possible score of z=2.3 (percentile 99). These 3 participants never made the mistake of revisiting a box in which a token had previously been found. The sample's median percentile was 15.5 (range 3-99), with only 23% (5/22) of the sample in percentile 42 or higher. In total, 36% (8/22) of the participants were in the lower range (percentile 10).

In the RVP task, both the z scores of how good the participants were at detecting target sequences (RVP A', prime; RVP prime, the accuracy to detect correctly the targets [RVPA}) and the probability of false alarm (RVP probability of false alarm; RVPPFA) were relatively low (RVPA: mean –0.5, SD 0.6 and median –0.5, range –1.4 to 1.2; RVPPFA: mean 0.3, SD 0.9 and median 0.1, range –1.5 to 2.3). While the frequency of false alarms was similar to that of the norm (Cohen d=0.30; t_{21}=1.41; P=.17), our participants were slightly worse at detecting the target sequences (Cohen d=–0.81; t_{21}=–3.80; P=.001). Another 14% (3/22) of the participants had a probability of false alarm of 0 (percentile 99). In both measures, only 5% (1/22) of the participants were in a percentile of <10 (8 for the RVPA and 7 for the RVPPFA) who were also in percentile 5 in the SWM task. The sample's median percentile for the RVPA measure was 29.5 (range 8-89), indicating that half (11/22, 50%) of the sample had equal or worse capability to detect target sequences than the bottom 29.5% in the CANTAB normative data.

Finally, the sample's mean efficiency in performing the IED task (more detailed information about the outcome measures used in this study can be found in Multimedia Appendix 1) was not different from the norm mean z scores (mean 0.0, SD 0.7; median 0.1, range –1.3 to 0.8; Cohen d=–0.06; t_{21}=–0.28; P=.78). However, the participants' performance was heterogeneous, with a median percentile of 49.6 (IQR 28.8-68.5). In total, 9% (2/22) of the participants were in percentile 9, which indicates that only 9% in the normative data made more errors in this task than these participants.

No correlation was found between the CANTAB z scores and any relevant sociodemographic data (ie, sex, age, number of children, failing in school, alcohol and caffeine consumption, years in service, and hours worked per week as a firefighter; sex: CANTAB_SWM: P=.81; CANTAB_RVPA: P=.499; CANTAB_RVPPFA: P=.31; CANTAB_IED: P=.50; age: CANTAB_SWM: P=.46; CANTAB_RVPA: P=.29; CANTAB_RVPPFA: P=.49; CANTAB_IED: P=.75; number of children: CANTAB_SWM: P=.57; CANTAB_RVPA: P=.55; CANTAB_RVPPFA: P=.73; CANTAB_IED: P=.60; failing in school: CANTAB_SWM: P=.34; CANTAB_RVPA: P=.19; CANTAB_RVPPFA: P=.84; CANTAB_IED: P=.12; alcohol: CANTAB_SWM: P=.37; CANTAB_RVPA: P=.61; CANTAB_RVPPFA: P=.37; CANTAB_IED: P=.17; caffeine consumption: CANTAB_SWM: P=.31; CANTAB_RVPA: P=.44; CANTAB_RVPPFA: P=.66; CANTAB_IED: P=.43; years in service: CANTAB_SWM: P=.496; CANTAB_RVPA: P=.63; CANTAB_RVPPFA: P=.84; CANTAB_IED: P=.998; hours per week: CANTAB_SWM: P=.42; CANTAB_RVPA: P=.82; CANTAB_RVPPFA: P=.49; CANTAB_IED: P=.20).

ITC-SOPI Results

The results for the ITC-SOPI scales were as expected. The highest mean was obtained for the Sense of Physical Space scale (mean 3.9, SD 0.8; median 3.8, range 2.9-6.4) followed by the Ecological Validity scale (mean 3.8, SD 0.6; median 3.8, range 2.8-4.8) and the Engagement scale (mean 3.8, SD 0.6; median 3.8, range 2.2-4.7). These scores were higher than the middle point of the scale (3) for all subscales ($P<.001$ for all comparisons).

Finally, the Negative Effects scale had the lowest scores, with a mean of 1.7 (SD 0.6; median 1.7, range 1.0-3.0). This mean, although small, was higher than the lowest possible score of 1 on this scale (Cohen *d*=1.25; t_{20}=5.71; *P*<.001). The most frequently mentioned negative effects were "feeling tired" (mean 2.1, SD 1.1; median 2.0, range 1-5) and "feeling disoriented" (mean 2.1, SD 1.0; median 2.0, range 1-4), followed by "ocular fatigue" (mean 2.0, SD 1.0; median 2.0, range 1-4).

Correlations Between Measures

At baseline (T0), stress as measured using the DASS-21 correlated positively with the NACM (*r*=0.44; *P*=.04) and AAR (*r*=0.43; *P*=.04) subscales of the PCL-5 but not with the full-scale scores (*r*=0.21; *P*=.34). In addition, the stress scale correlated positively with the negative affect scale of the PANAS at baseline (T0; *r*=0.42; *P*=.049) and after the second control condition (*r*=0.46; *P*=.05). On the other hand, the depression scale of the DASS-21 correlated negatively with the positive affect scale of the PANAS at baseline (T0; *r*=–0.51; *P*=.01) and after each condition (control condition 1: *r*=–0.63 and *P*=.002; experimental condition: *r*=–0.54 and *P*=.009; control condition 2: *r*=–0.70 and *P*<.001). The anxiety dimension of the DASS-21 did not show any correlation with the PANAS for any of the moments measured.

Posttraumatic stress symptoms, as measured using the PCL-5, did not correlate with affect measures (PANAS) except after the arousing experimental condition. Participants who held higher full-scale (*r*=0.43; *P*=.047), intrusion symptom (*r*=0.53; *P*=.01), and avoidance (*r*=0.71; *P*<.001) scores (measured using the PCL-5 at baseline [T0]) had increased negative affect scores measured using the PANAS after the experimental condition. A partial correlation controlling for anxiety, depression, and stress scores (DASS-21) confirmed this result (*r*=0.49 and *P*=.04, *r*=0.51 and *P*=.03, and *r*=0.75 and *P*<.001, respectively). Conversely, even though it appears that there is a negative relation between NACM (as measured using the PCL-5) and positive affect (as measured using the PANAS) after the experimental condition (*r*=–0.48; *P*=.02), this relationship dwindles after controlling for the 3 DASS-21 scales (*r*=–0.45; *P*=.06). A similar result is found when correlating the Avoidance subscale (PCL-5) and the number of adverse events experienced (as measured using the QEPAT). While there appears to be a correlation (*r*=0.44; *P*=.047), it becomes nonsignificant after controlling for the DASS-21 scales (*r*=0.43; *P*=.08).

There was no significant correlation between any of the CANTAB and QASA measures.

Direct Impact of the VR Training Task

Impact on Affect

Performing the VR task had a positive impact on positive affect compared to the baseline (partial η^2=0.52; Greenhouse-Geisser $F_{1.82, 32.78}$=19.73; *P*<.001). Baseline (T0) positive affect was lower than that measured after control condition 1 (*P*=.004), the experimental condition (*P*<.001), and control condition 2 (*P*<.001). The negative affect scores remained consistent throughout the experiment, with no differences found between or after the VR tasks (partial η^2=0.02; Greenhouse-Geisser $F_{2.13, 38.4}$=0.39; *P*=.69). The results found for positive and negative affect remained true after controlling for PCL-5 (full-scale) and DASS-21 scale (depression, anxiety, and stress) scores.

Impact on Sense of Presence

The mean results on the ITC-SOPI scales were positive, as mentioned previously. These positive results are highlighted when related to the affect felt throughout the VR tasks and after the experiment ended, controlling for positive and negative affect at T0 (baseline).

When considering the positive affect reported after all VR tasks were completed (ie, at the end of the experiment, after the second control condition) and its relationship to the ITC-SOPI scales, a positive correlation emerges solely with the participants' reported engagement (*r*=0.55; *P*=.02). Contrarily, the negative affect reported correlated positively and significantly with the negative effects of the VR task as measured using the ITC-SOPI, both after the first nonarousing control condition (*r*=0.61; *P*=.01) and at the end of the experiment (ie, after control condition 2; *r*=0.53; *P*=.03). This is in accordance with the correlation found between the negative effect scale of the ITC-SOPI and other scales. Specifically, a positive and moderate correlation was found between the negative effect and avoidance scores (*r*=0.73; *P*<.001) controlling for baseline levels of depression, anxiety, and stress (DASS-21) and for the number of adverse events experienced (measure obtained using the QEPAT questionnaire). No significant correlations were found between any ITC-SOPI scale and the positive or negative affect reported after the arousing experimental condition. We recognize how idiosyncratic differences in baseline affect could bias results, and therefore, all the aforementioned results do represent partial correlations.

Another noteworthy correlation found was between the number of adverse events that the participants reported having experienced (QEPAT) and the engagement subscale (ITC-SOPI; *r*=0.49; *P*=.02). It appears that participants who experienced more adverse events while in active duty also considered the VR setting more engaging.

Impact on Situational Awareness

Regarding the QASA tool, the sample had overall positive results, obtaining a mean rate of correct responses (ie, hit and correct rejections) of 68% (SD 14.6%; median 64.3%, range 43%-93%). This mean rate was similar for the hits (mean 68.2%, SD 21.1%; median 71.4%, range 29%-100%) and correct rejections (mean 67.5%, SD 12.6%; median 71.4%, range 43%-86%). In accordance, the mean rate of incorrect responses was relatively low at 32% (SD 14.6%; median 35.7%, range 7%-57%) for the rate of both misses (mean 31.8%, SD 21.1%; median 28.6%, range 0%-71%) and false alarms (mean 32.5%, SD 12.6%; median 28.6%, range 14%-57%).

Comparably, the mean actual situational awareness, as measured using A' was relatively high (mean 46.4, SD 34.5; median 45.0, range –25.0 to 92.9), with only 5% (1/22) of the participants having an A′ value of <0 (A′=–25). This indicates that the overall sample had a good situational awareness regarding the elements present in the VR task simulation and was able to distinguish true from false information (with the exception of 1/22, 5% of the participants).

These results were in tandem with the perceived confidence of the participants in their answers (perceived situational awareness

[PSA]). The sample's mean PSA was 47.2 (SD 25.0; median 47.6, range 0-100). All but one participant (21/22, 95%) had confidence in their ability to detect the correct response as they obtained positive values, and only 5% (1/22) of the participants reported not being very confident (ie, PSA value of 0).

When comparing the difference between A′ and PSA scores, it is possible to see that, although no difference between the means was found (mean difference=–0.8; Cohen *d*=–0.02; t_{21}=–0.08; *P*=.94), there was a great amount of variation between participants (SD 45.0; median –12.6, range –61.9 to 83.3 in this difference between A' and PSA). A total of 59% (13/22) of the participants perceived their situational awareness as being higher than it was (reflected by a negative difference), whereas 41% (9/22) had higher situational awareness than they themselves perceived having (as reflected by the positive difference).

The sample presented a slight information acceptance bias, with a mean of –13.6 (SD 30.3; median 0.0, range –100.0 to –9.1). In total, 9% (2/22) of the participants had B″ values of –100, indicating a great tendency to accept information as true even if it was false. A total of 36% (8/22) of the participants did not have any bias as they had a B″ value of 0.

Concordantly with the aforementioned results, no correlation was found between the actual situational awareness (A′) and the PSA (*r*=–0.12; *P*=.59). The same result was found when controlling for experience (ie, years as a firefighter; *r*=–0.11; *P*=.63). PSA values did not correlate with the bias scores (B″; *r*=0.19; *P*=.41) either. The only correlation found was between A′ and B″ scores (*r*=–0.51; *P*=.02) even after controlling for years in service (*r*=–0.50; *P*=.02).

A summary of the main findings of this study can be found in Multimedia Appendix 5.

Discussion

Principal Findings

This study showed the viability of a VR firefighter training tool as a valid serious game approach to act as an alternative for in-person real-life training, which can endanger a firefighter's life and is subject to limitations in the training situations available. Furthermore, the VR setup used in this study allows for ecologically valid and immersive scenarios, which is an advantage for realistic training that prepares for real-life situations.

Participants' level of acceptability is a crucial aspect to consider for VR use. Many reasons can explain the low tendency to integrate VR settings, namely, skepticism, vulnerability, insecurity, lack of ease of interaction with the equipment [71], or absence of realism [25]. Kari and Kosa [72] presented an interesting model joining hedonic and utilitarian or inconvenience measures to explain use and acceptance of VR games supported by the participants' behavioral intention to use that was influenced by curiosity and enjoyment. Another relevant aspect is that perceived ease of use was positively correlated with perceived usefulness, curiosity, enjoyment, and control and negatively associated with discomfort; in turn, utilitarian health and well-being factors were less significant aspects in terms of use of VR game applications.

In this study, the VR task was well received by the participants, in concordance with a previous study using the same VR apparatus [68]. Positive affect increased compared to the baseline just by performing the VR task. This was independent of the type of scenario. That is, it appears that being immersed in the VR setting itself improves positive affect, whether the scenario was neutral and did not require immediate action (preparing the surroundings of a house for a forest fire) or required urgent action and was arousing (putting out a house fire). This is corroborated by the relationship found between positive affect and engagement, both reported at the end of the overall VR task and both yielding highly positive mean scores. This reflects that engagement with the overall VR task increased as positive affect increased, and vice versa. This is true regardless of idiosyncratic differences in baseline positive and negative affect. For VR training to be successful, firefighters have to be engaged with the task, which could be related to positive rather than negative affect. Indeed, a recent systematic review and meta-analysis demonstrated the positive impact of VR environments on positive and negative affect measured using the PANAS [73].

Another aspect of the success of the VR firefighter training is the positive feedback from participants regarding the simulation itself. At the end of the study, participants reported that the simulations were perceived as real and similar to the real world, which was reflected by the high mean scores regarding the sense of physical space and ecological validity, as well as the reported usefulness for training and learning purposes. This adds to the existing literature that has demonstrated how VR increases the ecological validity of experiments [6] and allows for the transfer of skills from VR to real-world scenarios [7-9].

In this study, the realness (measured using the ecological validity scores) of the VR simulations was independent of both the positive and negative affect felt by the participants, which attests to the validity of the simulation being independent of the participant performing it. Similar results were found with VR applied for higher-education purposes in another study. The study showed that the level of immersion is crucial for students' sense of presence and positive affect (also measured using the PANAS). Our results contribute to the generally accepted premise that VR has a positive impact on users' learning and training outcomes [73]. In the same vein, Shafer et al [74] presented a model explaining how enjoyment is generated in VR applications. The authors argued that participants' perceived interactivity predicts spatial presence and realism in the task, whereas realism independently predicts spatial presence and, in turn, spatial presence predicts enjoyment; cybersickness was included as a covariate influencing enjoyment, and previous experience using VR or similar experiences had a weak impact on cybersickness.

Another point to consider regarding the viability of VR firefighting training is the dampening of negative effects. It could be argued that these might surpass positive effects if the use of VR leads to greater distress or discomfort or puts into question the health of the participants. In this study, we found no validation for this tenet. On the contrary, participants' negative affect remained low at baseline and throughout the VR task, with no impact of the arousing, urgent house fire scenario. These results (demonstrated by the firefighters' positive feedback and low negative emotions) are congruent with the principal dimensions highlighted in a very recent study—usability, usefulness, fun, and joy—as critical aspects to the adoption and recommendation of VR experiences [75].

Nevertheless, a positive correlation was found between the participants' avoidance, a symptom of PTSD, and the negative affect reported after the arousing scenario. Similarly, participants who reported experiencing symptoms of avoidance more frequently were also the ones who indicated having felt more negative effects due to the VR simulation. This is true regardless of the number of adverse events experienced and of depression, anxiety, or stress symptomatology. This finding should be further explored to ensure that even firefighters who demonstrate some PTSD symptomatology can perform the VR task and are not vulnerable. In addition, while one might expect that experiencing a greater number of adverse events would deter firefighters from engaging in a VR-simulated firefighting environment, the opposite was found. The participants who reported feeling greater engagement with the VR tasks were those who had experienced an increased number of adverse events. Some explanations can be posited in this regard. Indeed, having experienced stress levels during firefighting activities and confrontation with adversity could lead to greater motivation to improve this state of affairs and mobilize participants to act and manage situations, suggesting specific coping strategies and resilience mechanisms [75]. This topic gains importance with respect to implications for intervention strategies, namely, the relevance of including approaches to boost mechanisms of emotion regulation in firefighters, for example, stress buffer interventions or improving distress tolerance programs, as proposed by Stanley et al [40], who highlighted the positive role played by distress tolerance in occupation stress and suicide risk among 831 firefighters.

Nevertheless, the participants reported feeling scarce negative effects due to the VR task alone. The most frequently mentioned effect was "feeling tired," which might be attributed to the complex implementation of this study itself. Due to the plurality of physiological and psychological measures taken throughout the VR task, the participants had more equipment on them than they would in other VR training contexts and expended more time than usual between scenarios to complete questionnaires. Moreover, this negative effect can be reduced by not performing several VR scenarios in a single training session, thus diminishing fatigue. Regarding "ocular fatigue," the second most mentioned negative effect, it is something that is expected to subside as technology improves. As expected, negative *affect* measured after most of the VR tasks (control conditions) was related to the negative *effects* reported, both being minute.

The positive and negative emotional experience found in this study was congruent with the results found in a sample of trainees using a VR training system for fire extinction [25].

To further explore the reliability of VR training, it is relevant to assess firefighters' performance in this task. For this purpose, the QASA tool was developed. It has a 2-fold objective: it both allows for the assessment of whether firefighters are aware of and attentive to the elements present in the simulation and validates that the firefighters are really attentive to the VR task as if it were a real-life scenario. Participants were successful in this task as they managed to correctly identify most of the elements that were true (ie, hit) and most of the elements that were false (ie, correct rejections). However, this alone does not allow us to infer their awareness of the simulation situation. To assess their actual situational awareness, the emphasis should be on the A′ measure, which can be interpreted as a measure of "an individual's knowledge of the situation as compared to the 'ground truth" [63]. The participants' situational awareness was high, indicating that the overall sample was attentive to the simulation and could identify which statements addressing the simulation were true and which were false. This supports the fact that the participants were paying attention to the simulation. However, were the participants aware of their own situational awareness? Although the sample's mean PSA was similar to their actual situational awareness, there appears to be no correlation between the two, suggesting that performance self-appraisal is not accurate in this setting. In fact, approximately half (11/22, 50%) of the participants overestimated how good their awareness was, whereas the other half (11/22 50%) underestimated it. This is especially surprising considering how all but one participant (21/22, 95%) were confident that most of their responses were correct, some even indicating that all their responses were correct. This result, although expected, has major implications as awareness of the situation is a necessity when it comes to firefighting. Not being attentive to the details of a situation can be the difference between life and death in a real-life scenario. Finally, most participants (11/22, 50%) showed almost no bias in information acceptance. That is, they were considerate when responding to the questions and did not simply accept all the information provided as being true or false.

Concerning the firefighters' emotional and cognitive characterization, a few depressive, stress, and anxiety symptoms showed some correlation with other measures. For instance, stress levels were correlated with some PTSD symptoms. Participants who had higher values of NACM and AAR also had higher stress values, which was to be expected. Indeed, in the systematic review conducted by Jones [36], a plurality of studies found high comorbidities in firefighters.

Furthermore, participants with higher stress scores additionally reported increased feelings of negative affect (when the study started, before the VR task, and after the VR task was finished). While anxiety scores appear to not have any correlation with any of the other measures, the depression level reported had an inverse correlation with positive affect throughout the experiment. That is, the more the participants felt depressed, the less positive their emotional state was. However, this seems to reflect more on the participants' typical emotional state than on the VR task's success and impact. Nonetheless, our sample showed few indicators of psychopathology (as measured using the DASS-21 and PCL-5), which is in concordance with results from a sample of 312 Portuguese firefighters who reported low scores of depression, anxiety, and stress using the same measure used in this study and high levels of happiness [76]. The mental health of firefighters is an issue of particular relevance [37], which cannot be neglected and is emphasized by the aforementioned findings.

In this regard, serious games applied to interventions have revealed promising results in clinical settings, namely, by promoting treatment of and recovery from serious mental illness [77]. Regarding trauma, a recent study presented a contribution toward showing increasing users' awareness of psychological trauma, their sense of security, and promotion of seeking specialized help, using a serious game combined with psychological principles as theorical frameworks [78]. Another study using a computer game showed potential to reduce intrusive memories of a traumatic event [79]. Positive results using a VR simulator were exhibited by 2 truck drivers with PTSD [80]. This literature suggests a high potential for PTSD treatment using the innovative elements of VR.

Regarding the participants' cognitive performance, none demonstrated having any sensorimotor deficits or lack of comprehension, measured using the MOT as recommended [62]. Regarding their performance on the remaining cognitive tasks, the capability to not be swayed by false alarm sequences in the RVP task (RVPPFA measure) and their ability to shift between the intra- and extradimensions of a stimulus (IED test; ie, the ability to discriminate between visual set formation and maintain or shift attention in accordance with environmental clues) were no different than those of the sample norm. On the other hand, on average, the firefighters performed worse on the SWM task (ie, remembering which boxes were previously searched) and the ability to detect rapidly shown number sequences (RVPA measure) during the RVP task.

These measures were explored as they convey cognitive functions that are at play when putting out fires. SWM, strategic planning capacity via trial and error, attention and rapid visual processing, abstract thinking, and cognitive flexibility, as the ability to shift attention between elements of a stimulus following environmental clues through an organized search, are all competencies needed when firefighting as rapid decision-making and problem-solving are required. Although questions could be raised regarding the ecological validity of transposing this concise cognitive assessment to a real-life situation, it also brings into deliberation the need to assess a firefighter's strengths and vulnerabilities when appointing them to a position in action. Questions of ecology are highlighted by the fact that no correlation was found between any of the CANTAB and QASA measures.

Comparisons With Prior Work

According to previous work, VR serious game systems have high potential for training purposes, giving the possibility to customize them directly in accordance with the users' needs [2]. In this study, 2 realistic scenarios (and widely related to the daily duties of firefighters) were selected to ascertain the impact on induced positive and negative affect, on the one hand through neutral and nonurgent duties (preventive tasks) and, on the other hand, through urgent, stressful, and arousing activities, namely, extinguishing a house fire and rescuing victims.

During these tasks, firefighters had the opportunity to practice and develop their skill set (eg, dynamic thinking, stress management, decision-making under pressure, risk assessment, fire control, hose handling technique, and nozzle control) and learn in and experience a safe immersive environment, which might not be possible to replicate in real life. These are advantages of VR compared to traditional training according to several authors [2,10,11]. Other features of these systems demonstrated during this study were the capacity to maintain motivation and user engagement [13]. Indeed, participants reported adequate levels of physical presence, ecological validity, and engagement at the end of the experience and residual negative effects, in accordance with previous studies with Portuguese firefighters [15,35]. Furthermore, while performing the tasks, several firefighters provided additional feedback, which supported the system's similarity to real-life work, as recommended in the literature [15,23], and reinforces its acceptability and usefulness, in agreement with other studies [72,75]. In addition, emotions, a crucial dimension to consider in VR experiences as argued by Saghafian et al [25], were monitored at baseline, between scenarios, and immediately after, showing consistent effects. Collectively, these results demonstrate the reliability of the experiment.

Regarding cognitive performance, some studies have shown the negative impact of a shift work regimen on cognitive functions, which is frequent among Portuguese firefighters (19/22, 86% of our sample were shift workers). In general, our sample revealed normative performance across most tests, which is in line with the considerations of another study conducted with firefighters [44]. These authors assessed cognitive functions using 3 different tests (paired-associate learning, reaction time, and spatial span) of the CANTAB software in simulated hot conditions and further noted that some of these tasks were probably quite easy for their healthy sample of wildland firefighters.

Finally, another key aspect in the firefighters' performance was their actual situational awareness and how good they perceived it to be, that is, their performance self-appraisal (a metacognitive ability). As Edgar et al [63] note, the correlation between the actual situational awareness and PSA is inconsistent in the literature, sometimes being negligible, as in this study. They further highlight several studies that show that performance is moderated by how good the participants believe their situational awareness to be independently of whether their assessment is correct. They hypothesize that the inconsistency between studies might be explained by a lag between the measurement of actual situational awareness and PSA. This explanation does not fit this study as both measures were collected in the same moment. Nevertheless, efforts should be made not only to improve the actual situational awareness of firefighters but also to align their perception with their actual awareness.

Limitations and Future Work

This experiment was conducted for research purposes. Although precise instructions were provided for participants to act exactly as in real-life firefighting, these findings should be viewed with this caveat. Despite the greater levels of ecology and presence revealed by the experiment, fighting a fire and rescuing victims in simulated situations will never be exactly the same as in a real-life situation. In the same vein, although the cognitive dimensions assessed are valuable skills for performing fire duties, some caution is needed when transferring these outputs to real practice as they were obtained through computerized abstract tasks. In addition, considerations should be taken regarding the experimental setting itself. As previously mentioned, the experiment involved performing a complex set of diverse tasks (neuropsychological session followed by 3 moments of immersive VR scenarios with simultaneous measurement of biosignals) requiring that participants wear nonusual equipment (eg, electrocardiography, electroencephalography cap, and respiration band) during task performance and involving a wide range of resources. These constraints may challenge ecological transposition to real firefighting activity and could cause fatigue among participants. Despite this limitation, these were requirements of the experimental design, and pauses between procedures were implemented. However, to reduce this potential concern, future work should consider reviewing the duration of the experiment, including the duration of the simulation.

The impact of the psychological assessment on the participants' performance on the VR task should also be considered in future work. Although the protocol was designed to avoid bias due to mood-congruent memory recall [54], completing the PCL-5, DASS-21, and CANTAB before the VR task could induce fatigue and negative associations from past experiences and have impacted the participants' VR experience. It is recommended that this interplay between the psychological assessment and VR task is further explored in future research.

One other aspect not accounted for is the level of experience the firefighters had with VR specifically and gaming in general. In our case, participants had a moderate level of experience with immersiveness. Previous research has found that the level of experience and age of the participants influences their immersion in the VR setting [81]. Work on Outcome Measure instruments and Statistical analysis are reported at the Multimedia Appendix 1 [82-91]. Even though this was not the focus of this study, it is an aspect that could lead to a heterogeneous experience for individuals of different experience levels and, thus, should be further investigated in future research.

A final aspect that is worth mentioning is that many firefighters have stated that within a firefighter brigade, each member has a specific role assigned inside the team when deployed in field operations has a specific role assigned inside the team when deployed in field operations. In future work, this issue might be mitigated by including multiplayer tasks to achieve the most realistic situation, similar to a prototypical VR system using a multiuser training application among maritime firefighters [19]. Collaboration between members, communication skills, and collaborative work could also be addressed even more as target abilities of being a firefighter. These are skills easily promoted by opting for VR training settings [14] rather than in-person methodologies.

Conclusions

An innovative multicomponent approach was applied to a sample of Portuguese healthy firefighters considering psychological variables and cognitive indicators including situational awareness ability combined with a haptic realistic VR training system and simultaneously coupled with physiological measures. Considering that firefighters are a population at high risk, negative affective states of depression, anxiety, and stress were considered, as well as the severity of PTSD symptoms and adverse events witnessed as a firefighter. This is relevant because Portugal is a country annually ravaged by wildfires, with uncalculated costs at various levels.

Promising results were found as participants were engaged during all tasks, revealing high levels of motivation and acceptance related to the VR software. The ecological validity of both immersive scenarios was supported (irrespectively of facing neutral or arousing stress-inducing conditions), with a positive impact (positive affect was higher after the experiment) that could buffer the minor negative effects reported when PTSD symptom severity and anxiety, stress, and depression were controlled for. In general, the sample demonstrated adequate cognitive abilities and awareness of significant elements of the situations even though they did not correspond exactly with their perception.

These results attest to VR translatability to the real world as a valid cost-effective alternative to traditional in-person training closer to the specificity of firefighter activity. The use of innovative technologies such as the haptic and immersive VR software used in this study was attested for firefighter training and learning purposes and complemented prior work, combining several methodologies in a multicomponent approach. In addition, serious games and VR immersive tools might be put at the disposal of professional corpora to improve work conditions.

Valid contributions to this field of study emerged. Implications for real-world practice are important in relation to the impact of being a firefighter as well as the consequences of daily duties and identification of variables that have to be considered for improving training programs fit to the demands of firefighting work and the daily challenges faced.

Acknowledgments

This study was supported by grants from the Foundation for Science and Technology (in Portuguese, Fundação para a Ciência e Tecnologia) cofinanced by the European Regional Development Fund, PCIF/SSO/0082/2018, DSAIPA/DS/0041/2020, and FCT/UIDB&P/4950/2020. The authors would like to thank the Portuguese National Fire Service School (Sintra division) team and all firefighters who collaborated on this study, as well as the multimedia and reporting team, authors of the 2 images of firefighters' performance during the experiments, which were included in this work. The authors are grateful to Ana Dionísio (a former Coimbra Institute for Biomedical Imaging and Translational Research collaborator) who collaborated in experiments and supported the team during the initial virtual reality experiences. Also worthy of mention are psychologist Salomé Caldeira, PhD, and psychiatrist João Redondo, MD, from the Centre for Prevention and Treatment of Psychological Trauma, Coimbra University Hospital Centre, who provided expertise in the stress and trauma field. The authors also thank the Portuguese authors who gave permission for using the instruments used in this study. Finally, the authors would like to acknowledge the FLAIM systems team, who provided technical support for the FLAIM Trainer simulator system, in particular to James Mullins and Martijn Boosman. No generative artificial intelligence was used in any portion of manuscript writing.

Data Availability

The datasets generated during and analyzed during this study are not publicly available due the use of personal psychological evaluation data but are available from the corresponding author on reasonable request.

Authors' Contributions

VR and RP were responsible for recruitment by contacting the fire stations and scheduling participants' sessions at the Portuguese National Fire Service School. JO, JAD, RC, MS, and MC-B designed the study with validation from VR and RP. JO designed the neuropsychological protocol with direct collaboration from JAD. JO, JAD, RC, DS, and DA collected the data. JAD organized the database with direct collaboration from JO. JAD performed the statistical analysis. MC-B, JO, and JAD wrote sections of the manuscript. All authors contributed to manuscript revision and read and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed information about the outcome measures used in this study, the internal consistency of the scales, and detailed information about the statistical analysis performed in this study.

[DOCX File , 109 KB - games_v12i1e53683_app1.docx]

Multimedia Appendix 2

Descriptive statistics of scales, Cronbach α values, and detailed information about the duration of the experiment (conditions and resting intervals).

[DOCX File , 18 KB - games_v12i1e53683_app2.docx]

Multimedia Appendix 3

Participants' detailed sociodemographic data.

[DOCX File , 16 KB - games_v12i1e53683_app3.docx]

Multimedia Appendix 4

Positive and Negative Affect Schedule scores throughout the study (baseline, control condition 1, experimental condition, and control condition 2).

[DOCX File , 156 KB - games_v12i1e53683_app4.docx]

Multimedia Appendix 5

Summary of the main findings of this study.

[DOCX File , 13 KB - games_v12i1e53683_app5.docx]

References

1. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Dec;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
2. Engelbrecht H, Lindeman RW, Hoermann S. A SWOT analysis of the field of virtual reality for firefighter training. Front Robot AI 2019 Oct 16;6:101 [FREE Full text] [doi: 10.3389/frobt.2019.00101] [Medline: 33501116]
3. Jeon S, Paik S, Yang U, Shih PC, Han K. The more, the better? Improving VR firefighting training system with realistic firefighter tools as controllers. Sensors (Basel) 2021 Oct 29;21(21):7193 [FREE Full text] [doi: 10.3390/s21217193] [Medline: 34770500]
4. Keitel A, Ringleb M, Schwartges I, Weik U, Picker O, Stockhorst U, et al. Endocrine and psychological stress responses in a simulated emergency situation. Psychoneuroendocrinology 2011 Jan;36(1):98-108. [doi: 10.1016/j.psyneuen.2010.06.011] [Medline: 20650570]

5. Kihlstrom JF. Ecological validity and "ecological validity". Perspect Psychol Sci 2021 Mar;16(2):466-471. [doi: 10.1177/1745691620966791] [Medline: 33593121]
6. Parsons TD. Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. Front Hum Neurosci 2015;9:660 [FREE Full text] [doi: 10.3389/fnhum.2015.00660] [Medline: 26696869]
7. de Amorim JS, Leite RC, Brizola R, Yonamine CY. Virtual reality therapy for rehabilitation of balance in the elderly: a systematic review and META-analysis. Adv Rheumatol 2018 Jul 31;58(1):18. [doi: 10.1186/s42358-018-0013-0] [Medline: 30657081]
8. Corrigan N, Păsărelu C, Voinescu A. Immersive virtual reality for improving cognitive deficits in children with ADHD: a systematic review and meta-analysis. Virtual Real 2023 Feb 18;27(4):1-20 [FREE Full text] [doi: 10.1007/s10055-023-00768-1] [Medline: 36845650]
9. Jacobsen LF, Mossing Krogsgaard-Jensen N, Peschel AO. Shopping in reality or virtuality? A validation study of consumers' price memory in a virtual vs. physical supermarket. Foods 2022 Jul 15;11(14):2111 [FREE Full text] [doi: 10.3390/foods11142111] [Medline: 35885352]
10. Bellemans M, Lamrnens D, De Sloover J, De Vleeschauwer T, Schoofs E, Jordens W, et al. Training firefighters in virtual reality. In: Proceedings of the 2020 International Conference on 3D Immersion. 2020 Presented at: IC3D '20; December 15, 2020; Brussels, Belgium p. 1-6 URL: https://ieeexplore.ieee.org/document/9376336 [doi: 10.1109/ic3d51119.2020.9376336]
11. Cha M, Han S, Lee J, Choi B. A virtual reality based fire training simulator integrated with fire dynamics data. Fire Saf J 2012 May;50:12-24. [doi: 10.1016/j.firesaf.2012.01.004]
12. Chiu CY. The modern way of training situation awareness: virtual reality for firefighters. Dissertation. University of Twente. 2020. URL: https://essay.utwente.nl/80766/1/Chiu_MA_BMS.pdf [accessed 2023-07-04]
13. Corelli F, Battegazzorre E, Strada F, Bottino A, Cimellaro GP. Assessing the usability of different virtual reality systems for firefighter training. In: Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 2020 Presented at: VISIGRAPP '20; February 27-29, 2020; Valletta, Malta p. 27-29 URL: https://www.scitepress.org/PublishedPapers/2020/89624/89624.pdf [doi: 10.5220/0008962401460153]
14. Menin A, Torchelsen R, Nedel L. An analysis of VR technology used in immersive simulations with a serious game perspective. IEEE Comput Grap Appl 2018 Mar;38(2):57-73. [doi: 10.1109/mcg.2018.021951633]
15. Narciso D, Melo M, Raposo JV, Cunha J, Bessa M. Virtual reality in training: an experimental study with firefighters. Multimed Tools Appl 2020 Dec 13;79(9-10):6227-6245. [doi: 10.1007/S11042-019-08323-4]
16. Andrews KL. An evaluation and validation of simulated emergency tasks for firefighting. Dissertation. University of Limerick. 2021. URL: https://researchrepository.ul.ie/articles/thesis/An_evaluation_and_validation_of_simulated_emergency_tasks_for_firefighting/19822654 [accessed 2023-07-04]
17. Bayouth ST, Keren N, Franke WD, Godby KM. Examining firefighter decision-making: how experience influences speed in process and choice. Int Fire Serv J Leadersh Manag 2013;7:60 [FREE Full text]
18. Grabowski A, Jach K. The use of virtual reality in the training of professionals: with the example of firefighters. Comput Animat Virtual 2020 Dec 16;32(2):e1981. [doi: 10.1002/cav.1981]
19. Braun P, Grafelmann M, Gill F, Stolz H, Hinckeldeyn J, Lange A. Virtual reality for immersive multi-user firefighter-training scenarios. Virtual Real Intell Hardw 2022 Oct;4(5):406-417. [doi: 10.1016/j.vrih.2022.08.006]
20. Wheeler SG, Engelbrecht H, Hoermann S. Human factors research in immersive virtual reality firefighter training: a systematic review. Front Virtual Real 2021 Oct 12;2:671664. [doi: 10.3389/frvir.2021.671664]
21. Sharma S, Devreaux P, Sree S, Scribner D, Grynovicki J, Grazaitis P. Artificial intelligence agents for crowd simulation in an immersive environment for emergency response. Electron Imaging Soc Imaging Sci Technol 2019 Jan 13;31(2):176-1-176-8. [doi: 10.2352/issn.2470-1173.2019.2.ervr-176]
22. Cohen-Hatton SR, Honey RC. Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. J Exp Psychol Appl 2015 Dec;21(4):395-406. [doi: 10.1037/xap0000061] [Medline: 26523338]
23. Reis VM, Neves C. Application of virtual reality simulation in firefighter training for the development of decision-making competences. In: Proceedings of the 21st Conference on International Symposium on Computers in Education. 2019 Presented at: SIIE '19; November 21-23, 2019; Tomar, Portugal p. 21-23. [doi: 10.1109/siie48397.2019.8970143]
24. Reis V, Neves C. Simulations in virtual reality: assessment of firefighters' decision-making competence. Rev Iberoam Inform Educ 2020;31:28-39 [FREE Full text]
25. Saghafian M, Laumann K, Akhtar RS, Skogstad MR. The evaluation of virtual reality fire extinguisher training. Front Psychol 2020 Nov 9;11:593466 [FREE Full text] [doi: 10.3389/fpsyg.2020.593466] [Medline: 33240182]
26. Barbosa L, Monteiro P, Pinto M, Coelho H, Melo M, Bessa M. Multisensory virtual environment for firefighter training simulation: study of the impact of haptic feedback on task execution. In: Proceedings of the 2017 24º Encontro Português de Computação Gráfica e Interação. 2017 Presented at: EPCGI '17; October 12-13, 2017; Guimaraes, Portugal p. 1-7 URL: https://ieeexplore.ieee.org/document/8124316 [doi: 10.1109/epcgi.2017.8124316]
27. Pinheiro R, Reis V, Curral L, Chambel M. Firefighters' leadership and well-being in rural fires: study in virtual reality environments. In: Viegas DX, Ribeiro LM, editors. Advances in Forest Fire Research. Lisbon, Portugal: Imprensa da Universidade de Coimbra; 2022:1622-1626. [doi: 10.14195/978-989-26-2298-9_248]

28. Pinto D, Peixoto B, Goncalves G, Melo M, Amorim V, Bessa M. Developing training applications for hydrogen emergency response training. In: Proceedings of the 2019 International Conference on Graphics and Interaction. 2019 Presented at: ICGI '19; November 21-22, 2019; Faro, Portugal p. 21-22 URL: https://ieeexplore.ieee.org/document/8955091 [doi: 10.1109/icgi47575.2019.8955091]
29. Narciso D, Melo M, Rodrigues S, Cunha JP, Vasconcelos-Raposo J, Bessa M. Studying the influence of multisensory stimuli on a firefighting training virtual environment. IEEE Trans Vis Comput Graph 2024 Jul;30(7):4122-4136. [doi: 10.1109/TVCG.2023.3251188] [Medline: 37028005]
30. Butler PC, Bowers A, Smith AP, Cohen-Hatton SR, Honey RC. Decision making within and outside standard operating procedures: paradoxical use of operational discretion in firefighters. Hum Factors 2023 Nov;65(7):1422-1434 [FREE Full text] [doi: 10.1177/00187208211041860] [Medline: 34543138]
31. Monteiro P, Melo M, Valente A, Vasconcelos-Raposo J, Bessa M. Delivering critical stimuli for decision making in VR training: evaluation study of a firefighter training scenario. IEEE Trans Human Mach Syst 2021 Apr;51(2):65-74. [doi: 10.1109/thms.2020.3030746]
32. Catherwood D, Edgar GK, Sallis G, Medley A, Brookes D. Fire alarm or false alarm?! Situation awareness and decision-making "bias" of firefighters in training exercises. Int J Emerg Serv 2012;1(2):158. [doi: 10.1108/20470891211275920]
33. Sallis G, Catherwood D, Edgar GK, Baker S, Brookes D. Situation awareness and habitual or resting bias in high-pressure fire-incident training command decisions. Fire Safety Journal 2022 Mar;128:103539. [doi: 10.1016/j.firesaf.2022.103539]
34. Bhattarai M, Martinez-Ramon M. A deep learning framework for detection of targets in thermal images to improve firefighting. IEEE Access 2020;8:88308-88321. [doi: 10.1109/access.2020.2993767]
35. Narciso D, Melo M, Rodrigues S, Cunha JP, Vasconcelos-Raposo J, Bessa M. Using heart rate variability for comparing the effectiveness of virtual vs real training environments for firefighters. IEEE Trans Vis Comput Graph 2023 Jul;29(7):3238-3250. [doi: 10.1109/TVCG.2022.3156734] [Medline: 35254983]
36. Jones S. Describing the mental health profile of first responders: a systematic review. J Am Psychiatr Nurses Assoc 2017 May;23(3):200-214. [doi: 10.1177/1078390317695266] [Medline: 28445653]
37. Oliveira J, Aires Dias J, Duarte IC, Caldeira S, Marques AR, Rodrigues V, et al. Mental health and post-traumatic stress disorder in firefighters: an integrated analysis from an action research study. Front Psychol 2023 Oct 27;14:1259388 [FREE Full text] [doi: 10.3389/fpsyg.2023.1259388] [Medline: 37965662]
38. Vara N, Queirós C, Gonçalves SP. Bombeiros: o papel das emoções e do coping na satisfação com a profissão. Territorium 2015 Aug 23(22):267-276 [FREE Full text] [doi: 10.14195/1647-7723_22_20]
39. Godfrey DA, Zegel M, Babcock JC, Vujanovic AA. Posttraumatic stress disorder and relationship satisfaction among firefighters: the role of emotion regulation difficulties. J Aggress Maltreat Trauma 2022 Feb 24;31(3):356-369 [FREE Full text] [doi: 10.1080/10926771.2022.2043973] [Medline: 35602924]
40. Stanley IH, Boffa JW, Smith LJ, Tran JK, Schmidt NB, Joiner TE, et al. Occupational stress and suicidality among firefighters: examining the buffering role of distress tolerance. Psychiatry Res 2018 Aug;266:90-96 [FREE Full text] [doi: 10.1016/j.psychres.2018.05.058] [Medline: 29857292]
41. Evans TR. Emotions in the fire service: decision-making, risk, and coping. In: Evans TR, Steptoe-Warren G, editors. Applying Occupational Psychology to the Fire Service: Emotion, Risk, and Decision-Making. Coventry, UK: Palgrave Macmillan Chan; 2019:13-57. [doi: 10.1007/978-3-030-14588-0_2]
42. Robinson SJ, Leach J, Owen-Lynch PJ, Sünram-Lea SI. Stress reactivity and cognitive performance in a simulated firefighting emergency. Aviat Space Environ Med 2013 Jun;84(6):592-599. [doi: 10.3357/asem.3391.2013] [Medline: 23745287]
43. Sumińska S, Nowak K, Łukomska B, Cygan HB. Cognitive functions of shift workers: paramedics and firefighters - an electroencephalography study. Int J Occup Saf Ergon 2021 Sep;27(3):686-697. [doi: 10.1080/10803548.2020.1773117] [Medline: 32436781]
44. Williams-Bell FM, Aisbett B, Murphy BA, Larsen B. The effects of simulated wildland firefighting tasks on core temperature and cognitive function under very hot conditions. Front Physiol 2017 Oct 24;8:815 [FREE Full text] [doi: 10.3389/fphys.2017.00815] [Medline: 29114230]
45. Bradley DJ. Cognitive effects on firefighters in Oklahoma from their initial start of service till the present. In: Burrell DN, editor. Applied Research Approaches to Technology, Healthcare, and Business Hershey. New York, NY, USA: IGI Global; 2023:103-120. [doi: 10.4018/979-8-3693-1630-6.ch008]
46. Stout JW, Beidel DC, Brush D, Bowers C. Sleep disturbance and cognitive functioning among firefighters. J Health Psychol 2021 Oct 04;26(12):2248-2259. [doi: 10.1177/1359105320909861] [Medline: 32126834]
47. Swami V, Voracek M, Stieger S, Tran US, Furnham A. Analytic thinking reduces belief in conspiracy theories. Cognition 2014 Dec;133(3):572-585. [doi: 10.1016/j.cognition.2014.08.006] [Medline: 25217762]
48. Dionísio A, Duarte IC, Correia R, Oliveira J, Simões M, Redondo J, et al. A multimodal approach to understand and improve cognitive decision-making during firefighting. In: Viegas DX, Ribeiro LM, editors. Advances in Forest Fire Research. Coimbra, Portugal: Imprensa da Universidade de Coimbra; 2022:1559-1565. [doi: 10.14195/978-989-26-2298-9_237]
49. Ernst M, Paulus MP. Neurobiology of decision making: a selective review from a neurocognitive and clinical perspective. Biol Psychiatry 2005 Oct 15;58(8):597-604. [doi: 10.1016/j.biopsych.2005.06.004] [Medline: 16095567]

50. Baddeley A. Working memory. Curr Biol 2010 Feb 23;20(4):R136-R140 [FREE Full text] [doi: 10.1016/j.cub.2009.12.014] [Medline: 20178752]
51. Scott WA. Cognitive complexity and cognitive flexibility. Sociometry 1962 Dec;25(4):405. [doi: 10.2307/2785779]
52. Cooper S, Porter J, Peach L. Measuring situation awareness in emergency settings: a systematic review of tools and outcomes. Open Access Emerg Med 2013 Dec:1-7. [doi: 10.2147/oaem.s53679]
53. Rholes WS, Riskind JH, Lane JW. Emotional states and memory biases: effects of cognitive priming and mood. J Pers Soc Psychol 1987 Jan;52(1):91-99. [doi: 10.1037//0022-3514.52.1.91] [Medline: 3820082]
54. Faul L, LaBar KS. Mood-congruent memory revisited. Psychol Rev 2023 Nov;130(6):1421-1456 [FREE Full text] [doi: 10.1037/rev0000394] [Medline: 36201828]
55. Galinha IC, Pais-Ribeiro JL. Contribuição para o estudo da versão portuguesa da Positive and Negative Affect Schedule (PANAS): II – estudo psicométrico. Análise Psicológica 2005 Dec 10;23(2):219-227. [doi: 10.14417/ap.84]
56. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
57. Pais-Ribeiro JL, Honrado A, Leal I. Contribuição para o estudo da adaptação portuguesa das Escalas de Ansiedade, Depressão e Stress (EADS) de 21 Itens de Lovibond e Lovibond. Psicol Saúde Doenças 2004;5(1):239 [FREE Full text]
58. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the beck depression and anxiety inventories. Behav Res Ther 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u] [Medline: 7726811]
59. Carvalho C, Ângela M. Perturbação pós-stress traumático e indicadores de (in)adaptação em bombeiros Portugueses. In: Proceedings of the 2009 Congress Conference on Health and Behaviour of Portuguese-speaking Countries. 2009 Presented at: CHBPSC '09; May 6-11, 2009; Braga, Portugal p. 277-290 URL: https://repositorium.sdum.uminho.pt/handle/1822/11328
60. Silva JCDD. Posttraumatic stress disorder checklist for DSM-5 (PCL-5): validação e invariância da medida numa amostra de bombeiros voluntários portugueses. Dissertation. Universidade do Minho. 2018. URL: http://repositorium.sdum.uminho.pt/bitstream/1822/55699/1/dissertação_final.pdf [accessed 2023-07-04]
61. Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP. The PTSD checklist for DSM-5 (PCL-5). National Center for PTSD. 2013. URL: https://www.ptsd.va.gov/ [accessed 2023-07-04]
62. CANTAB® [cognitive assessment software]. Cambridge Cognition. 2023. URL: http://www.cantab.com [accessed 2023-07-04]
63. Edgar GK, Catherwood D, Baker S, Sallis G, Bertels M, Edgar HE, et al. Quantitative Analysis of Situation Awareness (QASA): modelling and measuring situation awareness using signal detection theory. Ergonomics 2018 Jun 29;61(6):762-777. [doi: 10.1080/00140139.2017.1420238] [Medline: 29286253]
64. Vasconcelos-Raposo J, Melo M, Teixeira C, Cabral L, Bessa M. Adaptation and validation of the ITC - sense of presence inventory for the Portuguese language. Int J Hum Comput Stud 2019 May;125:1-6. [doi: 10.1016/j.ijhcs.2018.12.005]
65. Lessiter J, Freeman J, Keogh E, Davidoff J. A cross-media presence questionnaire: the ITC-Sense of Presence Inventory. Presence Teleoperators Virtual Environ 2001 Jun;10(3):282-297. [doi: 10.1162/105474601300343612]
66. FLAIM Trainer™, the world's first immersive technology enabled firefighter training solution. FLAIM Systems. URL: https://flaimsystems.com/products/trainer [accessed 2023-08-28]
67. Levin T, Chessum S, Mullins J, Yoshihashi N, Havashi K. Ensuring greater safety for our firefighters and our communities: integrating FLAIM TrainerTM and hitoeTM. NTT Tech Rev 2019;17(2):31. [doi: 10.53829/ntr201902fa3]
68. De França Bail R, Orlei Michaloski A, Augusto Bortolassi de Oliveira R, Jose Slomp Aguiar E. Usability of immersive technology for education and training of firefighters in Brazil. Int J Innov Educ Res 2022 Sep 01;10(9):365-380. [doi: 10.31686/ijier.vol10.iss9.3921]
69. FLAIM TrainerTM - user manual. FLAIM Systems. 2022. URL: https://ccg-rsi.com/images/training/UserManual_FlaimTrainer_Feb22.pdf [accessed 2024-04-29]
70. VIVE focus vision. HTC. URL: https://www.vive.com/us/ [accessed 2023-10-06]
71. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]
72. Kari T, Kosa M. Acceptance and use of virtual reality games: an extension of HMSAM. Virtual Real 2023 Jan 31:1-21 [FREE Full text] [doi: 10.1007/s10055-023-00749-4] [Medline: 36742344]
73. Bolouki A. The impact of virtual reality natural and built environments on affective responses: a systematic review and meta-analysis. Int J Environ Health Res 2024 Jan;34(1):73-89. [doi: 10.1080/09603123.2022.2130881] [Medline: 36201684]
74. Shafer DM, Carbonara CP, Korpi MF. Factors affecting enjoyment of virtual reality games: a comparison involving consumer-grade virtual reality technology. Games Health J 2019 Feb;8(1):15-23. [doi: 10.1089/g4h.2017.0190] [Medline: 30199273]
75. van Gisbergen MS, Rashnoodi SR, Doicaru M, Campos F. Do you recommend this?: exploring the role of presence, self-efficacy, and usability in the willingness to adopt and recommend a VR application. In: Dieck MC, Jung TK, editors. R and Metaverse. Cham, Switzerland: Springer; 2023:3-15. [doi: 10.1007/978-3-031-50559-1_1]

76. Afonso N, Galvão A, Pinheiro M, Gomes MJ. Felicidade, ansiedade, depressão e stress em bombeiros portugueses. Rev Port Enferm Saúde Mental 2020 Oct(spe7):37-42 [FREE Full text] [doi: 10.19131/rpesm.0245]
77. Fitzgerald M, Ratcliffe G. Serious games, gamification, and serious mental illness: a scoping review. Psychiatr Serv 2020 Feb 01;71(2):170-183. [doi: 10.1176/appi.ps.201800567] [Medline: 31640521]
78. Ammannato G, Chiesi F. Game over, trauma! Empowering trauma healing through gaming. In: Proceedings of the 2023 International Conference on eXtended Reality. 2023 Presented at: XR Salento' 23; September 6-9, 2023; Lecce, Italy p. 454-465 URL: https://link.springer.com/chapter/10.1007/978-3-031-43404-4_31
79. Holmes EA, James EL, Kilford EJ, Deeprose C. Key steps in developing a cognitive vaccine against traumatic flashbacks: visuospatial Tetris versus verbal Pub Quiz. PLoS One 2010 Nov 10;5(11):e13706 [FREE Full text] [doi: 10.1371/journal.pone.0013706] [Medline: 21085661]
80. Menelas BA, Haidon C, Ecrepont A, Girard B. Use of virtual reality technologies as an action-cue exposure therapy for truck drivers suffering from post-traumatic stress disorder. Entertain Comput 2018 Jan;24:1-9 [FREE Full text] [doi: 10.1016/j.entcom.2017.10.001]
81. Lorenz M, Brade J, Klimant P, Heyde C, Hammer N. Age and gender effects on presence, user experience and usability in virtual environments-first insights. PLoS One 2023 Mar 27;18(3):e0283565 [FREE Full text] [doi: 10.1371/journal.pone.0283565] [Medline: 36972245]
82. Lovibond SH, Lovibond PF. Manual for the depression anxiety stress scales. 2nd edition. Sydney, Australia: Psychology Foundation of Australia; 1995.
83. American Psychiatric Association, DSM-5 Task Force. Diagnostic and Statistical Manual of Mental Disorders: DSM-5TM. 5th edition. Washington, DC: American Psychiatric Association; 2013.
84. PTSD: National Center for PTSD. US Department of Venterans Affairs . URL: https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp#:~:text=The%20PCL%2D5%20is%20a%2020%2Ditem%20questionnaire%2C%20corresponding,in%20the%20DSM%2DIV%20version [accessed 2023-07-04]
85. Makhani A, Akbaryan F, Cernak I. Cognitive performance improvement in Canadian Armed Forces personnel during deployment. J Mil Veteran Fam Health 2015 Feb;1(1):59-67. [doi: 10.3138/jmvfh.2014-04]
86. Olff M, Polak AR, Witteveen AB, Denys D. Executive function in posttraumatic stress disorder (PTSD) and the influence of comorbid depression. Neurobiol Learn Mem 2014 Jul;112:114-121. [doi: 10.1016/j.nlm.2014.01.003] [Medline: 24440596]
87. Zhu H, Li Y, Yuan M, Ren Z, Yuan C, Meng Y, et al. Increased functional segregation of brain network associated with symptomatology and sustained attention in chronic post-traumatic stress disorder. J Affect Disord 2019 Mar 15;247:183-191. [doi: 10.1016/j.jad.2019.01.012] [Medline: 30684892]
88. Robbins TW, James M, Owen AM, Sahakian BJ, Lawrence AD, McInnes L, et al. A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging. Cambridge Neuropsychological Test Automated Battery. J Int Neuropsychol Soc 1998 Sep 1;4(5):474-490. [doi: 10.1017/s1355617798455073] [Medline: 9745237]
89. Heaton RK. A Manual for The Wisconsin Card Sorting Test. Odessa, FL: Psychological Assessment Resources; 1981.
90. Stanislaw H, Todorov N. Calculation of signal detection theory measures. Behav Res Methods Instrum Comput 1999 Feb;31(1):137-149. [doi: 10.3758/bf03207704] [Medline: 10495845]
91. Kline RB. Principles and Practice of Structural Equation Modeling. 4th edition. New York, NY: The Guilford Press; 2016.

Abbreviations

AAR: Alterations in Arousal and Reactivity
CANTAB: Cambridge Neuropsychological Test Automated Battery
DASS-21: 21-item Depression, Anxiety, and Stress Scales
IED: intra/extradimensional set shift
ITC-SOPI: ITC–Sense of Presence Inventory
MOT: motor screening task
NACM: Negative Alterations in Cognitions and Mood
PANAS: Positive and Negative Affect Schedule
PCL-5: Posttraumatic Stress Disorder Checklist for the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
PSA: perceived situational awareness
PTSD: posttraumatic stress disorder
QASA: Quantitative Analysis of Situation Awareness
QEPAT: Questionário de Exposição e Perturbação dos Acontecimentos Traumáticos (in Portuguese, a questionnaire related to the exposure to and disturbance of traumatic events)
RVP: rapid visual information processing
RVPA: rapid visual information processing prime, the accuracy to detect correctly the targets
RVPPFA: rapid visual information processing probability of false alarm

SWM: spatial working memory
VR: virtual reality

Edited by T de Azevedo Cardoso; submitted 15.10.23; peer-reviewed by A Thunström, J Buele; comments to author 09.02.24; revised version received 15.06.24; accepted 29.06.24; published 24.10.24.

Please cite as:
Oliveira J, Aires Dias J, Correia R, Pinheiro R, Reis V, Sousa D, Agostinho D, Simões M, Castelo-Branco M
Exploring Immersive Multimodal Virtual Reality Training, Affective States, and Ecological Validity in Healthy Firefighters: Quasi-Experimental Study
JMIR Serious Games 2024;12:e53683
URL: https://games.jmir.org/2024/1/e53683
doi:10.2196/53683
PMID:39446479

Attentional Bias, Pupillometry, and Spontaneous Blink Rate: Eye Characteristic Assessment Within a Translatable Nicotine Cue Virtual Reality Paradigm

Kelly Elizabeth Courtney[1], PhD; Weichen Liu[2], BS; Gianna Andrade[1], BS; Jurgen Schulze[2], PhD; Neal Doran[1,3], PhD

1

2

3

Corresponding Author:
Kelly Elizabeth Courtney, PhD

Abstract

Background: Incentive salience processes are important for the development and maintenance of addiction. Eye characteristics such as gaze fixation time, pupil diameter, and spontaneous eyeblink rate (EBR) are theorized to reflect incentive salience and may serve as useful biomarkers. However, conventional cue exposure paradigms have limitations that may impede accurate assessment of these markers.

Objective: This study sought to evaluate the validity of these eye-tracking metrics as indicators of incentive salience within a virtual reality (VR) environment replicating real-world situations of nicotine and tobacco product (NTP) use.

Methods: NTP users from the community were recruited and grouped by NTP use patterns: nondaily (n=33) and daily (n=75) use. Participants underwent the NTP cue VR paradigm and completed measures of nicotine craving, NTP use history, and VR-related assessments. Eye-gaze fixation time (attentional bias) and pupillometry in response to NTP versus control cues and EBR during the active and neutral VR scenes were recorded and analyzed using ANOVA and analysis of covariance models.

Results: Greater subjective craving, as measured by the Tobacco Craving Questionnaire–Short Form, following active versus neutral scenes was observed ($F_{1,106}$=47.95; P<.001). Greater mean eye-gaze fixation time ($F_{1,106}$=48.34; P<.001) and pupil diameter ($F_{1,102}$=5.99; P=.02) in response to NTP versus control cues were also detected. Evidence of NTP use group effects was observed in fixation time and pupillometry analyses, as well as correlations between these metrics, NTP use history, and nicotine craving. No significant associations were observed with EBR.

Conclusions: This study provides additional evidence for attentional bias, as measured via eye-gaze fixation time, and pupillometry as useful biomarkers of incentive salience, and partially supports theories suggesting that incentive salience diminishes as nicotine dependence severity increases.

(JMIR Serious Games 2024;12:e54220) doi:10.2196/54220

KEYWORDS

nicotine; craving; cue exposure; virtual reality; attentional bias; pupillometry; spontaneous blink rate; eye-tracking; tobacco; VR; development; addiction; eye; pupil; craving; biomarker; biomarkers; tobacco product

Introduction

Automatic appetitive motivational processes are emphasized as critical components in the development and maintenance of substance addiction (eg, dual-process theories [1,2], incentive salience theory [3,4], Tiffany's model [5], and incentive-habit model [6]). Preclinical and human investigations frequently rely on the use of cue exposure paradigms to elicit these motivational processes in the laboratory. The cue exposure paradigm is largely grounded in associative learning principles, which posit that repeated pairing of specific stimuli and substance consumption produces conditioned reinforcement such that the stimuli become conditioned cues capable of eliciting motivational or incentive salience for the substance [7]. Incentive salience can be a conscious or unconscious process and is defined as the motivation for a reward resulting from the integration of one's current physiological state and previously learned associations about the reward cue [8]. Subjective craving for substances is thought to reflect the conscious product of high levels of incentive salience [3,4].

Despite this conceptual coherence, a lack of ecological validity in traditional cue exposure paradigms limits our ability to accurately test and interpret incentive salience outcomes. Attempts have been made to improve the potency of cues and the ecological validity of cue-reactivity designs (eg, [9,10]), yet cue exposure studies typically present the cues in isolation, outside of the context of usual use in natural environments (eg, 2D images or single cigarettes). This isolation of cues limits the ability to invoke a *true* craving state in the lab [11,12] and potentially contributes to poor generalization to the real world [13]. Through greater immersion and interaction within typical contexts of use (eg, the presence of others within a setting where the substance is commonly taken), paradigms using virtual reality (VR) technology have greatly enhanced our ability to elicit craving for various substances in the laboratory [12,14-18], including tobacco [19-22]. Further, VR cue exposure paradigms show great promise as treatment platforms by promoting individualized and accessible care, and allowing for the experience of social immersion and reaction to cues within relevant contexts [23]. Thus, VR cue exposure paradigms represent generalizable tasks with substantial potential for utilization within addiction-related research and clinical settings.

Recent technological advances in VR implementation also allow for precise inline assessment of eye-related measures during cue exposure. The integration of eye-tracking technology into the VR headset is a substantial improvement from previous eye-tracking applications that require inadequate camera placement for precision eye tracking, resulting in partial blockage of the field of view. With this improved technology, it is possible to extract several eye-related measurements that are theoretically related to automatic appetitive motivational processes such as incentive salience and subjective craving; these are attentional bias, pupillary responses, and spontaneous eyeblink rate (EBR).

Attentional bias, or the allocation of a disproportionate amount of time attending to substance-related stimuli relative to neutral stimuli, is thought to either cause or index critical processes responsible for substance-seeking behavior [24]. Several theoretical models suggest that cue-induced subjective craving and attentional bias reflect closely linked underlying processes [3,25,26], such that the degree of attentional bias toward reward cues correlates with the motivational, as opposed to the hedonic, qualities of the reward [27]. Clinically, attentional bias to smoking cues is linked to relapse following smoking cessation [28,29] and was found to be even more predictive of relapse than withdrawal symptoms, subjective craving, and low mood during acute abstinence [29]. Recently, the use of direct eye-tracking indices of attentional bias has shown substantial improvements in bias estimate reliability [30-33]. Assessment within naturalistic settings has also independently improved the reliability [34] and validity [35] of attentional bias measurement, yet the naturalistic constraints of these methods prohibit advanced clinical application. Thus, eye-tracking indices of attentional bias within naturalistic, yet clinically feasible settings, may be especially useful as biomarkers of the incentive salience/craving phenomenon in substance addiction.

Pupillary responses and EBR represent two lesser-studied eye characteristics with theoretical ties to incentive salience processes that warrant further study as potential biomarkers of addiction. Pupil diameter has been associated with engagement of cognitive resources [36], sensitivity to rewards [37], and reward processing broadly [38]. Pupil diameter changes indicate fluctuations in attention allocation and are suggested as a measure of attention-related constructs that do not reach the threshold of overt behavior or conscious appraisal [39]. Only one study has investigated pupillometry as a measure of response to substance cue exposure in humans and found that pupillary bias toward alcohol versus neutral cues, but not subjective craving reports, predicted relapse to alcohol use in a sample of detoxified patients with alcohol dependence [40].

EBR has been linked with striatal dopaminergic function in preclinical models and has been advanced by some as a reliable alternative to the assessment of dopaminergic functioning via positron emission tomography [41]. Dopamine release in the basal ganglia is theorized to inhibit the spinal trigeminal complex, consequently triggering increased EBRs [42]. Given the observed modulation of striatal dopamine during cue exposure [43], it may be possible to detect these dopaminergic fluctuations through EBR measurement. Yet, outside of our preliminary report on this sample [44], this hypothesis has not yet been tested.

This study sought to investigate the validity of these eye characteristics as markers of incentive salience acquired during a novel real-world VR nicotine and tobacco product (NTP) cue exposure paradigm across NTP users with varying degrees of use. An initial report was published by our group early on during data collection (N=31) [44] that described the development of the NTP cue VR paradigm and provided preliminary results supporting the potential of this paradigm as an effective lab-based cue exposure task, including its ability to elicit subjective craving and a sense of presence in the virtual world. The present study provides an update to this preliminary report with a larger sample of daily and nondaily users of NTPs (N=108). It was hypothesized that eye-based markers of attentional bias, pupillometry, and EBR would be greater in response to NTP cues compared with control cues presented during the VR NTP cue exposure paradigm and that these measures would correlate with subjective craving and measures of past NTP use.

Methods

Participant Recruitment and Screening Procedures

As previously described [44], participants were recruited through flyers and social media posts (eg, Facebook, Craigslist, and San Diego Reader) targeting the San Diego community. A brief telephone screening interview was used to determine initial eligibility. Inclusion criteria for the study were ages ≥18 years, at least weekly NTP use during the past 3 months, and NTP use history ≥1 year. Exclusionary criteria were nonfluency in English, medical or psychiatric history affecting brain development (ie, current severe *Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition; *DSM-5*] psychiatric disorders other than tobacco use disorders, severe head trauma with loss of consciousness >2 minutes, or history or treatment of neurologic disorders), and (3) visual problems that interfere with task completion (eg, severe motion sickness and blindness). NTP use was defined as use of any tobacco (eg, cigarette, cigar, or hookah) or electronic nicotine delivery system (eg, e-cigarette or vaporizer). NTP use groups were defined as daily users (average use of 7 days per week in the past 3 months) and nondaily users (average use of 4 - 27 days per month in the past

3 months). The distinction between daily and nondaily users is supported by the literature, confirming that regular, voluntary, nondaily users of tobacco do not smoke often enough to regulate nicotine levels and evince less tobacco dependence and cue-induced craving as compared to daily users [45-47].

Eligible participants were invited into the laboratory and instructed to bring their NTPs with them for use immediately after the visit to control for effects related to expectations of imminent substance availability [48]. Participants were asked to abstain from NTP use for at least 1 hour prior to their visit, resulting in VR testing at least 2 hours post use (the average half-life of nicotine in body tissues [49]), and all other substance use (including alcohol and cannabis use) for at least 24 hours prior to testing. Abstinence was self-reported as COVID-19 restrictions did not allow for biological verification.

Ethical Considerations

Participants received a detailed explanation of study procedures and provided written informed consent consistent with the University of California, San Diego Institutional Review Board policies upon arrival to the laboratory (UCSD IRB #180719). Participant data were deidentified. Participants received US $50 cash for completing the in-person session and up to US $60 in gift cards for completing the follow-up portion of the research (not presented here).

Psychological and Substance Use Measures

Prior to undergoing the NTP cue VR paradigm, participants underwent a clinical interview to assess psychological health (Mini International Neuropsychiatric Interview for *DSM-5* [50]) and completed self-report questionnaires encompassing basic demographic information, previous VR experience, and other measures of psychological functioning not reported here.

The 90-day timeline follow-back (TLFB) [51] and Customary Drinking and Drug Use Record [52] interviews were administered to assess substance use history (including recency since last NTP use in minutes). The TLFB has high test-retest reliability for intervals ranging from 30 to 360 days prior to the interview date, with an intraclass correlation coefficient of 0.92 for "Total number of cigarettes smoked per interval" [53]. The Population Assessment of Tobacco and Health (PATH) tobacco dependence index [54], with a range of 0 - 80, was administered to assess nicotine dependency across nicotine products. Subjective craving before and after the VR paradigm was assessed via the Tobacco Craving Questionnaire–Short Form (TCQ-SF) [55], modified to reference participants' preferred nicotine product (eg, e-cigarettes and tobacco cigarettes). The TCQ-SF has demonstrated reliability (Cronbach α coefficients >0.69 across subscales) and validity, and has been shown to reliably measure the same multidimensional aspects of tobacco craving as the original TCQ when tested following overnight abstinence and during ad libitum smoking [55]. Pre- and post-VR TCQ-SF scores and previous 90-day NTP use episode count from the TLFB (logged transformed due to skewness) were used in the quantitative analyses presented below. REDCap electronic data capture tools hosted at the University of California, San Diego were used for interview and self-report data collection.

Following completion of the NTP cue VR paradigm, participants were assessed on VR presence (Igroup Presence Questionnaire [IPQ] [56]) and VR-related simulator motion sickness (Simulator Sickness Questionnaire [SSQ] [57]). The IPQ total score was calculated using a simple averaging method to obtain a perceived presence score ranging from 0 to 100. The SSQ was scored in concordance with procedures outlined to assess VR-specific sickness (Virtual Reality Sickness Questionnaire) [58], involving a simple averaging method to obtain a score ranging from 0 to 100.

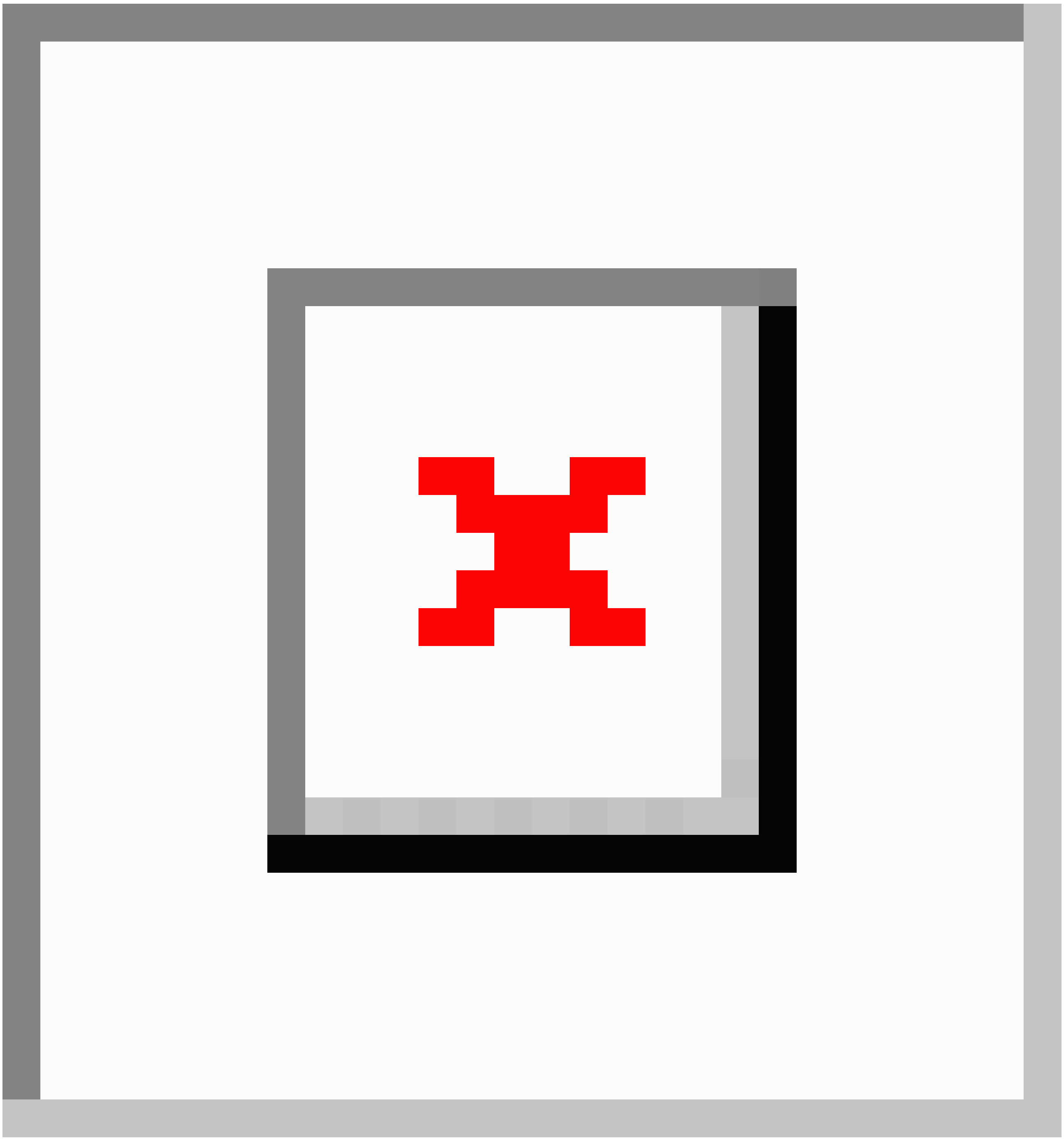
NTP Cue VR Paradigm

As previously detailed [44], the HTC VIVE Pro Eye VR headset (HTC, Taoyuan City, Taiwan) was used to enable VR capabilities and collect eye-related data during the NTP cue VR paradigm (built in Unity). HTC's SRanipal SDK [59] was used in conjunction with Tobii's XR SDK (Tobii Technology, Stockholm, Sweden) to provide access to data from the eye tracker. Specifically, Tobii's XR SDK and Gaze-to-Object-Mapping (G2OM) algorithm were applied to determine object selections, while the remaining data were retrieved from the SRanipal SDK.

Initially, 3 active scenes containing control and NTP-related cues (driving, patio, outdoor BBQ) and 3 neutral scenes containing only control cues (bus, waiting room, library) were developed. However, after preliminary testing of the paradigm, 1 active and 1 neutral scene were removed due to inconsistent eye-gaze effects and increased VR-related sickness (driving, bus; see Liu et al [44] for additional details). All active scenes contained multiple types of NTPs (see Liu et al [44] for a detailed description of the scenes). Thus, the data presented below are derived from the remaining 2 active (patio, outdoor BBQ) and 2 neutral (waiting room, library) scenes (Figure 1). Importantly, the selection of the study outcomes was done prior to any data acquisition and thus was not affected by removal of the scenes.

During the paradigm, participants were encouraged to move around in the virtual scenes via teleportation and interact with cue objects using two handheld VIVE controllers. Virtual visual analog scales assessing subjective craving ("How much are you craving nicotine right now?") and scene relevance ("How relevant was that scene to your own life?") were presented between scenes, and responses were made by adjusting a slide bar using one of the controllers. Participants were instructed to "Just explore everything around you until the scene changes."

Figure 1. Screenshots of the four final scenes from the nicotine and tobacco product cue virtual reality paradigm. Neutral scenes include the (A) library and (B) waiting room. Active scenes include the (C) outdoor BBQ and (D) patio. The figure is adapted from the initial task development report [44], which is published under Creative Commons Attribution 4.0 International License [60].



Gaze Statistics Calculation

A combination of the G2OM algorithm provided by Tobii's XR SDK, a machine learning–based mapping algorithm that aims to improve small object and fast-moving object tracking, and naive ray-casting was used to enable object selection in the direction of the gaze [61]. Specifically, to ensure adequate performance without detrimentally affecting the frame rate, the G2OM algorithm provided by Tobii's XR SDK was used only for the detection of the interactable objects (including all NTPs and control cues), and the naive ray-casting was used for the detection of background and other nonmovable large objects. In addition, when a virtual object was interacted with via the controllers, the object selection was "locked" until the object was released, thus reducing eye-gaze errors due to rapid movement and microsaccades.

Given the complexity of the dynamic virtual environment, eye fixations were defined based on functionality—the duration of eye gaze intersection with the selected object of interest. The total object fixation number and total object fixation time (dwell time) were summed within each cue category (NTP and control) for each scene. Mean fixation time (total fixation time/object fixation number) indices were then created within each cue category for each scene and averaged across the scenes. Mean NTP versus control cue fixation and fixation time contrast scores from the active scene metrics were calculated for use in the exploratory analyses described below.

Pupil Diameter and Blink Detection

Pupil diameter was recorded continuously throughout the paradigm and mapped to each object identified via Tobii's G2OM algorithm. Pupil diameter was summed within each cue category (NTP and control) for each scene. Mean pupil diameter indices were then created by averaging over the mapped pupil diameter samples within each cue category for each scene and averaged across the scenes. Mean NTP versus control cue pupil diameter contrast scores (mean NTP cue diameter – mean control cue diameter) from the active scene metrics were calculated for use in exploratory analyses.

Consistent with previous studies, an eyeblink was defined as complete eyelid closure (or missing pupil diameter) with the pupil covered for 50 - 500 milliseconds [62,63]. Total EBRs were summed within each scene and averaged within scene type (active and neutral). Mean active versus neutral scene EBR contrast scores (mean active scene EBR – mean neutral scene EBR) were calculated for use in exploratory analyses.

Statistical Analysis

Statistical analyses for demographic differences between NTP use groups were conducted using one-way ANOVA models. Analyses for the main outcomes were conducted using repeated measures ANOVAs, followed by analyses of covariance controlling for age and sex. Interactions between NTP use group and cue/scene type as well as their main effects were estimated. Estimated marginal means (EMMs) are reported for the main effects that control for the other variable of interest (ie, NTP use group or cue/scene type) in the model. Analyses of preliminary reliability estimates across scenes were conducted using Pearson correlations. A significance threshold of $P<.05$ was set for all primary analyses.

Exploratory investigations of relationships between the objective outcomes (ie, total fixations, mean gaze fixation time, pupil diameter, and EBR) and subjective craving (pre-VR, during VR, post-VR), recency of NTP use (minutes since last NTP use at time of testing), and previous 90-day NTP use utilized Pearson correlations and partial correlations. Bonferroni-corrected *P* value thresholds that corrected for the tests of the 3 subjective craving and 2 NTP use variables per objective outcome were calculated ($P_{corr}<.01$). Follow-up analyses computed correlations within NTP use groups, transformed the *r* values into *z* scores using Fisher *r*-to-*z* transformation, and compared the *z* values by determining the observed *z* test statistic. SPSS Statistics for Windows, version 28 (IBM Corp) software was used for all analyses.

Results

Study Sample

A total of 303 phone screenings were completed, with 193 individuals deemed eligible. The primary reasons for ineligibility were low/no NTP use (32 screenings) and severe psychiatric comorbidity/psychotropic medication use (39 screenings). Many eligible screenings were not enrolled due to COVID-19 restrictions/cancellations at the time. Of the 115 participants who completed the protocol, 108 participants had usable eye fixation data, 104 had pupillometry data, and 106 had EBR data (excluded participants had calibration or technical issues with the eye-tracking hardware/software).

Demographic information for the sample of 108 with eye fixation data is presented in Table 1. In general, the sample contained slightly more male participants (n=61, 56.5%) and predominately self-identified as White (n=60, 55.6%), and 58.3% (n=63) had no or very limited (one time) previous experience with VR. Of the full sample, 56.5% (n=61) were predominately e-cigarette or nicotine vaporizer users; however, 68.5% (n=74) of the sample reported smoking a tobacco cigarette, and 77.8% (n=84) reported use of any combustible tobacco product (cigarette, cigar, pipe, or hookah) within the previous 6 months. Daily and nondaily NTP use groups were not found to differ in type of NTP use (*P* values >.25).

Table . Sample demographics by nicotine and tobacco product (NTP) use group and total sample.

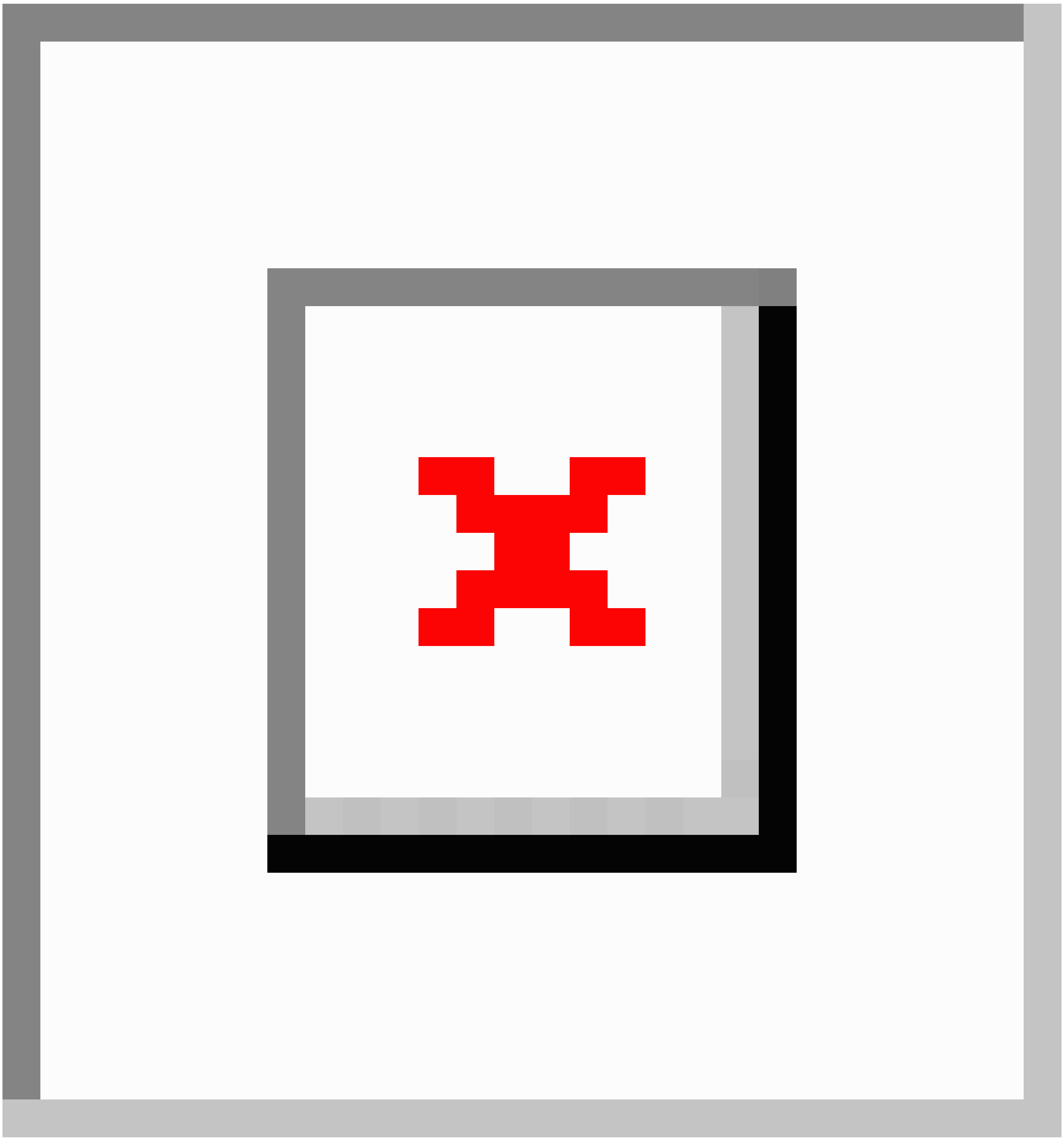
Variable		NTP use group		Total (N=108)
		Nondaily (n=33)	Daily (n=75)	
Age (years), mean (SD)		30.76 (12.58)	31.92 (12.75)	31.56 (12.65)
Sex, n (%)				
	Female	14 (42.4)	33 (44.0)	47 (43.5)
	Male	19 (57.6)	42 (56.0)	61 (56.5)
Ethnicity: White, n (%)		21 (63.6)	39 (52.0)	60 (55.6)
Education: college level, n (%)		31 (93.9)	63 (84.0)	94 (87.0)
Previous VR[a] experience, n (%)				
	Never	17 (51.5)	26 (34.7)	43 (39.8)
	Once	5 (15.2)	15 (20.0)	20 (18.5)
	A few times	7 (21.2)	27 (36.0)	34 (31.5)
	Many times	4 (12.1)	7 (9.3)	11 (10.2)
Combustible tobacco product user (predominately), n (%)		18 (54.5)	29 (38.7)	47 (43.5)
NTP use days (previous 90 days)[b], mean (SD)		33.91 (21.57)	89.79 (0.76)	72.71 (28.43)
NTP use episodes (previous 90 days)[b], mean (SD)		161.82 (180.85)	2145.92 (2633.99)	1539.67 (2376.99)
PATH[c] tobacco dependence index[b], mean (SD)		15.94 (11.70)	46.15 (17.54)	36.92 (21.19)
Tobacco Craving Questionnaire (baseline)[b], mean (SD)		30.09 (15.71)	47.55 (14.43)	42.17 (16.84)
VR presence (IPQ[d]), mean (SD)				
	Spatial presence	64.06 (18.16)	65.56 (17.80)	65.10 (17.84)
	Involvement	59.47 (21.30)	56.72 (22.16)	57.56 (21.84)
	Experienced realism	44.44 (23.07)	48.67 (22.39)	47.38 (22.58)
	Total	55.99 (8.65)	56.98 (9.60)	56.68 (9.29)
VR-related sickness (VRSQ [e]), mean (SD)				
	Oculomotor	18.43 (14.09)	17.22 (18.90)	17.59 (17.51)
	Disorientation	15.91 (14.64)	12.11 (17.67)	13.27 (16.82)
	Total	17.17 (13.21)	14.67 (17.25)	15.43 (16.10)

[a]VR: virtual reality.
[b]Denotes significant group differences (*P*<.001).
[c]PATH: Population Assessment of Tobacco and Health.
[d]IPQ: Igroup Presence Questionnaire.
[e]VRSQ: Virtual Reality Sickness Questionnaire.

Subjective Craving

Results of the ANOVA investigating subjective craving during the paradigm revealed significant effects of scene condition ($F_{1,106}$=47.95; P<.001; η_p^2=0.31) and NTP use group ($F_{1,106}$=16.91; P<.001; η_p^2=0.14) on craving. No interaction between scene condition and NTP use group was observed ($F_{1,106}$=0.03; P=.87; η_p^2<0.001). Specifically, active scenes (EMM 46.50, SE 3.11) elicited greater subjective craving than neutral scenes (EMM 31.89, SE 2.92; Figure 2), and daily users reported greater subjective craving across scenes (EMM 50.81, SE 3.12) than nondaily users (EMM 27.59, SE 4.70). Controlling for age and sex in this analysis reduced the main effect of scene condition ($F_{1,104}$=4.16; P=.04; η_p^2=0.04) but not NTP use group ($F_{1,104}$=17.63; P<.001; η_p^2=0.15). Age and sex were not found to be significant predictors of subjective craving either via direct effects or interactions (*P* values >.05). Subjective craving ratings were found to positively correlate between the ratings following the 2 active scenes (r=0.844; P<.001) and between the ratings following the 2 neutral scenes (r=0.816; P<.001).

Figure 2. Mean subjective craving ratings averaged across active and neutral scenes by NTP use group. Error bars indicate a 95% CI. NTP: nicotine and tobacco product.



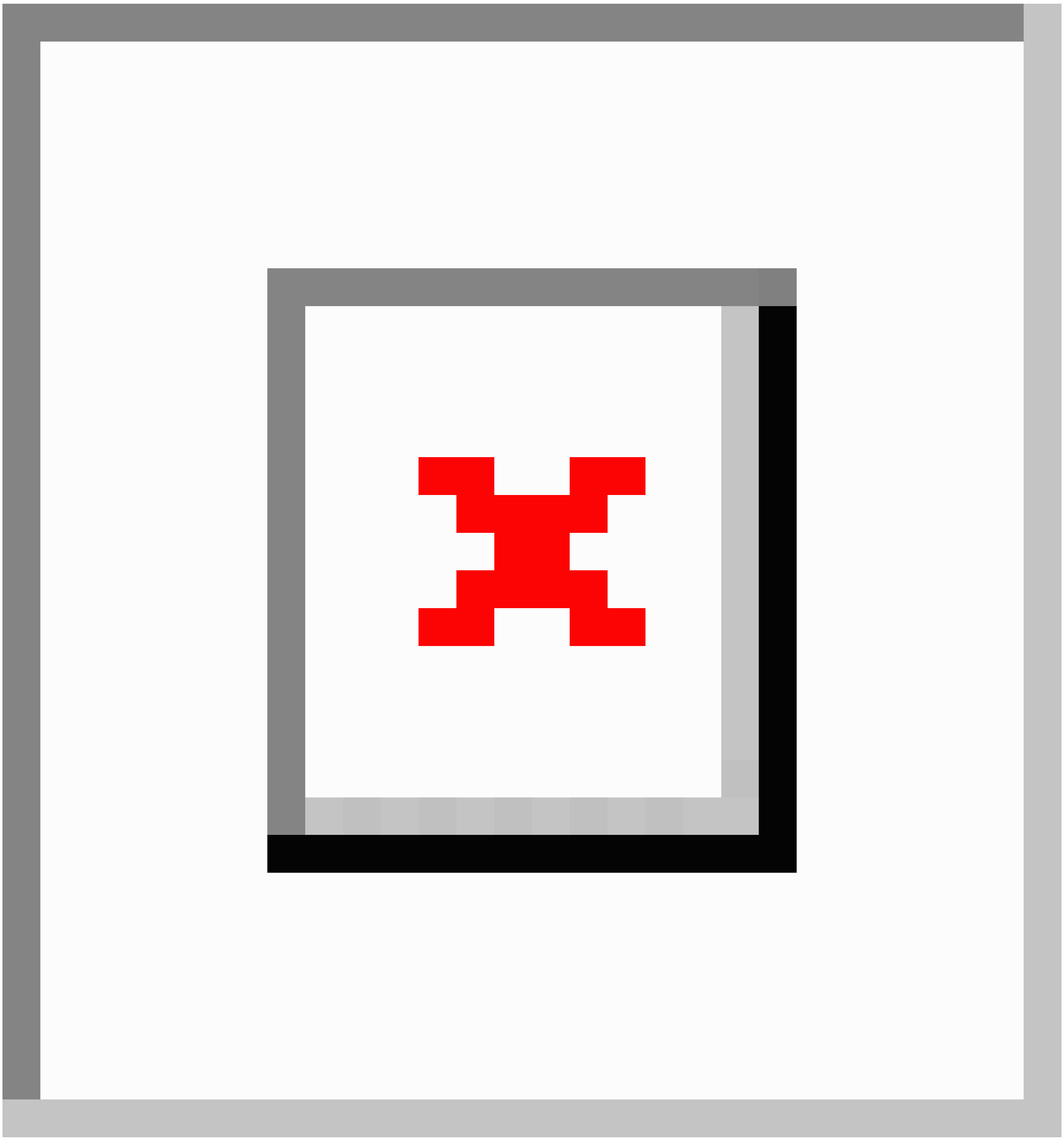
Total Cue Eye-Gaze Fixations

Results of the ANOVA investigating total cue eye-gaze fixations during the paradigm revealed a significant effect of cue type on number of fixations during the active scenes ($F_{1,106}$=1353.18; *P*<.001; η_p^2=0.93). No effect of NTP use group ($F_{1,106}$=0.03; *P*=.85; η_p^2<0.001) or interaction between cue type and NTP use group ($F_{1,106}$<0.001; *P*=.98; η_p^2<0.001) were observed. Specifically, NTP cues were associated with fewer total fixations (EMM 12.91, SE 0.21) as compared to control cues (EMM 48.11, SD 1.00). Controlling for age and sex in this analysis reduced, but did not eliminate, the main effect of cue type ($F_{1,104}$=175.04; *P*<.001; η_p^2=0.63). Further, there was a significant interaction observed between cue type and sex ($F_{1,104}$=8.40; *P*=.005; η_p^2=0.07), whereby male participants engaged in more total fixations toward control cues (EMM 50.07, SE 1.19) as compared to female participants (EMM 45.67, SE 1.35), yet their total fixations toward NTP cues were similar (male participants: EMM 12.69, SE 0.24; female participants: EMM 13.28, SE 0.28). Borderline main effects of age ($F_{1,104}$=3.69; *P*=.06; η_p^2=0.03) and sex ($F_{1,104}$=3.81; *P*=.05; η_p^2=0.03) were also observed. When comparing the 2 active scenes, total cue eye-gaze fixations were found to positively correlate for the NTP cues (r=0.347; P<.001) and control cues (r=0.657; P<.001).

Mean Eye-Gaze Fixation Time (Attentional Bias)

Results of the ANOVA investigating mean eye-gaze fixation time (attentional bias) during the paradigm revealed a significant effect of cue type on fixation time during the active scenes ($F_{1,106}$=48.34; P<.001; η_p^2=0.31) and some support for a NTP use group effect ($F_{1,106}$=3.31; P=.07; η_p^2=0.03; Figure 3). No interaction between cue type and NTP use group was observed ($F_{1,106}$=0.31; P=.58; η_p^2=0.003). Specifically, NTP cues were associated with greater mean fixation times (EMM 3557.79, SE 163.36 ms) as compared to control cues (EMM 2225.54, SE 87.34 ms), and daily users demonstrated greater mean fixation times across cue type (EMM 3054.09, SE 98.75) as compared to nondaily users (EMM 2729.24, SE 148.87). Controlling for age and sex in this analysis reduced, but did not eliminate, the main effect of cue type ($F_{1,104}$=14.32; P<.001; η_p^2=0.12), and the NTP use group effect was largely unchanged ($F_{1,104}$=3.53; P=.06; η_p^2=0.03). Age and sex were not found to be significant predictors of mean eye-gaze fixation time either via direct effects or interactions (P values >.05). When comparing the 2 active scenes, mean eye-gaze fixation times were found to positively correlate for the NTP cues (r=0.261; P=.006) and control cues (r=0.462; P<.001).

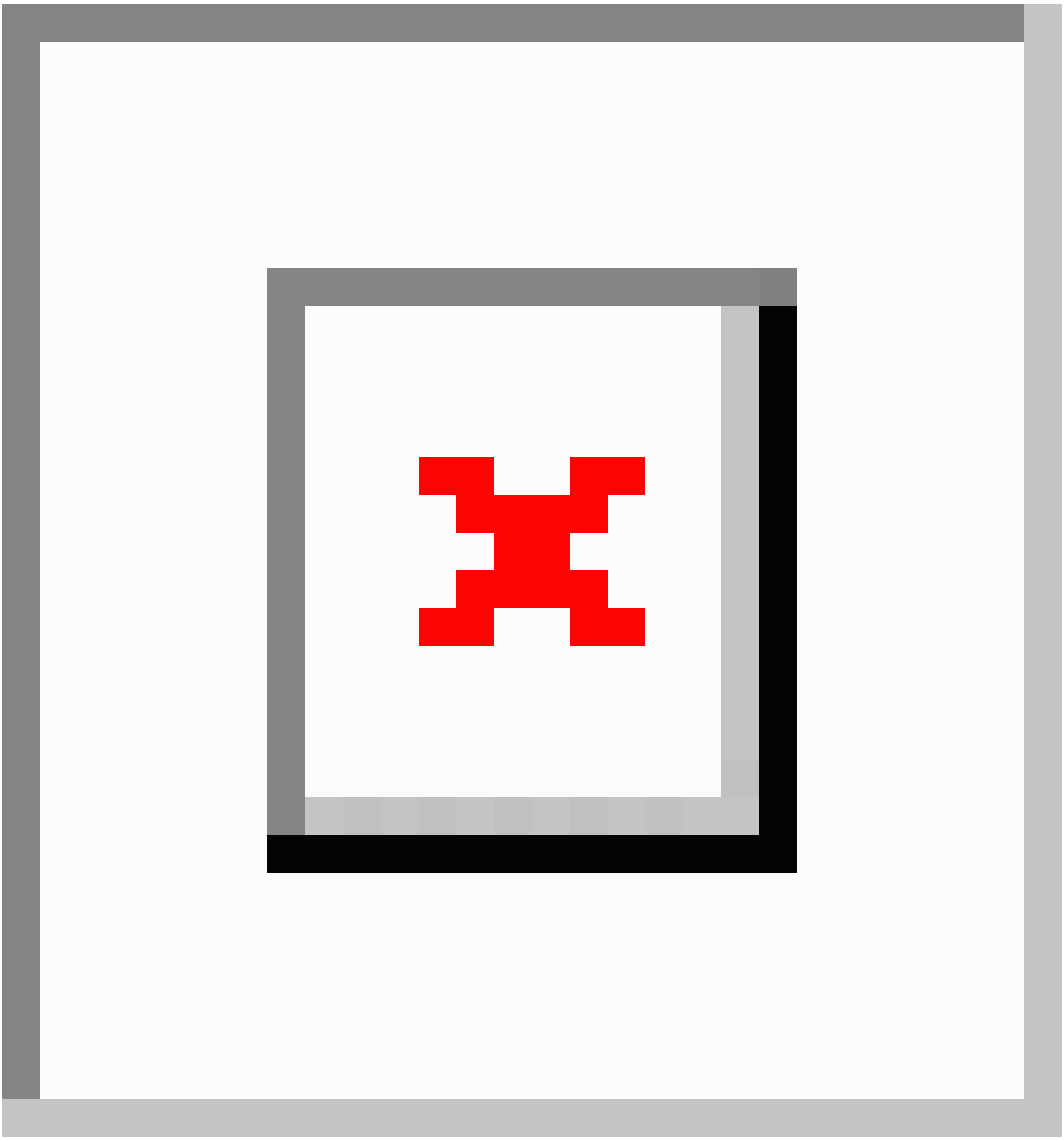
Figure 3. Mean eye-gaze fixation time averaged across NTP and control cues by NTP use group. Error bars indicate a 95% CI. NTP: nicotine and tobacco product.



Pupil Diameter

Results of the ANOVA investigating mean pupil diameter during the paradigm revealed a significant effect of cue type on mean pupil diameter during the active scenes ($F_{1,102}$=5.99; *P*=.02; η_p^2=0.05) and some support for an interaction between cue type and NTP use group ($F_{1,102}$=3.38; *P*=.07; η_p^2=0.03; Figure 4). No NTP use group main effect was observed ($F_{1,102}$<0.001; *P*=.99; η_p^2<0.001). Specifically, NTP cues were associated with greater pupil diameter (EMM 3.86, SE 0.07 mm) as compared to control cues (EMM 3.81, SE 0.07 mm) across groups, but only the nondaily use group displayed a significant difference between cue types (nondaily mean difference 0.09, SE 0.03; *P*=.01; daily mean difference 0.01, SE 0.02; *P*=.57). Controlling for age and sex in this analysis reduced the main effect of cue type ($F_{1,100}$=2.25; *P*=.14; η_p^2=0.02) and the interaction between cue type and NTP use group ($F_{1,100}$=3.26; *P*=.07; η_p^2=0.03). A main effect of age on pupil diameter was also observed ($F_{1,100}$=18.36; *P*<.001; η_p^2=0.16). When comparing the 2 active scenes, pupil diameters were found to positively correlate for the NTP cues (*r*=0.793; *P*<.001) and control cues (*r*=0.765; *P*<.001).

Figure 4. Mean pupil diameter averaged across NTP and control cues by NTP use group. Error bars indicate a 95% CI. NTP: nicotine and tobacco product.



Spontaneous Eyeblink Rate

Results of the ANOVA investigating mean EBR found no significant differences between EBR during active and neutral scenes ($F_{1,104}$=0.50; *P*=.48; η_p^2=0.005) or between NTP use groups ($F_{1,104}$=0.17; *P*=.68; η_p^2=0.002), nor a significant interaction ($F_{1,104}$=0.37; *P*=.54; η_p^2=0.004). Controlling for age and sex in this analysis resulted in no change to the relationships. Age and sex were not found to be significant predictors of EBR either via direct effects or interactions (*P* values >.05). EBRs were found to positively correlate between the 2 active scenes (*r*=0.635; *P*<.001) and between the 2 neutral scenes (*r*=0.567; *P*<.001).

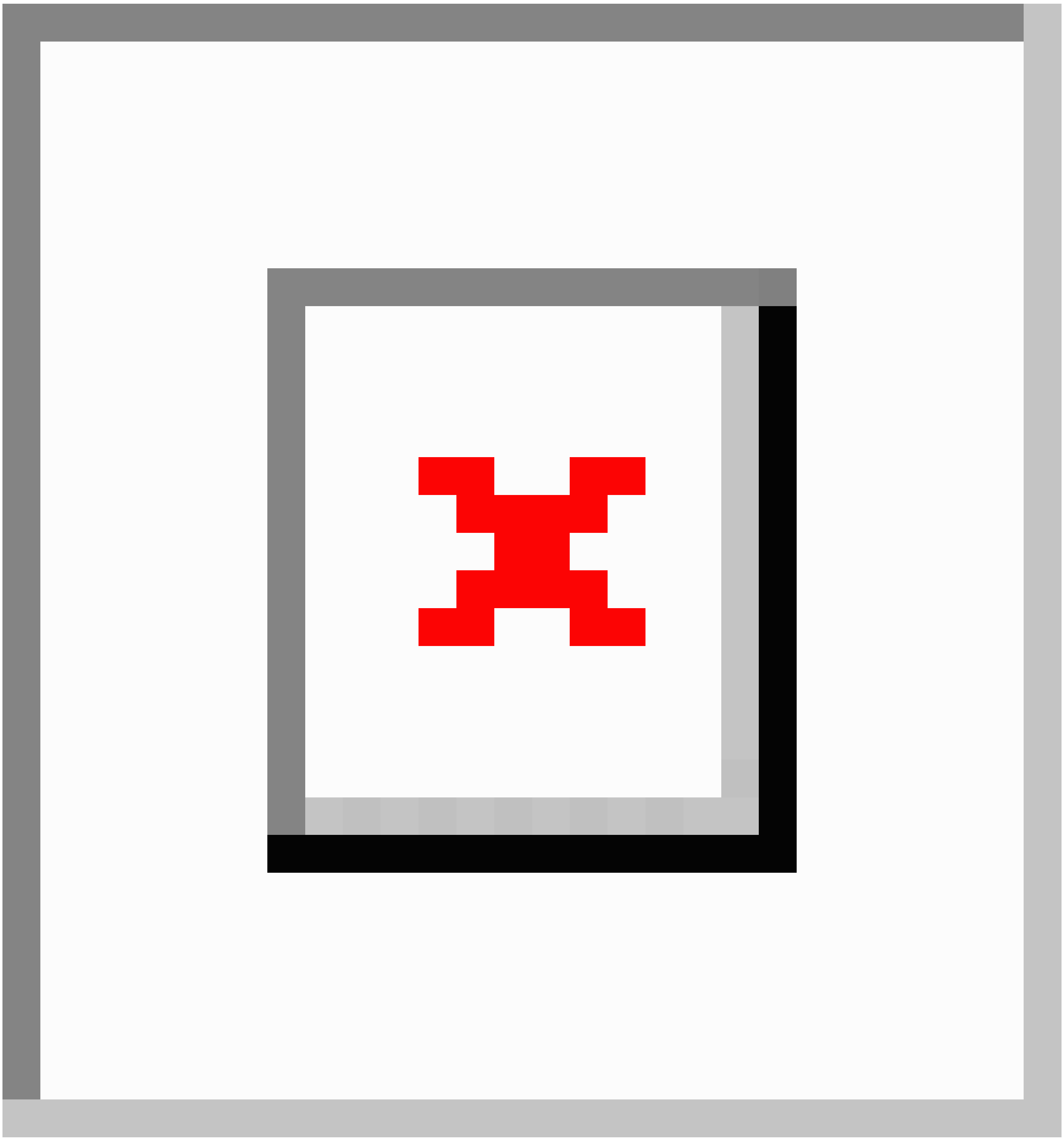
Relationship to NTP Subjective Craving and Use

Mean NTP versus control cue fixation time contrast scores (attentional bias scores) were found to positively correlate with subjective craving assessed within the paradigm (*r*=0.19; *P*=.04) and with the TCQ-SF administered pre- (*r*=0.18; *P*=.06) and post-VR paradigm (*r*=0.22; *P*=.02). Comparison of correlations between NTP use groups demonstrated a significant group difference (Z_{obs}=2.48; *P*=.007), with the nondaily group demonstrating a stronger positive correlation (*r*=0.57; *P*=.001) compared to the daily group (*r*=0.10; *P*=.38) in TCQ-SF scores post-VR paradigm (Figure 5). Similar group relationships held for the other craving metrics. Follow-up analyses investigating mean cue fixation time separately by cue type (NTP and control) in the full sample revealed that the positive correlations with all three subjective craving ratings were driven primarily by mean NTP cue fixation times (*r* values=0.21-0.28; *P* values <.03), as opposed to control cue fixation times (*r* values=0.02-0.03; *P* values >.76). The mean NTP versus control cue fixation time contrast score (attentional bias) was not found to correlate with previous 90-day NTP use (*r*=0.05; *P*=.58), yet significant positive correlations with previous 90-day NTP use were observed for mean NTP cue fixation time (*r*=0.20; *P*=.03) and control cue fixation time (*r*=0.26; *P*=.007) when analyzed independently.

Mean NTP versus control cue pupil diameter contrast scores were not found to correlate with subjective craving measures across the full sample (*P* values >.20) or NTP use groups independently (*P* values >.07). Mean NTP versus control cue pupil diameter was found to negatively correlate with previous 90-day NTP use (*r*=–0.20; *P*=.04), yet no significant correlations with previous 90-day NTP use were observed for the NTP (*r*=0.09; *P*=.37) or control cue (*r*=0.15; *P*=.12) pupil diameters independently.

Partial correlations controlling for age and sex resulted in similar estimates for all analyses described above except for the correlation between mean NTP versus control cue pupil diameter contrast scores and previous 90-day NTP use, which was eliminated when age and sex were controlled for ($r_{partial}$=−0.05; *P*=.64). Total cue fixations and eyeblink rates were not found to correlate with any subjective craving measures or previous 90-day NTP use (*P* values >.05). None of the objective measures significantly correlated with recency of NTP use (*P* values >.05). None of the first-level correlations survived multiple comparison correction (Bonferroni-corrected per objective measure *P* values <.01) and, as such, must be interpreted with caution.

Figure 5. Scatterplot depicting the linear relationships between mean NTP versus control cue eye gaze fixation time contrast scores (in milliseconds) from the active scenes and subjective craving from the TCQ-SF administered post-VR paradigm by NTP use group. NTP: nicotine and tobacco product; TCQ-SF: Tobacco Craving Questionnaire—Short Form; VR: virtual reality.



Discussion

This study provides updated results on the utility of a novel VR NTP cue exposure paradigm to index incentive salience via three eye characteristic markers: eye-gaze fixation time (attentional bias), pupil diameter, and EBR in response to NTP versus control cues. Overall, the results are largely consistent with our preliminary report [44] and support two of our initial hypotheses, suggesting that measures of eye-gaze fixation time and pupil diameter, but not EBR, during VR cue exposure could be useful objective indicators of the incentive salience process in nicotine addiction.

Consistent with previous VR cue exposure investigations across a variety of substances [14-21], active VR scenes with NTP cues elicited greater subjective craving compared to neutral control scenes. Further, daily NTP users endorsed greater overall levels of subjective craving compared to nondaily users across scenes. Together these findings suggest the VR NTP cue exposure paradigm elicits subjective phasic craving in response to NTP cues and can discriminate by frequency of NTP use on this metric.

Given that the intention of the paradigm was to provide a more naturalistic and translatable context of use than standard cue exposure and attentional bias paradigms, it follows that more control versus NTP cues were present in the active scenes. As a result, greater total eye-gaze fixations toward control versus NTP cues were observed. Yet, the average gaze fixation time was found to be 1.33 seconds longer for the NTP cues compared with the control cues across the full sample, thus demonstrating attentional bias toward the NTP cues regardless of NTP use frequency. The attentional bias contrast score (mean NTP vs control cue fixation time) was also modestly, yet consistently, associated with measures of subjective craving assessed before, during, and after the VR paradigm. This was primarily driven by response to the NTP cues, as opposed to the control cues, supporting the previously established link between attentional bias, as indexed by fixation time, and subjective craving [10,30,64-66]. The culmination of fixation time results also supports the validity of the VR NTP cue exposure paradigm as suitable for measuring attentional bias toward NTP cues in a free-viewing, translatable, and ecologically valid context.

Interestingly, although no interaction between NTP use group and cue type was observed, the daily NTP users were found to fixate on all cues (NTP and control) longer (325 ms) than the nondaily users. Furthermore, no association was observed between the attentional bias contrast score and previous NTP use frequency, yet greater mean gaze fixation time to NTP and control cues were *independently* associated with greater NTP use in the previous 90 days. These results are somewhat contradictory to the findings by Mogg and colleagues [67], where greater smoking versus control fixation times were inversely associated with nicotine dependence. This discrepancy could relate to differences in tasks, as Mogg et al [67] used a visual probe task to assess eye-gaze fixation time, which presented cues in isolation, devoid of context and additional competing cues. Additionally, independent associations between cue types and dependence severity were not reported in their paper; thus, it remains unknown whether a similar relationship would have been observed in their data. Regardless, our results suggest that in the presence of additional naturalistic context and the absence of any researcher-directed task demands, individuals with varying levels of nicotine dependence evince attentional bias toward NTP cues and more frequent/dependent NTP users demonstrate prolonged attentional engagement with all salient visual cues present.

Consistent with Mogg and colleagues [67] and with subjective craving associations broadly [68], we observed stronger correlations between the attentional bias contrast score and subjective craving levels within the nondaily NTP users, as compared to the daily users. Relatedly, greater pupillary diameter was observed in response to NTP cues compared to control cues, particularly within the nondaily users. Interestingly, the NTP versus control cue pupillary response contrast was found to be negatively associated with previous NTP use, although these effects were essentially eliminated after controlling for age, a known correlate of pupil size [69]. These findings are in line with theories suggesting that appetitive motivational processes (ie, incentive salience) reduce in importance as addiction becomes more severe and habitual [5,6,67].

Taken together with our attentional bias and NTP use data, there may be additional nonselective attentional processes occurring in individuals with more severe nicotine dependence that are not routinely captured by traditional subjective and procedural tasks of attention. For example, our results may reflect an effect of prolonged nicotine use on general attentional processing in the absence of task demands and trial durations, whereby individuals with more prolonged NTP use may have delayed disengagement from any salient cue in their visual environment. These correlations appear to hold even after controlling for total number of cue fixations, suggesting this is not a product of orientation bias. Traditional tasks used to investigate attentional bias (eg, Stroop and visual probe tasks) are thought to index the delayed disengagement of attention; yet their ability to do so is limited by trial carryover effects, short durations of stimulus onset asynchrony (SOA), and task demands (eg, to shift attention based on cue location [24]). Even with tasks thought to explicitly measure disengagement, the SOA is often only 500 - 2000 milliseconds [24], yet our data suggest that when free-viewing a complex scene with many cues present, individuals spend on average 2955 milliseconds engaged and attending to one object cue irrespective of NTP use history and cue type. Thus, further investigations using these paradigms to assess attentional disengagement may benefit from increasing SOA beyond 3000 milliseconds to ensure they are capturing the entire disengagement process. Given that acute nicotine administration facilitates attention disengagement from a cued location [70], it may be possible that the reverse effect is occurring during the state of acute withdrawal in heavier users of NTPs.

This study has several strengths and limitations. Strengths include the inline assessment of eye characteristics during a translatable real-world VR NTP cue exposure paradigm with no imposed task demands, thus more accurately indexing naturalistic attentional and incentive salience processes. The inclusion of light to heavy users of various NTPs and ages increases the generalizability of the findings to the majority of current nicotine users. Limitations include the absence of biological verification to confirm self-reported NTP use due to COVID-19–related precautions, absence of prospective NTP use data, and the short duration of abstinence at the time of testing. However, substantial variability in abstinence was reported and abstinence time was not found to substantially impact the results. Still, studies investigating these effects at much longer durations of abstinence and in treatment-seeking populations may observe differing results as the salience of cues may change based on extended abstinence. Given that the active scenes differ in the amount and nature of the cues (both NTP and control), additional studies with multiple identical administrations and with prospective NTP use data are needed to adequately assess the reliability and validity of these eye-tracking indices. Lastly, given the relationship between increasing age and potential for greater addiction severity (eg, allowing for greater years of use with increasing age), future studies are needed to identify the independent contributions of age and eye-related variables (especially pupil size [69]) on NTP use outcomes.

In summary, this study represents an update to our initial paper [44], provides validation of the utility of the VR NTP cue exposure paradigm for the assessment of attentional bias as measured via eye-gaze fixation time and pupillometry, and highlights areas for further consideration in other attentional bias paradigms (eg, increasing SOA). Given that attentional bias has been shown to predict relapse following smoking cessation [28], these markers may prove useful in clinical settings by facilitating the matching of individuals who exhibit greater attentional bias with interventions targeting incentive salience processes (eg, varenicline [16,18] and mindfulness [71]). The validation of reliable biomarkers of addiction such as attentional bias could also greatly benefit treatment development by providing an earlier identification of treatment efficacy (a "fast fail" marker) in clinical trials. Broadly, markers such as attentional bias and pupil diameter have the potential to provide much needed objective measures of addiction phenotypes, thus reducing error associated with phenotyping and outcomes measurement based solely on subjective assessments.

Acknowledgments

This research was supported by the California Tobacco-Related Disease Research Grants Program Office of the University of California grant T30IP0962 and National Institutes of Health grant UL1TR001442.

Conflicts of Interest

JS is a paid employee at Qualcomm. The remaining authors have no relevant financial or nonfinancial interests to disclose.

References

1. Wiers RW, Bartholow BD, van den Wildenberg E, et al. Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. Pharmacol Biochem Behav 2007 Feb;86(2):263-283. [doi: 10.1016/j.pbb.2006.09.021] [Medline: 17116324]
2. Stacy AW, Wiers RW. Implicit cognition and addiction: a tool for explaining paradoxical behavior. Annu Rev Clin Psychol 2010;6:551-575. [doi: 10.1146/annurev.clinpsy.121208.131444] [Medline: 20192786]
3. Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Res Brain Res Rev 1993;18(3):247-291. [doi: 10.1016/0165-0173(93)90013-p] [Medline: 8401595]
4. Robinson TE, Berridge KC. Addiction. Annu Rev Psychol 2003 Jun;54:25-53. [doi: 10.1146/annurev.psych.54.101601.145237] [Medline: 12185211]
5. Tiffany ST. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. Psychol Rev 1990 Apr;97(2):147-168. [doi: 10.1037/0033-295x.97.2.147] [Medline: 2186423]
6. Di Chiara G. Role of dopamine in the behavioural actions of nicotine related to addiction. Eur J Pharmacol 2000 Mar 30;393(1-3):295-314. [doi: 10.1016/s0014-2999(00)00122-9] [Medline: 10771025]
7. Carter BL, Tiffany ST. Meta-analysis of cue-reactivity in addiction research. Addiction 1999 Mar;94(3):327-340. [Medline: 10605857]
8. Berridge KC. From prediction error to incentive salience: mesolimbic computation of reward motivation. Eur J Neurosci 2012 Apr;35(7):1124-1143. [doi: 10.1111/j.1460-9568.2012.07990.x] [Medline: 22487042]
9. Wray JM, Godleski SA, Tiffany ST. Cue-reactivity in the natural environment of cigarette smokers: the impact of photographic and in vivo smoking stimuli. Psychol Addict Behav 2011 Dec;25(4):733-737. [doi: 10.1037/a0023687] [Medline: 21553947]
10. Creswell KG, Skrzynski CJ. The influence of smoking motivation on the associations among cigarette craving, attentional bias to smoking cues, and smoking behavior. Nicotine Tob Res 2021 Aug 29;23(10):1727-1734. [doi: 10.1093/ntr/ntab028] [Medline: 33599777]
11. Wilson SJ, Sayette MA. Neuroimaging craving: urge intensity matters. Addiction 2015 Feb;110(2):195-203. [doi: 10.1111/add.12676] [Medline: 25073979]
12. Hone-Blanchet A, Wensing T, Fecteau S. The use of virtual reality in craving assessment and cue-exposure therapy in substance use disorders. Front Hum Neurosci 2014 Oct 17;8:844. [doi: 10.3389/fnhum.2014.00844] [Medline: 25368571]
13. Bouton ME. A learning theory perspective on lapse, relapse, and the maintenance of behavior change. Health Psychol 2000 Jan;19(1S):57-63. [doi: 10.1037/0278-6133.19.suppl1.57] [Medline: 10709948]
14. Bordnick PS, Traylor A, Copp HL, et al. Assessing reactivity to virtual reality alcohol based cues. Addict Behav 2008 Jun;33(6):743-756. [doi: 10.1016/j.addbeh.2007.12.010] [Medline: 18282663]
15. Ryan JJ, Kreiner DS, Chapman MD, Stark-Wroblewski K. Virtual reality cues for binge drinking in college students. Cyberpsychol Behav Soc Netw 2010 Apr;13(2):159-162. [doi: 10.1089/cyber.2009.0211] [Medline: 20528271]
16. Ghiţă A, Gutiérrez-Maldonado J. Applications of virtual reality in individuals with alcohol misuse: a systematic review. Addict Behav 2018 Jun;81:1-11. [doi: 10.1016/j.addbeh.2018.01.036] [Medline: 29421343]
17. Bordnick PS, Copp HL, Traylor A, et al. Reactivity to cannabis cues in virtual reality environments. J Psychoactive Drugs 2009 Jun;41(2):105-112. [doi: 10.1080/02791072.2009.10399903] [Medline: 19705672]
18. Culbertson C, Nicolas S, Zaharovits I, et al. Methamphetamine craving induced in an online virtual reality environment. Pharmacol Biochem Behav 2010 Oct;96(4):454-460. [doi: 10.1016/j.pbb.2010.07.005] [Medline: 20643158]

19. Traylor AC, Parrish DE, Copp HL, Bordnick PS. Using virtual reality to investigate complex and contextual cue reactivity in nicotine dependent problem drinkers. Addict Behav 2011 Nov;36(11):1068-1075. [doi: 10.1016/j.addbeh.2011.06.014] [Medline: 21783326]
20. Acker J, MacKillop J. Behavioral economic analysis of cue-elicited craving for tobacco: a virtual reality study. Nicotine Tob Res 2013 Aug;15(8):1409-1416. [doi: 10.1093/ntr/nts341] [Medline: 23322768]
21. Baumann SB, Sayette MA. Smoking cues in a virtual world provoke craving in cigarette smokers. Psychol Addict Behav 2006 Dec;20(4):484-489. [doi: 10.1037/0893-164X.20.4.484] [Medline: 17176184]
22. Tamburin S, Dal Lago D, Armani F, et al. Smoking-related cue reactivity in a virtual reality setting: association between craving and EEG measures. Psychopharmacology (Berl) 2021 May;238(5):1363-1371. [doi: 10.1007/s00213-020-05733-3] [Medline: 33263158]
23. Lebiecka Z, Skoneczny T, Tyburski E, Samochowiec J, Kucharska-Mazur J. Is virtual reality cue exposure a promising adjunctive treatment for alcohol use disorder? J Clin Med 2021 Jul 1;10(13):2972. [doi: 10.3390/jcm10132972] [Medline: 34279455]
24. Field M, Munafò MR, Franken IHA. A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. Psychol Bull 2009 Jul;135(4):589-607. [doi: 10.1037/a0015843] [Medline: 19586163]
25. Ryan F. Detected, selected, and sometimes neglected: cognitive processing of cues in addiction. Exp Clin Psychopharmacol 2002 May;10(2):67-76. [doi: 10.1037//1064-1297.10.2.67] [Medline: 12022800]
26. Franken IHA. Drug craving and addiction: integrating psychological and neuropsychopharmacological approaches. Prog Neuropsychopharmacol Biol Psychiatry 2003 Jun;27(4):563-579. [doi: 10.1016/S0278-5846(03)00081-2] [Medline: 12787841]
27. Anselme P, Robinson MJF. From sign-tracking to attentional bias: implications for gambling and substance use disorders. Prog Neuropsychopharmacol Biol Psychiatry 2020 Apr 20;99:109861. [doi: 10.1016/j.pnpbp.2020.109861] [Medline: 31931091]
28. Waters AJ, Shiffman S, Sayette MA, Paty JA, Gwaltney CJ, Balabanis MH. Attentional bias predicts outcome in smoking cessation. Health Psychol 2003 Jul;22(4):378-387. [doi: 10.1037/0278-6133.22.4.378] [Medline: 12940394]
29. Powell J, Dawkins L, West R, Powell J, Pickering A. Relapse to smoking during unaided cessation: clinical, cognitive and motivational predictors. Psychopharmacology (Berl) 2010 Dec;212(4):537-549. [doi: 10.1007/s00213-010-1975-8] [Medline: 20703450]
30. Soleymani A, Ivanov Y, Mathot S, de Jong PJ. Free-viewing multi-stimulus eye tracking task to index attention bias for alcohol versus soda cues: satisfactory reliability and criterion validity. Addict Behav 2020 Jan;100:106117. [doi: 10.1016/j.addbeh.2019.106117] [Medline: 31522132]
31. Marks KR, Pike E, Stoops WW, Rush CR. The magnitude of drug attentional bias is specific to substance use disorder. Psychol Addict Behav 2015 Sep;29(3):690-695. [doi: 10.1037/adb0000084] [Medline: 25961149]
32. Marks KR, Roberts W, Stoops WW, Pike E, Fillmore MT, Rush CR. Fixation time is a sensitive measure of cocaine cue attentional bias. Addiction 2014 Sep;109(9):1501-1508. [doi: 10.1111/add.12635] [Medline: 24894879]
33. Marks KR, Pike E, Stoops WW, Rush CR. Test-retest reliability of eye tracking during the visual probe task in cocaine-using adults. Drug Alcohol Depend 2014 Dec 1;145:235-237. [doi: 10.1016/j.drugalcdep.2014.09.784] [Medline: 25456573]
34. Spanakis P, Jones A, Field M, Christiansen P. A stroop in the hand is worth two on the laptop: superior reliability of a smartphone based alcohol stroop in the real world. Subst Use Misuse 2019;54(4):692-698. [doi: 10.1080/10826084.2018.1536716] [Medline: 30572780]
35. Monem RG, Fillmore MT. Measuring heightened attention to alcohol in a naturalistic setting: a validation study. Exp Clin Psychopharmacol 2017 Dec;25(6):496-502. [doi: 10.1037/pha0000157] [Medline: 29251979]
36. Bijleveld E, Custers R, Aarts H. The unconscious eye opener: pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues. Psychol Sci 2009 Nov;20(11):1313-1315. [doi: 10.1111/j.1467-9280.2009.02443.x] [Medline: 19788532]
37. Muhammed K, Manohar S, Ben Yehuda M, et al. Reward sensitivity deficits modulated by dopamine are associated with apathy in Parkinson’s disease. Brain 2016 Oct;139(Pt 10):2706-2721. [doi: 10.1093/brain/aww188] [Medline: 27452600]
38. Anderson BA, Yantis S. Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. Atten Percept Psychophys 2012 Nov;74(8):1644-1653. [doi: 10.3758/s13414-012-0348-2] [Medline: 22810561]
39. Laeng B, Sirois S, Gredebäck G. Pupillometry: a window to the preconscious? Perspect Psychol Sci 2012 Jan 5;7(1):18-27. [doi: 10.1177/1745691611427305] [Medline: 26168419]
40. Kvamme TL, Pedersen MU, Overgaard M, Rømer Thomsen K, Voon V. Pupillary reactivity to alcohol cues as a predictive biomarker of alcohol relapse following treatment in a pilot study. Psychopharmacology (Berl) 2019 Apr;236(4):1233-1243. [doi: 10.1007/s00213-018-5131-1] [Medline: 30607476]
41. Jongkees BJ, Colzato LS. Spontaneous eye blink rate as predictor of dopamine-related cognitive function-a review. Neurosci Biobehav Rev 2016 Dec;71:58-82. [doi: 10.1016/j.neubiorev.2016.08.020] [Medline: 27555290]
42. Kaminer J, Powers AS, Horn KG, Hui C, Evinger C. Characterizing the spontaneous blink generator: an animal model. J Neurosci 2011 Aug 3;31(31):11256-11267. [doi: 10.1523/JNEUROSCI.6218-10.2011] [Medline: 21813686]

43. Zijlstra F, Booij J, van den Brink W, Franken IHA. Striatal dopamine D2 receptor binding and dopamine release during cue-elicited craving in recently abstinent opiate-dependent males. Eur Neuropsychopharmacol 2008 Apr;18(4):262-270. [doi: 10.1016/j.euroneuro.2007.11.002] [Medline: 18077142]
44. Liu W, Andrade G, Schulze J, Doran N, Courtney KE. Using virtual reality to induce and assess objective correlates of nicotine craving: paradigm development study. JMIR Serious Games 2022 Feb 15;10(1):e32243. [doi: 10.2196/32243] [Medline: 35166685]
45. Shiffman S, Ferguson SG, Dunbar MS, Scholl SM. Tobacco dependence among intermittent smokers. Nicotine Tob Res 2012 Nov;14(11):1372-1381. [doi: 10.1093/ntr/nts097] [Medline: 22529224]
46. Shiffman S, Tindle H, Li X, Scholl S, Dunbar M, Mitchell-Miland C. Characteristics and smoking patterns of intermittent smokers. Exp Clin Psychopharmacol 2012 Aug;20(4):264-277. [doi: 10.1037/a0027546] [Medline: 22390208]
47. Shiffman S, Dunbar MS, Kirchner TR, et al. Cue reactivity in non-daily smokers: effects on craving and on smoking behavior. Psychopharmacology (Berl) 2013 Mar;226(2):321-333. [doi: 10.1007/s00213-012-2909-4] [Medline: 23142992]
48. Jones A, Hogarth L, Christiansen P, Rose AK, Martinovic J, Field M. Reward expectancy promotes generalized increases in attentional bias for rewarding stimuli. Q J Exp Psychol (Hove) 2012 May;65(12):2333-2342. [doi: 10.1080/17470218.2012.686513] [Medline: 22631033]
49. Le Houezec J. Role of nicotine pharmacokinetics in nicotine addiction and nicotine replacement therapy: a review. Int J Tuberc Lung Dis 2003 Sep;7(9):811-819. [Medline: 12971663]
50. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33. [Medline: 9881538]
51. Sobell L, Sobell M. Timeline Followback: User’s Guide: Addiction Research Foundation; 1996.
52. Brown SA, Myers MG, Lippke L, Tapert SF, Stewart DG, Vik PW. Psychometric evaluation of the customary drinking and drug use record (CDDR): a measure of adolescent alcohol and drug involvement. J Stud Alcohol 1998 Jul;59(4):427-438. [doi: 10.15288/jsa.1998.59.427] [Medline: 9647425]
53. Robinson SM, Sobell LC, Sobell MB, Leo GI. Reliability of the timeline followback for cocaine, cannabis, and cigarette use. Psychol Addict Behav 2014 Mar;28(1):154-162. [doi: 10.1037/a0030992] [Medline: 23276315]
54. Strong DR, Messer K, Hartman SJ, et al. Measurement of multiple nicotine dependence domains among cigarette, non-cigarette and poly-tobacco users: insights from item response theory. Drug Alcohol Depend 2015 Jul 1;152:185-193. [doi: 10.1016/j.drugalcdep.2015.03.040] [Medline: 26005043]
55. Heishman SJ, Singleton EG, Pickworth WB. Reliability and validity of a short form of the tobacco craving questionnaire. Nicotine Tob Res 2008 Apr;10(4):643-651. [doi: 10.1080/14622200801908174] [Medline: 18418787]
56. Schubert T. The sense of presence in virtual environments: a three-component scale measuring spatial presence, involvement, and realness. Z Medienpsychologie 2003 Apr;15:69-71. [doi: 10.1026//1617-6383.15.2.69]
57. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aerosp Psychol 1993 Nov;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
58. Kim HK, Park J, Choi Y, Choe M. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl Ergon 2018 May;69:66-73. [doi: 10.1016/j.apergo.2017.12.016] [Medline: 29477332]
59. Eye and Facial Tracking SDK (legacy). VIVE. URL: https://developer-express.vive.com/resources/vive-sense/eye-and-facial-tracking-sdk/ [accessed 2024-06-17]
60. Creative Commons. Attribution 40 International (CC BY 40). URL: http://www.w3.org/1999/xlink [accessed 2024-06-24]
61. Liu W, Andrade G, Schulze J, Courtney KE. Evaluation of gaze-to-object mapping algorithms for use in “real-world” translatable neuropsychological paradigms. Psychol Neurosci 2023 Dec;16(4):339-348. [doi: 10.1037/pne0000324] [Medline: 38298524]
62. Caffier PP, Erdmann U, Ullsperger P. Experimental evaluation of eye-blink parameters as a drowsiness measure. Eur J Appl Physiol 2003 May;89(3-4):319-325. [doi: 10.1007/s00421-003-0807-5] [Medline: 12736840]
63. Schleicher R, Galley N, Briest S, Galley L. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics 2008 Jul;51(7):982-1010. [doi: 10.1080/00140130701817062] [Medline: 18568959]
64. Lochbuehler K, Wileyto EP, Tang KZ, Mercincavage M, Cappella JN, Strasser AA. Do current and former cigarette smokers have an attentional bias for e-cigarette cues? J Psychopharmacol 2018 Mar;32(3):316-323. [doi: 10.1177/0269881117728418] [Medline: 28927321]
65. Bollen Z, Masson N, Salvaggio S, D’Hondt F, Maurage P. Craving is everything: an eye-tracking exploration of attentional bias in binge drinking. J Psychopharmacol 2020 Jun;34(6):636-647. [doi: 10.1177/0269881120913131] [Medline: 32202459]
66. Kvamme TL, Pedersen MU, Overgaard M, Thomsen KR, Voon V. Pilot study: improving attention bias modification of alcohol cues through concealed gaze-contingent feedback in alcohol dependence. Addict Behav Rep 2019 Nov 6;10:100231. [doi: 10.1016/j.abrep.2019.100231] [Medline: 31832536]
67. Mogg K, Field M, Bradley BP. Attentional and approach biases for smoking cues in smokers: an investigation of competing theoretical views of addiction. Psychopharmacology (Berl) 2005 Jul;180(2):333-341. [doi: 10.1007/s00213-005-2158-x] [Medline: 15696322]

68. Gass JC, Motschman CA, Tiffany ST. The relationship between craving and tobacco use behavior in laboratory studies: a meta-analysis. Psychol Addict Behav 2014 Dec;28(4):1162-1176. [doi: 10.1037/a0036879] [Medline: 25134054]
69. Guillon M, Dumbleton K, Theodoratos P, Gobbe M, Wooley CB, Moody K. The effects of age, refractive status, and luminance on pupil size. Optom Vis Sci 2016 Sep;93(9):1093-1100. [doi: 10.1097/OPX.0000000000000893] [Medline: 27232893]
70. Murphy FC, Klein RM. The effects of nicotine on spatial and non-spatial expectancies in a covert orienting task. Neuropsychologia 1998 Nov;36(11):1103-1114. [doi: 10.1016/s0028-3932(98)00012-8] [Medline: 9842757]
71. de Souza ICW, de Barros VV, Gomide HP, et al. Mindfulness-based interventions for the treatment of smoking: a systematic literature review. J Altern Complement Med 2015 Mar;21(3):129-140. [doi: 10.1089/acm.2013.0471] [Medline: 25710798]

Abbreviations

***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition)
EBR: spontaneous eyeblink rate
EMM: estimated marginal mean
G2OM: Gaze-to-Object-Mapping
IPQ: Igroup Presence Questionnaire
NTP: nicotine and tobacco product
PATH: Population Assessment of Tobacco and Health
SOA: stimulus onset asynchrony
SSQ: Simulator Sickness Questionnaire
TCQ-SF: Tobacco Craving Questionnaire–Short Form
TLFB: timeline follow-back
VR: virtual reality

Edited by G Eysenbach; submitted 01.11.23; peer-reviewed by M Hirota, R Meng, Z Duan; revised version received 05.03.24; accepted 01.06.24; published 27.06.24.

Please cite as:
Courtney KE, Liu W, Andrade G, Schulze J, Doran N
Attentional Bias, Pupillometry, and Spontaneous Blink Rate: Eye Characteristic Assessment Within a Translatable Nicotine Cue Virtual Reality Paradigm
JMIR Serious Games 2024;12:e54220
URL: https://games.jmir.org/2024/1/e54220
doi:10.2196/54220

Original Paper

An Escape Game on University Students' Mental Health During the COVID-19 Pandemic: Cocreation Study

David Labrosse[1*], MD; Clara Vié[2*], MSc; Hana Hajjam[1], BSc; Clément Tisseron[2], BSc; Dimitri Thellier[1], MD; Ilaria Montagni[2], PhD

[1]Tricky, Bordeaux, France

[2]University of Bordeaux, Inserm, Bordeaux Population Health Research Center U1219, Bordeaux, France

[*]these authors contributed equally

Corresponding Author:
Ilaria Montagni, PhD
University of Bordeaux
Inserm
Bordeaux Population Health Research Center U1219
146 rue Léo Saignat
Bordeaux, 33000
France
Phone: 33 06 42 19 33 63
Email: ilaria.montagni@u-bordeaux.fr

Abstract

Background: The COVID-19 pandemic has had a severe impact on students' mental health. Interventions are needed to promote their psychological well-being and prevent mental illnesses in the aftermath of this unprecedented situation. Digital escape games can be an effective tool to support students' mental health. A cocreation approach can improve the acceptability of these interventions by involving different stakeholders (eg, end users, game designers, and health professionals) to obtain audience-specific games.

Objective: This study aims to describe the process of testing and optimizing the game "EscapeCovid" on students' mental health, to serve as a model for the cocreation of future similar interventions.

Methods: The PRODUCES (Problem, Objective, Design, End Users, Cocreators, Evaluation, Scalability) framework was used. Cocreation steps (test and optimization) were detailed for replicability. A total of 45 students tested a pilot version of the game, with 10 undergoing a semistructured interview. Meetings with a group of stakeholders and brainwriting were organized to optimize the game.

Results: We produced a new version of the game incorporating the suggestions provided by student testers and following the stakeholders' guidelines. Improvements were made to both the content and the form of the new version of the pilot game. The storyline, including the protagonist and the scenes, was adapted to the student population.

Conclusions: Our results suggested that cocreation can contribute to the design of more widely accepted interventions aimed at promoting mental health and preventing psychological disorders. Results also suggest that an end user–centered approach can facilitate intervention tailoring. When conceiving a health-related escape game for students, we recommend using the cocreation approach to enhance players' experience, thus positively influencing their learning process and overall well-being.

(JMIR Serious Games 2024;12:e48545) doi:10.2196/48545

KEYWORDS

students; game; mental health; COVID-19; cocreation; university student; promote; psychological well-being; effective tool; tool; acceptability; testing; questionnaire; learning; motivation; user centered

Introduction

University Students' Mental Health and the COVID-19 Pandemic

University students are often likely to experience serious mental health problems during their studies because they are exposed to several stressors including academic pressure, taking on more adult-like responsibilities, or having limited financial resources [1].

The COVID-19 pandemic has exacerbated students' mental health, as demonstrated by a skyrocketing incidence of mental health disorders during the repeated lockdowns between 2020 and 2021 [2,3]. Shifting to online courses, uncertainty about academic and professional future, and a dramatic reduction of social interactions have largely contributed to compromising the mental health of the student population [4]. Restrictive measures, in particular, were associated with high levels of depression, anxiety, and stress in students [5-7]. In 2020, nearly one-fifth of students experienced suicidal ideation as a result of COVID-19 [8]. The prevalence of these mental health problems was more than 50% higher among students than in the general population [9]. Students' mental distress due to the COVID-19 pandemic persisted in the aftermath of the peak of the pandemic [10].

Against this background, there has been a high demand for mental health prevention programs addressed to students during and after the lockdowns [11]. Several studies have described the development and application of interventions aimed at supporting students during the health crisis. In particular, digital psychological interventions have produced positive effects on students by promoting resilience and well-being [12]. During the lockdowns, when face-to-face contact was limited, digital interventions had the advantage of reaching a larger audience with no time or space limit. Examples of interventions were video clips, online booklets, mobile applications, virtual doctor appointments, etc. University students, in particular, actively sought this type of online help and interventions, probably because they are digital natives [13].

Gamification in Mental Health–Related Interventions

To optimize interventions, gamification is considered an essential strategy, including in the mental health field [14]. Gamification relies on the full involvement of the player and exploits several psychosocial determinants affecting the learning process (eg, self-efficacy, social interaction, and a positive learning environment) [15]. Enhancing these factors can facilitate the recall of abstract concepts, such as the concept of mental health [16]. Introducing gamified elements (eg, step-by-step sequencing, rewarding systems, and puzzle-solving activities) in a health promotion and prevention tool can influence the aforementioned psychosocial determinants and, consequently, stimulate participants' learning loop and their cognitive capacity [17].

Over the last few years, increasing attention has been paid to the possibility of games improving well-being [18]. Games can engage players, especially young target populations, in understanding and retaining information in a more attractive and acceptable way [19]. This means that they can increase their mental health literacy and their knowledge of both the symptoms of psychological problems and the different solutions to overcome them. By providing tips and skills to face psychological difficulties, games might also contribute to positive changes in individuals' behaviors and attitudes.

Escape Games as a Tool to Improve Psychological Well-Being and Prevent Mental Disorders

Escape games are a type of digital intervention based on gamification where players collaborate to find clues, complete tasks, and solve puzzles with the aim of achieving a specific, time-bound goal, which is usually to escape from a room. Previous research corroborated the constructive impact of escape games in improving health-related knowledge in players [20,21] using a learning-by-doing approach [21].

Escape games can contribute to delivering health-related messages by fostering motivation for behavioral change through an enjoyable and playful approach, according to the PRIME (Plan, Response, Impulses, Motives, and Evaluation) theory of motivation [22]. Based on this theory, a decision to engage in an activity will not result in action unless it generates the desire and the impulse to do it at the relevant moment. Thus, the stimuli generated from the act of playing a game—including different tasks, lights, sounds, and colors—trigger feelings, ideas, and brain activities for positive decision-making. In other words, the game gives the input to change.

Additionally, gamification has the potential to increase motivation, engagement, and self-awareness, and even diminish symptoms of diseases such as depression and anxiety [23]. Indeed, gamification stimulates several components of good mental health. As an example, achieving goals in a game can result in a sense of satisfaction, accomplishment, and increased self-esteem, all of which improve overall well-being. Furthermore, game enjoyment is associated with positive well-being and social and emotional support [24].

The Cocreation of Escape Games

Cocreation occurs when end users and service providers, often along with other participants, work together in the early phases of the development of an intervention cycle [25]. Cocreation is a process facilitating the acceptability of an intervention because it primarily considers the needs and preferences of end users during the intervention development. Thus, the benefits of gamification tend to increase when cocreation is used [26], and cocreating an escape game can foster its adoption [27]. Based on this approach, game producers and end users must first exchange views to achieve a shared goal [28,29]. Indeed, when developing a public health–related game, players' experience and needs are relevant for enhancing its effectiveness in promoting health and prevention. Ideally, players work alongside designers, health professionals, and researchers, to produce the intervention. The cooperation of players and other stakeholders is therefore essential to maximize end users' acceptability and adherence to the game. As a result, cocreation is usually recommended to produce a successful game, including in the mental health field. Accordingly, sensitive topics and taboos should be addressed using players' words and taking into account the levels of empathy and sympathy players display during the cocreation process. Including end users with lived experience of mental health disorders promotes a deeper understanding of the game topic [30].

The PRODUCES Framework

PRODUCES (Problem, Objective, Design, End Users, Cocreators, Evaluation, Scalability) is among the different existing frameworks facilitating the cocreation of health-related interventions [27]. It is well-known for using a systemic approach to participatory methodology. According to this framework, the *problem* is a narrowed-down behavioral issue that the researchers and the designers wish to address. The cocreation process has an *objective* (“what” and “how”) and follows a specific predefined *design*, engaging *cocreators* who represent end users (ie, a specified target population). For the latter, all characteristics must be considered, from age to socioeconomic status, to tailor the intervention coherently. Cocreated interventions can also be *evaluated* and their *scalability* can be assessed. The satisfaction of the end user as well as the effectiveness of the intervention are elements to consider for the final evaluation. A successful intervention can be scaled up to reach a wider public. Thus, the PRODUCES framework helps to guide the participatory methodology by providing specific instructions.

Objective

The objective of this study was to describe the process of cocreation of the escape game “EscapeCovid”. The end goal of the game was to promote university students’ mental health literacy, their beliefs about mental health, management of emotions, and positive coping strategies during the COVID-19 pandemic. An applied methodology is presented here to be used as a model for cocreating an acceptable gamified mental health intervention addressed to young people. Providing this example also has the aim to illustrate one cocreation process for the benefit of other researchers and designers.

Methods

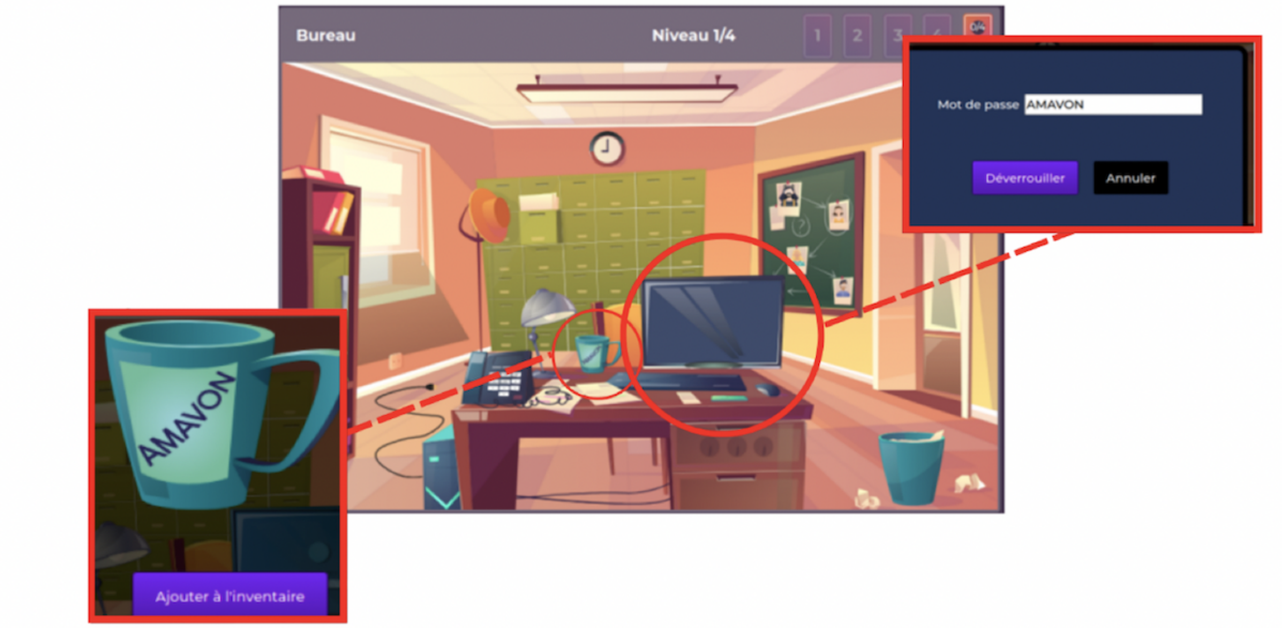
The First Pilot Version of “EscapeCovid”: The Escape Game “Manage Your Emotions”

“EscapeCovid” was based on an existing escape game that was used as the skeleton of the final game. “Manage Your Emotions” was created in 2021 during the pandemic by a start-up based in Bordeaux, France, specializing in producing both real-world and online escape games. Creators were game designers, programmers, and health care professionals aged under 35 years.

The game “Manage Your Emotions” is set in Tony’s room, a fictional university student living in a shared flat and experiencing the difficulties of the first lockdown. The goal of the game is to collect several tools to disclose emotion cards and combine them. The game session involves 4 players and a game guide. The role of the game guide is to coordinate the whole game, to give clues if the players are stuck, and to animate the debriefing session. The game lasts in total 2 hours: 45 minutes of play and 1 hour and 15 minutes of debriefing. During the debriefing, the game guide and the players discuss in more detail the concept of emotions (ie, how to identify and manage them). The game guide follows a predefined plot facilitating the interactions between participants.

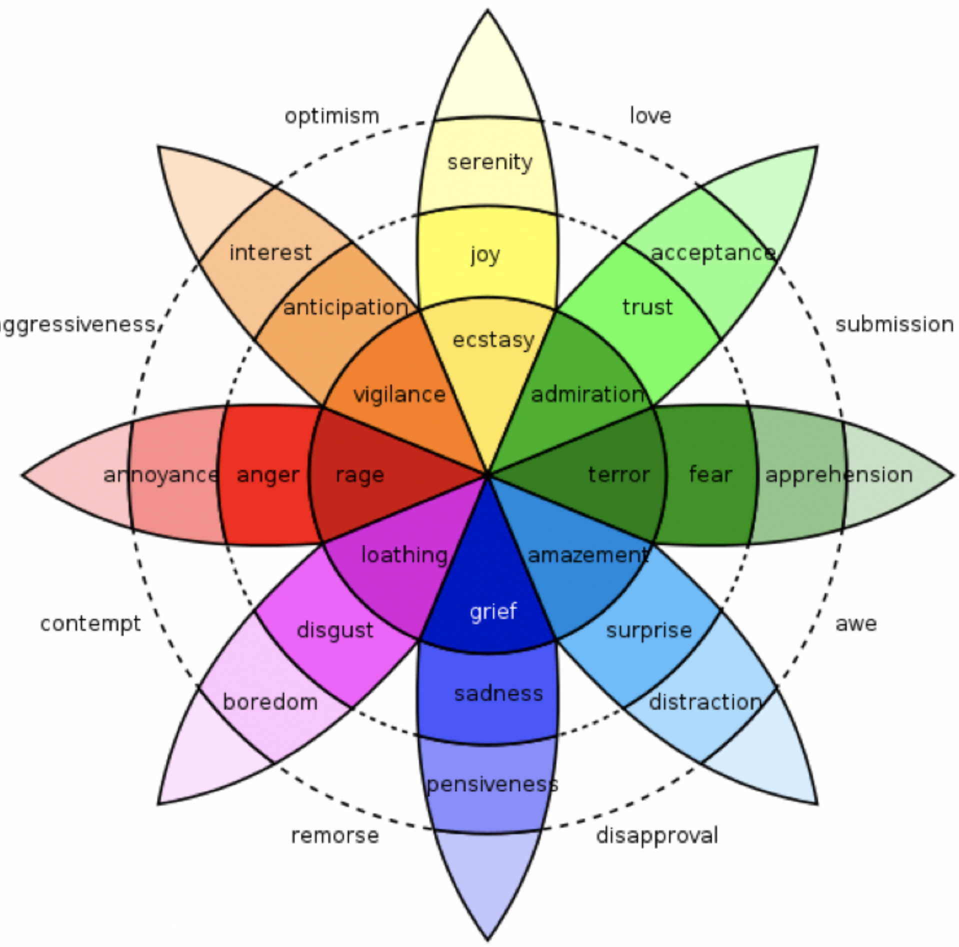
During the game, players follow Tony during a typical day through 3 rooms of his apartment (his home office in the bedroom, living room, and bedroom). By doing so, players discover his emotions and their consequences on his daily life. Figure 1 illustrates Tony’s home office in the bedroom with a set of clues for the players.

Figure 1. Tony's office in the escape game "Manage Your Emotions."



Players solve puzzles by clicking on the elements on the screen to uncover emotion cards. The definition of the different emotions is based on Plutchik’s wheel of emotions [31] (Figure 2) which is the theoretical framework of the game.

Figure 2. Plutchik's wheel of emotions.



Plutchik's wheel of emotions covers 8 primary emotions: joy, trust, fear, surprise, sadness, anticipation, anger, and disgust. They can be combined into more complex secondary emotions—for example, the combination of joy and trust can result in love. In the wheel, darker colors correspond to more intense emotions. All combinations and intensities are explained in the cards. Plutchik's theory posits that the more we know about emotions, the better we understand how various emotions are interlinked and how they can change over time. Plutchik's wheel of emotions has been used in several studies as a scientific instrument to interpret emotions [32,33]. In this escape game, playing cards had to be associated to identify Plutchik's emotions.

As the name suggests, the game "Manage Your Emotions" exclusively focused on emotions and therefore did not cover the full range of features of mental health (eg, mental health literacy and positive coping strategies). Furthermore, the software presented several bugs and the scenarios did not reflect the real-life conditions of a student during the pandemic.

Cocreating "EscapeCovid": Test and Optimization

Our cocreation process followed the PRODUCES framework [27]. The problem we chose was students' mental health during the COVID-19 pandemic. We particularly focused on mental health literacy, beliefs about mental health, management of emotions, and positive coping strategies as the levers to act upon for increasing students' psychological well-being. We addressed all types of mental health problems, but specifically anxiety and depression, among the most common troubles in young people [34]. These problems were exacerbated during the repeated COVID-19–related lockdowns [35].

Our objective was to develop the "EscapeCovid" game. The project was born during the third lockdown in France (from April 3 to May 3, 2021 [not included], ie, 29 days), where students were especially penalized because all educational institutions, but universities, were open. At that time, the plight of university students was prominently featured in the French media, which in turn heightened the pressure on French politicians [36].

As for the design aspect, we used a 2-step participatory methodology approach (test and optimization), as described below. Both steps involved students as players of the game before and during its improvement. Thus, through direct experimentation, cocreators were a sample of students representing all university students referred to as end users.

The evaluation was performed through questionnaires and semistructured interviews using a mixed methods approach. Students reported their opinions on the game allowing for an assessment of its qualities and defaults. In this sense, the design and the evaluation were strictly related.

In terms of scalability, our objective was to distribute the new game among additional universities catering to French-speaking students (eg, France, Africa, and Québec).

Table 1 reports the components of our study corresponding to the PRODUCES framework, including the phases and steps of the project.

Table 1. The PRODUCES[a] framework applied to the "EscapeCovid" study.

PRODUCES framework	Application in "EscapeCovid"	Corresponding element/phase
Problem	To address students' mental health during the COVID-19 pandemic	N/A[b]
Objective	To develop the "EscapeCovid" game	N/A
Design	Participatory methodology approach	Cocreation (test + optimization)
End users	Students	Optimization
Cocreators	45 health care students + 2 game guides + 1 developer + 1 project manager + 1 student intern + 1 designer+ 1 developer + 1 medical doctor + 1 researcher	Cocreation (test + optimization)
Evaluation	45 questionnaires and 10 semistructured interviews (mixed methods)	Test
Scalability	Disseminate the new game to other French students	Optimization

[a]PRODUCES: Problem, Objective, Design, End Users, Cocreators, Evaluation, Scalability.

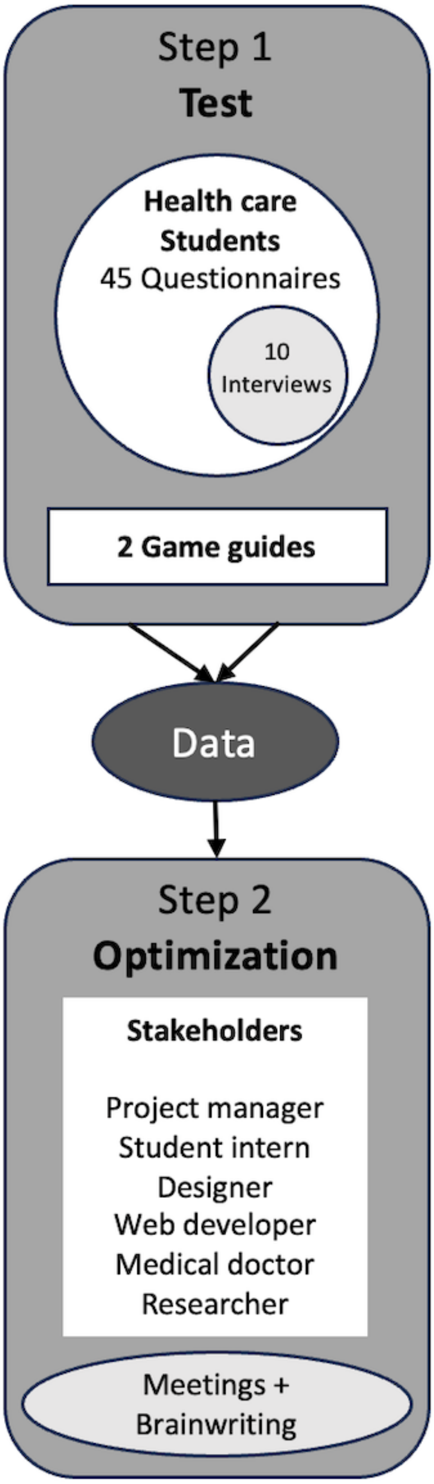
[b]N/A: not applicable.

In practice, referring more specifically to design, cocreation was implemented following 2 steps: the test and the optimization. In the first step, the escape game "Manage Your Emotions" was tested by a sample of health care students, 2 game guides (both public health experts), and 1 developer (a computer scientist). Game guides trained with colleagues and friends to annotate their first impressions of the game flow. They commented on the difficulties encountered while animating the game, including interactions with players and technical issues. The collected information was used to improve the debriefing and to solve bugs in collaboration with the developer.

Then, the official test was launched. Health care students played the game in groups of 3 or 4, each group supervised by 1 game guide. At the end of the game session, health care students were asked to complete an online questionnaire to rate their experience and provide feedback for improvement. Semistructured interviews were conducted with some voluntary respondents to obtain more in-depth advice (for improving the game).

In the second step, the game was optimized using the data collected from the test with the collaboration of a group of stakeholders. The latter included 1 project manager from a public health research center, 1 student intern completing a degree in cognitive engineering, 1 game designer, the developer of the test step specialized in computer science, 1 medical doctor, and 1 psychology researcher. Meetings were organized to reshape the pilot game and brainwriting techniques were used to collect ideas and pass them on to stakeholders. The brainwriting technique involves the written generation of ideas by different individuals on separate sheets, which are then collated by the project manager in the same shared file. Ideas are categorized and synthetized in the shared document where stakeholders can discuss them with written comments and paragraphs. Then, the team meets in person to agree on a common solution. Meetings and brainwriting were done in an iterative loop [37]. Figure 3 illustrates the cocreation process we adopted.

Figure 3. The cocreation process.



The Testers of "EscapeCovid"

The students testing the game were included exclusively if they were health care students registered at the University of Bordeaux (France). We accepted all specific health-related fields of study (eg, medicine, midwifery, speech therapy), genders, and ages. We opted for health care students to assess the relevance of the contents of the game given their expertise in medical and paramedical care. Furthermore, previous studies have shown that health care students are a population at risk of mental illness [38] and, as a consequence, tend to be more sensitive to this topic. Recruitment was conducted from April 27 to May 17, 2022. We used a snowball sampling approach starting with health care students doing their internship at the public health research center where the study was based. Additionally, health care–related student associations were asked to post a recruitment announcement on their social media pages. We aimed to recruit between 30 and 50 health care students, as this number would guarantee the feasibility of the study and the interactions required during the game. Thus, the recruitment was stopped when we reached a sample of 45 health care students and the recruitment strategies seemed to be no longer efficient (ie, no further responses). The final number of health care students who took part in the study was 45.

Through an email or by clicking on a link on the association's social media posts, participants in the study were directed to a form to schedule the game session and then randomly allocated to a session including 4 players each.

All 45 health care students received a €20 (US $22) gift card. Among them, 10 also volunteered to take part in a semistructured interview and received a supplementary €20 (US $22) gift card.

In addition to health care students, testers included 2 game guides and 1 developer employed by the start-up producing the game. No inclusion or exclusion criteria were considered for these testers who were all females and aged between 25 and 30 years.

The overall aim of the project was to coproduce a game and not to measure its impact on the mental health of health care students playing the game. Students were in charge of testing the intervention, as opposed to being on the receiving end. In other words, they were not the research sample but were actively engaged in designing and implementing the research process [39]. Nonetheless, they were provided with a list of mental health care services they could refer to if needed. The medical doctor and the researcher in psychology from the stakeholders' group were also available upon request. Finally, respondents were asked to electronically sign a consent form stating that their answers were completely anonymous without tracing. Interviewees also signed a form assuring that the recording of the interview would be deleted after 5 years until the final report and the last published paper, according to the policy of the involved research center.

Data Collection Instruments and Analysis During the Test

A mixed methods approach was applied using both questionnaires and semistructured interviews which were administered to our sample. The satisfaction questionnaire was sent by email to students 1 day after having played the game. It was created ad hoc by IM for this study and tested with 3 public health interns at the research center where the project was conducted. The interns played the escape game and answered the questions reporting to IM if they were adapted and appropriate, and whether it was easy to answer them. Some adjustments to the original items were made after this pilot testing. The final satisfaction questionnaire included 12 items on the degree of appreciation and relevance of the intervention. On a visual analog scale from 0 (not at all) to 10 (a lot), students had to rate the game in terms of how enjoyable it was, the quality of its content, its level of difficulty, the graphics, and the clarity and relevance of the objective. Students were also asked to state to whom they would recommend the game, whether they would pay to play it, whether they had understood the importance of talking about mental health, whether the game increased their knowledge about mental health, whether the game helped them speak more freely about mental health, and whether they felt the game destigmatized mental health. Participants were also asked to rate the game from 1 (very bad) to 5 (very good) stars. We included these specific questions because they provided concrete hints for improving the game. The start-up appraised the features with the lowest scores as the most important to consider when reshaping the game. For instance, they chose to work first and foremost on the graphics if players rated them low (ie, <5 points). Some questions helped understand if the game could have its own business model, with players advising and paying for it. Finally, the questions were aimed at assessing the impact the game had on students' mental health literacy, ranging from destigmatization to readiness to seek help [40]. The items of the satisfaction questionnaire are available in Multimedia Appendix 1. Sociodemographic characteristics were also collected, including students' gender, age, and year of study. Variables were described as counts and percentages. The questionnaire allowed us to obtain a large number of answers in a short time from a young population that is often difficult to reach [41].

Semistructured interviews were based on a grid composed of 3 macro themes and related 13 subthemes. The first macro theme, called "General Description of the Participant" included the following 3 subthemes: students' profile (sociodemographic characteristics), any previous experience with escape games, and the reason for participating in this study. The second macro theme was a "Brief Account of Participants' Experience" during the game session of "Manage Your Emotions," focusing on 5 subthemes: whether students enjoyed it, their satisfaction with the design and scenarios, the feasibility of the game, the learning outcomes, and any advice to improve the game. The last macro theme, "The Impact of the Game," included questions on the effectiveness of the game in teaching 5 topics (each corresponding to a specific subtheme): mental health, stigmatization, understanding and managing emotions, the importance of help-seeking, and techniques for mental health promotion. All subthemes were applied deductively, meaning they had been determined before the interviews.

Then, individual students' speeches were generalized to obtain an overall assessment of the game. Interviews were recorded, fully transcribed, and analyzed through qualitative coding. The framework method was used to cross-check results among individuals and within individuals to report common and consistent concepts [42]. This approach allowed us to list the guidelines for the optimization step.

Ethics Considerations

As the goal of the project was to collect satisfaction data with no repercussions on participants' health, no ethical approval was needed, in line with the French law for health-related research (Délibération n° 2018-155 du 3 mai 2018).

Results

Sociodemographic Characteristics

The quantitative sample of 45 students was purposely limited for a small-scale test. Among them, 34 were female students, 10 male, and 1 nonbinary. Their average age was 22 years (range 18-27 years). The years of study ranged from first-year students to PhD candidates, with the majority attending their fourth year (n=16). The qualitative sample (n=10) was composed of 7 female students and 3 male students. In the qualitative study, one-half of the sample declared having experienced a mental health problem and having seen a mental health specialist.

Students' Gaming Experience

The sample of 45 students who answered the questionnaire and the sample of 10 students interviewed reported enjoying the game session.

> *We discuss between us, why and how it is this emotion and not another [...] it was really good.* [B, female PhD candidate, Public Health, first year]
>
> *Yeah, I really liked the associations of emotions [...] frankly, we spoke with people we didn't know, so frankly it went well, it was cool.* [D, male students, Pharmacy, fourth year]

The majority of the sample (30/45, 67%) gave a high score (between 8 and 10) when asked whether they enjoyed the game. For 34/45 (76%) it was interesting (scores from 8 to 10). Twenty-one students considered that the game was easy. The most frequent overall score given to the game was 4 out of 5 stars (25/45, 56%).

In line with this finding, 21/45 (47%) respondents gave a positive score between 8 and 10 regarding the appeal of the graphics. Concerning the visual staging of the game, 1 student declared that the storyline was not coherent:

> *Tony's apartment is too big for being a student flat.* [MA, female student, Speech Therapy, second year]

Tony's character was also discussed, with some students questioning his relatability:

> *When Tony was talking, I didn't really get into the thing, in the end I found it very tricky, too tricky, a bit like a fake student.* [L, female student, Speech Therapy, third year]

Regarding the overall content of the game, 27/45 (60%) participants found the objective of the game clear (scores from 8 to 10) and 30/45 (67%) considered the content of the game suitable for students (scores from 8 to 10).

Half of the sample (24/45, 53%) would recommend the game to their close friends and family, especially their friends attending university (43/45, 96%). However, the vast majority of students (36/45, 80%) would not pay to play it.

The Knowledge Acquired During the Online Game Session

For 27/45 (60%) of the respondents, the game made them understand the importance of talking about mental health and 38/45 (84%) thought that the game was likely to increase their knowledge about mental health. However, interviewed students reported that knowledge about mental health was addressed in an unsuitable way.

For students, the game enabled users to better understand and identify different emotions, but the general concept of mental health was missing:

> *It was really more about identifying emotions, and self-reflection.* [L, female student, Speech Therapy, fourth year]
>
> *There would be a wealth of important information to address on mental health.* [A, female student, fourth year of international health]

The Development of the "EscapeCovid" Game

Quantitative and qualitative data from the test (step 1) informed the optimization of the game "Manage Your Emotions" (step 2) to produce the new game "EscapeCovid". Data were collated and analyzed by the stakeholders working on the development of the game. All results were considered to reshape the new escape game accordingly. The results of the mixed methods analyses were shared among stakeholders. This group of experts met 3 times to summarize the most important suggestions provided by the testers. Each meeting lasted from 3 to 4 hours. Then, 1 shared document was prepared and the stakeholders were asked to provide solutions for each suggested change. This was the beginning of the brainwriting process, where stakeholders updated the document once per week and met regularly every other week. Ideas were incorporated into a new working document, which was the basis for a new round of discussions (ie, five 2-hour meetings). Once consensus had been reached, the web developer revised the game following the guidelines written by the stakeholders on an online document and Figma (Figma, Inc.), a collaborative web application for interface design. Both the contents and the designs were discussed and modified.

Table 2 reports the modifications made from the first version of the game to the final one (also see Textboxes 1 and 2 for the topics addressed and educational content of versions 1 and 2).

Table 2. Comparison of features from the 2 versions of the game.

Features	Version 1: "Manage Your Emotions"	Version 2: "EscapeCovid"
Objectives	• To teach players to name, identify, and manage their emotions.	• To increase students' knowledge of mental health by familiarizing them with a range of emotions and symptoms of depression and anxiety.
Game flow	• 45-minute game session (3 rooms: home office in the bedroom, living room, and bedroom) + 45-minute debrief.	• Alternating game/debrief sessions in each room (home office in the bedroom, living room, and bedroom) and evaluation questions.
Topics addressed and educational content	• See Textbox 1.	• See Textbox 2.
Character(s)	• Tony, a confined student.	• Thomas, a confined student. • Thomas' roommate, "Hana." • A researcher who appears on the screen to give instructions and clues if the players need them.
Team competition	• Accumulation of points assigned according to the speed with which the player solves puzzles.	• Accumulation of points based on different criteria: speed in solving a puzzle, number of clicks used, time spent in each room and in the entire game, and correct answers to evaluation questions. Teams can also lose points if they choose to access clues to solve puzzles or if they answer assessment questions incorrectly.
Storyline	• Tony is a student living confined in his shared flat during the first lockdown. Game users follow him and the emotions he felt throughout lockdown. • Players must solve puzzles to access emotion cards.	• The same story plot as version 1. The presence of a new character changes the transition from 1 room to the other. • Players must solve puzzles to access emotion cards.
Game setting	• The graphics of the game are similar to an apartment of a young worker and not a student. • The vocabulary used by Tony is not adapted to the target audience.	• A student flat share; instead of having a home office in the bedroom and a separate 2-bedroom flat, the home office in the bedroom is on one side of Thomas' room. • Thomas' voice and vocabulary have also been adapted to meet the expectations/requests of the target audience.

Textbox 1. Topics addressed and educational content of version 1: "Manage Your Emotions."

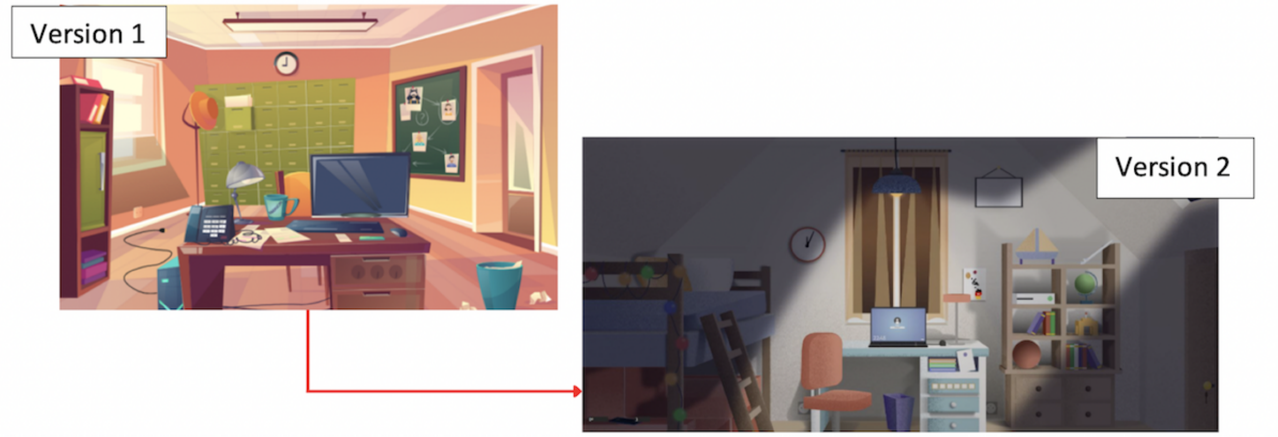
1. Game
   - This involved knowledge, identification, and management of emotions.
2. Debriefing Session
   - This involved tips and resources to identify and manage emotions.
   - In addition, the emotions felt during the examination period were discussed.

Textbox 2. Topics addressed and educational content of version 2: "EscapeCovid."

1. Game (examining symptoms of depression through emotions):
   - Stress and anxiety: other emotions, fear.
   - Anhedonia: sadness.
   - Self-devaluation: disgust.
   - Other topics related to mental health that were addressed, including the following:
     - Names of some mental health illnesses.
     - Stigmatization of mental health and mental illness.
     - Resources
2. Debriefing session
   - Room 1: Home office in the bedroom within the bedroom
     - Stress and anxiety: definitions and differences. Link with depression.
   - Room 2: Living room
     - Relationship between mental health and anhedonia. Link with depression.
   - Room 3: Bedroom
     - Relationship between mental health and self-devaluation. Association with depression.
3. Debriefing/end of the game
   - Emotions not previously addressed during the game.
   - Experienced symptoms and depression.
   - Stress on the importance of having good mental health.
   - Resources available in the case of a mental health problem.

In particular, substantial modifications concerned the introduction of mental health–related information in the game. Students had confirmed that the first version of the game was exclusively oriented toward emotions and their management. There were no specific elements of mental health described as either mental diseases or psychological well-being. In "EscapeCovid," the terms "depression" and "anxiety" were used by the main character. The symptoms and consequences of these mental health diseases were presented in the story. Depression and anxiety were selected because they are the most frequent psychological issues among young adults [43]. Players were supposed to learn more about mental health problems (mental health literacy), to destigmatize them (positive beliefs about mental health), and be able to tackle them (positive coping strategies). Puzzles and enigmas were used to teach these concepts with debriefing sessions to reinforce the learning process. Given their expertise in the field of health care, interviewees (health care students) helped with the writing of the plot, from the enigmas to the summary sheets. An example of a cocreated scenario within the story is available in Multimedia Appendix 2. Thanks to a cocreation approach, the game content could be revised by all public health stakeholders who were experts in public health and psychology. The graphics were also modified as shown in Figure 4.

Figure 4. Evolution of the graphics of the first room of "Manage Your Emotions" versus "EscapeCovid".



Playing the “EscapeCovid” Game

The escape game takes place in Thomas’ apartment which he shares with another student, Hana, during the first COVID-19 lockdown. Thomas is a university student and is taking his classes remotely. Throughout the game, we follow him during a typical day in lockdown. There are 3 rooms in Thomas’ apartment—a home office in the bedroom, a living room, and a bedroom. To move from 1 room to another, players must solve all the puzzles by clicking on the objects spread out in Thomas’ room. When players click on an object, a riddle appears and must be solved to move on to the following riddle. Players can only move to another room if they have solved all of the enigmas by uncovering clues or cracking the codes hidden in the sofa, among books, on the floor, etc. There is a limited number of clicks per participant.

At the end of each room scenario, a set of cards is shown, with each card containing a mental health–related message linked to the puzzles. For instance, in the living room, Hana is sleeping on the sofa in the dark and the books around her have titles containing the words “depression,” “pain,” etc. By solving clues and clicking, players can switch the light on and tidy up the room to make her recover strength. The cards synthesize the messages transmitted through the puzzles in the room. In this case, they explain the symptoms of depression and give tips for coping with distress.

“EscapeCovid” can be played in groups of 4-6 players who help each other and discuss using their computer cameras and headphones. This encourages team spirit and mutual aid, which can be the reflection of real life in the case of mental suffering. All along the game, the group of players is guided by a game guide who explains the rules and answers any questions. The same guide concludes the game session with a final debrief where all participants share their experiences. This final stage is essential for understanding and retaining the mental health–related takeaway messages.

Discussion

Principal Findings

We described the process we used to cocreate a digital game promoting students’ mental health during the COVID-19 pandemic. We followed a 2-step procedure. First, we collected quantitative and qualitative data from a manageable sample of students testing a preexisting game. Second, a group of stakeholders used these data to refine and optimize the game to obtain the final user-centered version.

The cocreation approach was very informative for developing the “EscapeCovid” game. In particular, during the test, students felt free to express their opinions openly and give feedback. They mostly appreciated the fact that they could support the development of an intervention addressed directly to them and their peers. Students were also motivated to cocreate the game because it was in line with their values. Students’ contribution to the design process nurtured new ideas following a collective creative approach [44] from the testing phase to the final optimization phase. Stakeholders’ work was facilitated by students’ guidelines while being creative and innovative.

The Rationale and the Usefulness of “EscapeCovid”

During the COVID-19 crisis, several mental health diseases emerged in the young, and digital games were among the most accepted solutions to overcome psychological difficulties [45]. With this rationale, we conceived "EscapeCovid". This game was designed to alleviate anxiety and depression by encouraging interaction with peers and fostering empathy. Participants in this study also confirmed the usefulness and appreciation of digital games during and after the COVID-19 crisis. Previous studies have shown that playing games is helpful in dealing with trauma and improves well-being [46]. This has also been observed in the context of the pandemic [47]. For this, "EscapeCovid" combines the pleasure and the entertainment of games, with a positive psychological effect. This might be due to teamwork, engagement, learning of coping strategies, and creativity, which are all at the root of our game. Indeed, the objective of "EscapeCovid" was to trigger the need to speak out about mental health after having experienced the psychological difficulties of COVID-19. "EscapeCovid" pioneered the discussion of mental health, making it a common topic, and provided advice on how to improve one’s mental health, especially in the aftermath of the crisis.

Guidelines for Successful Escape Games on Young’s People’s Mental Health

First, we confirmed that students enjoy playing escape games, which are linked to mental health. This was also found in other studies where health-related serious games were proven to facilitate experiential learning through an entertaining approach [21,48]. Thus, resorting to this type of intervention could be a good strategy to convey messages aimed at improving players' mental health. Engaging in playing games has been reported to promote the potential to enhance life satisfaction and improve individuals' mental well-being [18].

We observed that the plot was essential in capturing players' attention. During the game, testers were attracted by the messages and the scenario, feeling interested in following Thomas' story. They considered this aspect as crucial to transmit educational content, helping to convey new health-related topics, as shown in a previous study [21]. A meta-analysis on the gamification of learning confirmed that the use of personified narrative components is particularly effective in promoting behavioral learning [49].

Playing in groups was also a strategy to make connections and combat isolation, particularly experienced during the COVID-19 lockdown. The notion of interrelationship and mutual aid is a component to consider when developing games, even if they have a digital format.

According to testers, the "EscapeCovid" game had to be user-friendly, fun, and pedagogical. It had to present supplementary contents on mental health, with more specific details on mental health disorders and advice for preventing or treating them. We recommend that future game creators use precise and detailed content, providing accurate and uncensored mental health information and avoiding stigmatizing psychological disorders.

The Challenges of Cocreating Mental Health–Related Games With and For Young People

The involvement of end users entails a large proportion of subjectivity. This is especially true when handling topics such as mental health where feelings and emotions are at stake. End users give their opinions without any specific framework [50]. To mitigate this issue, the sample answering the questionnaire should be large enough to be representative of the target population. However, for the sake of feasibility in terms of time and financial efforts, it is not always possible to question more than 50-100 people. Qualitative interviews are meant to provide further information corroborating the quantitative data, but they still imply subjectivity. Per se, interviews cannot be representative [51]. The limited number of stakeholders has its share of arbitrariness. Nonetheless, regardless of their number, cocreators are the bridge between the whole target population and the stakeholders [27].

Cocreation is time-consuming. The 2-step development demands at least twice as long as the standard time to produce a game. Data collection and analysis add work to the producers who need to incorporate the results into their creative process. Discussions among stakeholders and brainwriting also slow down the production process. This is a limit of cocreation which cannot be overcome while being the best solution for producing an intervention that is well-tailored to the needs of the end user. Qualitative interviews in particular require time and effort, but they are a crucial tool for an in-depth analysis.

Technological issues should also be considered. Players' expectations might not be easily met because of software limitations. This could result in frustration from both parties and decrease adherence to the game. Start-ups and game industries should therefore keep up with new technologies and continuously update their services.

The Advantages of Cocreation

Cocreation has the advantage of considering the viewpoint of the end user, which might not be the case in classical processes of game development using a top–down approach. Collecting students' opinions before the development of the game allowed us to obtain several inputs and ideas that a limited number of web developers and project managers could not provide. The filter of expert health care professionals was also essential during the process. New knowledge was produced through sharing among parties.

The consultation with students having experienced mental health disorders allowed us to address the escape game's topic through a different lens. By considering their opinion, the game could be made more realistic and engaging. The disclosure of emotions and opinions can be facilitated through anonymous questionnaires and qualitative interviews, with students knowing that they are contributing to an intervention beneficial to them and their peers. The feeling of being useful to the community is another added value in the cocreation process [52].

Two different teams—one from an academic environment and the other from a start-up—collaborating to develop the game was also an asset. Indeed, researchers' scientific point of view informed the business goal of the start-up with the common will of creating an evidence-based marketable product.

Finally, cocreation provided useful information for the improvement of both the content and the format of the game. The latter was more contextually specific, adapted to a young population, namely, students, and bridging the gap between the preconceived ideas held by the start-up team and the real-world implementation of the game.

Recommendations for Researchers and Designers

The "EscapeCovid" game is an example of a digital game on mental health, which could be cocreated with young users. The guidelines we present might be applied to other similar interventions.

A 2-step approach is recommended with (1) an initial collection and analysis of combined quantitative and qualitative data, followed by (2) the integration of these data into the reflective and creative work of a group of experts and stakeholders. This approach, similar to a market survey, allows us to obtain clearer game instructions and broader insights, resulting in a more targeted and audience-specific final product.

We suggest basing the coconstruction process on an already existing pilot version because it facilitates the development of the final game. Although the game can be completely re-created, preliminary mock-ups will allow to save time and money. In our study, students were not required to design the game from scratch, and working on the first version of the game was an advantage for providing relevant, concrete, and realistic comments based on an existing version.

Stakeholders are also advised to take into account testers' opinions seriously and implement them accordingly. Testers represent the end users and their preferences must be carefully considered to obtain a fully satisfactory end product. For this reason, it is essential to collect as much information as possible during the testing stage.

We suggest to try out the game again once it has been modified. An iterative loop of test-optimization-test will increase the quality of the game. However, it must be pointed out that this process is expensive and time-consuming, despite being extremely informative. It is therefore recommended to end the loop once the comments are saturated, which effectively means limiting the number of additional changes suggested by users and resulting work for developers. This approach can serve as a blueprint for future work on creating gamified interventions on health-related topics addressed to students. Successfully cocreated games can have a wide outreach and improved scalability.

Study Limitations

Testers were mostly student interns at the research center where the study was conducted. This might have biased the results because participating students were already made aware of the project and willing to contribute to its progress as members of the same research laboratory. The test was performed by health care students, meaning that the modifications of the game might be relevant to them and not to students of other subjects. This particular student population also faces specific forms of stress not experienced by their peers. However, we considered their opinion to be of paramount importance in terms of the contents of the game, which benefited from their skills and experience with stress. The game was more realistic and other students could relate to Thomas' story imagined by young people their age.

Another limitation is that the students participating in the study were rewarded with a gift card, which could have significantly influenced the answers due to desirability bias. This phenomenon was even more likely in the case of interviewees who had received 2 gift cards. The gender balance among the participants was skewed in favor of female students, which may have influenced the results of the test.

Finally, we were not able to retest the game after its modification. Because of money and time constraints, we only produced a new version of the game without further refinement.

Conclusions

Our results suggest that cocreation contributes to improving the suitability of a health promotion and disease prevention intervention and that an end user–centered approach can facilitate intervention tailoring. When conceiving a health-related escape game, we recommend using a 2-step approach, including an initial collection of quantitative and qualitative data from end users testing the game (test), followed by the integration of these data into the development of the game by a restricted number of experts (optimization). This approach can serve as a model for future work on creating gamified interventions on health-related topics addressed to students.

Acknowledgments

The authors acknowledge and thank all the students who took part in the study. The authors are also indebted to the university student associations for helping recruit participants. Finally, the authors wish to thank Sara Garfield for revising, proofreading, and editing the paper. The project was funded by the French National Research Agency (ANR; grant number ANR-21-CE36-0011-02). IM was supported by the Digital Public Health Graduate Program within the framework of the PIA3 (Investment for the Future; Project reference: 17-EURE-0019). Further support was given by the AUF CO//ectif (DREO-7450).

Authors' Contributions

DL was responsible for conceptualization, investigation, and writing—reviewing. CV performed data curation and writing—original draft preparation. DT collected and analyzed the quantitative data. HH and CT took part in data curation and writing. IM contributed to conceptualization, methodology, investigation, writing—reviewing and editing, and supervision.

Conflicts of Interest

DL, HH, and DT are employees of the start-up Tricky, which developed the game described in this paper. Their jobs are paid by other projects, and they did not take any financial benefit from the development of "EscapeCovid." The game was developed for business, but data collection, analysis, and observation for this manuscript were conducted by CV, CT, and IM who are totally independent from Tricky. The study results are completely transparent and based on scientific integrity.

Multimedia Appendix 1
Satisfaction questionnaire.
[DOCX File , 15 KB - games_v12i1e48545_app1.docx]

Multimedia Appendix 2
Example of a cocreated scenario in "EscapeCovid".

[DOCX File , 329 KB - games_v12i1e48545_app2.docx]

References

1. Montagni I, Qchiqach S, Pereira E, Tully PJ, Tzourio C. Sex-specific associations between sleep and mental health in university students: a large cross-sectional study. J Am Coll Health 2020 Apr 07;68(3):278-285. [doi: 10.1080/07448481.2018.1546183] [Medline: 30615574]
2. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
3. Savage MJ, James R, Magistro D, Donaldson J, Healy LC, Nevill M, et al. Mental health and movement behaviour during the COVID-19 pandemic in UK university students: prospective cohort study. SSRN Journal 2020;19:100357. [doi: 10.2139/ssrn.3631268]
4. Arsandaux J, Montagni I, Macalli M, Texier N, Pouriel M, Germain R, et al. Mental health condition of college students compared to non-students during COVID-19 lockdown: the CONFINS study. BMJ Open 2021 Aug 19;11(8):e053231 [FREE Full text] [doi: 10.1136/bmjopen-2021-053231] [Medline: 34413111]
5. Vindegaard N, Benros ME. COVID-19 pandemic and mental health consequences: systematic review of the current evidence. Brain Behav Immun 2020 Oct;89:531-542 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.048] [Medline: 32485289]
6. Rossi R, Socci V, Talevi D, Mensi S, Niolu C, Pacitti F, et al. COVID-19 pandemic and lockdown measures impact on mental health among the general population in Italy. Front Psychiatry 2020 Aug 7;11:790 [FREE Full text] [doi: 10.3389/fpsyt.2020.00790] [Medline: 32848952]
7. Li Y, Zhao J, Ma Z, McReynolds LS, Lin D, Chen Z, et al. Mental health among college students during the COVID-19 pandemic in China: a 2-wave longitudinal survey. J Affect Disord 2021 Feb 15;281:597-604. [doi: 10.1016/j.jad.2020.11.109] [Medline: 33257043]
8. Jones L, Vereschagin M, Wang A, Munthali R, Pei J, Richardson C, et al. Suicidal ideation amongst university students during the COVID-19 pandemic: time trends and risk factors. Can J Psychiatry 2023 Jul;68(7):531-546 [FREE Full text] [doi: 10.1177/07067437221140375] [Medline: 36475311]
9. Macalli M, Texier N, Schück S, Côté SM, Tzourio C. A repeated cross-sectional analysis assessing mental health conditions of adults as per student status during key periods of the COVID-19 epidemic in France. Sci Rep 2021 Nov 09;11(1):21455 [FREE Full text] [doi: 10.1038/s41598-021-00471-8] [Medline: 34753945]
10. Pandya A, Lodha P. Mental health consequences of COVID-19 pandemic among college students and coping approaches adapted by higher education institutions: a scoping review. SSM Ment Health 2022 Dec;2:100122 [FREE Full text] [doi: 10.1016/j.ssmmh.2022.100122] [Medline: 35665095]
11. Theurel A, Witt A, Shankland R. Promoting university students' mental health through an online multicomponent intervention during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Aug 22;19(16):10442 [FREE Full text] [doi: 10.3390/ijerph191610442] [Medline: 36012078]
12. Riboldi I, Cavaleri D, Calabrese A, Capogrosso CA, Piacenti S, Bartoli F, et al. Digital mental health interventions for anxiety and depressive symptoms in university students during the COVID-19 pandemic: a systematic review of randomized controlled trials. Rev Psiquiatr Salud Ment 2023 Jan;16:47-58 [FREE Full text] [doi: 10.1016/j.rpsm.2022.04.005] [Medline: 35755491]
13. Montagni I, Tzourio C, Cousin T, Sagara JA, Bada-Alonzi J, Horgan A. Mental health-related digital use by university students: a systematic review. Telemed J E Health 2020 Feb 01;26(2):131-146. [doi: 10.1089/tmj.2018.0316] [Medline: 30888256]
14. Fleming T, Bavin L, Stasiak K, Hermansson-Webb E, Merry S, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. Front Psychiatry 2016;7:215. [doi: 10.3389/fpsyt.2016.00215] [Medline: 28119636]
15. Huotari K, Hamari J. A definition for gamification: anchoring gamification in the service marketing literature. Electron Markets 2016 Jan 15;27(1):21-31. [doi: 10.1007/s12525-015-0212-z]
16. DeSocio J, Stember L, Schrinsky J. Teaching children about mental health and illness: a school nurse health education program. The Journal of School Nursing 2006;22(2):81. [doi: 10.1622/1059-8405(2006)022[0081:tcamha]2.0.co;2]
17. Rodriguez-Ferrer JM, Manzano-León A, Cangas AJ, Aguilar-Parra JM. A web-based escape room to raise awareness about severe mental illness among university students: randomized controlled trial. JMIR Serious Games 2022 May 05;10(2):e34222 [FREE Full text] [doi: 10.2196/34222] [Medline: 35511232]
18. Jones CM, Scholes L, Johnson D, Katsikitis M, Carras MC. Gaming well: links between videogames and flourishing mental health. Front Psychol 2014 Mar 31;5:260 [FREE Full text] [doi: 10.3389/fpsyg.2014.00260] [Medline: 24744743]
19. Fleming T, Cheek C, Merry S, Thabrew H, Bridgman H, Stasiak K. Serious games for the treatment or prevention of depression: a systematic review. Revista de Psicopatologia y Psicologia Clinica 2014;19(3):227-242 [FREE Full text] [doi: 10.5944/rppc.vol.19.num.3.2014.13904]
20. Edwards T, Boothby J, Succheralli L. Escape room: using an innovative teaching strategy for nursing students enrolled in a maternity clinical course. Teaching and Learning in Nursing 2019 Oct;14(4):251-253. [doi: 10.1016/j.teln.2019.05.001]

21. Eukel HN, Frenzel JE, Cernusca D. Educational gaming for pharmacy students - design and evaluation of a diabetes-themed escape room. Am J Pharm Educ 2017 Sep;81(7):6265 [FREE Full text] [doi: 10.5688/ajpe8176265] [Medline: 29109566]
22. West R, Michie S. A brief introduction to the COM-B Model of behaviour and the PRIME Theory of motivation. Qeios 2020 Apr 07:1-6 [FREE Full text] [doi: 10.32388/ww04e6]
23. Cheng VWS. Recommendations for implementing gamification for mental health and wellbeing. Front Psychol 2020 Dec 7;11:586379 [FREE Full text] [doi: 10.3389/fpsyg.2020.586379] [Medline: 33365001]
24. Johannes N, Vuorre M, Przybylski AK. Video game play is positively correlated with well-being. R Soc Open Sci 2021 Feb 17;8(2):202049 [FREE Full text] [doi: 10.1098/rsos.202049] [Medline: 33972879]
25. Torfing J, Sørensen E, Røiseland A. Transforming the public sector into an arena for co-creation: barriers, drivers, benefits, and ways forward. Administration & Society 2016 Nov 28;51(5):795-825. [doi: 10.1177/0095399716680057]
26. Lopes JM, Gomes S, Rodrigues I. Playing the gamification and co-creation game: a bibliometric literature review. Manag Rev Q 2022 Nov 08;74(1):353-376. [doi: 10.1007/s11301-022-00302-8]
27. Leask CF, Sandlund M, Skelton DA, Altenburg TM, Cardon G, Chinapaw MJM, GrandStand, Safe StepTeenage Girls on the Move Research Groups. Framework, principles and recommendations for utilising participatory methodologies in the co-creation and evaluation of public health interventions. Res Involv Engagem 2019 Jan 9;5(1):2 [FREE Full text] [doi: 10.1186/s40900-018-0136-9] [Medline: 30652027]
28. Ind N, Coates N. The meanings of co-creation. European Business Review 2013;25(1):e1 [FREE Full text] [doi: 10.1108/09555341311287754]
29. OHern M, Rindfleisch A. Customer co-creation: a typology and research agenda. Review of Marketing Research 2010 Nov;6:84-106 [FREE Full text] [doi: 10.1108/s1548-6435(2009)0000006008]
30. Hanlon CA, McIlroy D, Poole H, Chopra J, Saini P. Evaluating the role and effectiveness of co-produced community-based mental health interventions that aim to reduce suicide among adults: a systematic review. Health Expect 2023 Feb 14;26(1):64-86 [FREE Full text] [doi: 10.1111/hex.13661] [Medline: 36377305]
31. Plutchik R. The Emotions: Facts, Theory and a New Model. New York, NY: Random House; 1962.
32. Arroyo-Barrigüete JL. Sentiment analysis of Lovecraft's fiction writings. Heliyon 2023 Jan;9(1):e12673 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e12673] [Medline: 36699284]
33. Kaklauskas A, Abraham A, Ubarte I, Kliukas R, Luksaite V, Binkyte-Veliene A, et al. A review of AI cloud and edge sensors, methods, and applications for the recognition of emotional, affective and physiological states. Sensors (Basel) 2022 Oct 14;22(20):7824 [FREE Full text] [doi: 10.3390/s22207824] [Medline: 36298176]
34. Pedrelli P, Nyer M, Yeung A, Zulauf C, Wilens T. College students: mental health problems and treatment considerations. Acad Psychiatry 2015 Oct 21;39(5):503-511 [FREE Full text] [doi: 10.1007/s40596-014-0205-9] [Medline: 25142250]
35. Arsandaux J, Montagni I, Macalli M, Texier N, Pouriel M, Germain R, et al. Mental health condition of college students compared to non-students during COVID-19 lockdown: the CONFINS study. BMJ Open 2021 Aug 19;11(8):e053231 [FREE Full text] [doi: 10.1136/bmjopen-2021-053231] [Medline: 34413111]
36. Trespeuch M, Tenret É. La crise sanitaire: une prise de conscience de la condition étudiante? Revue Française de Socio-Économie 2021;26(1):5-11 [FREE Full text] [doi: 10.3917/rfse.026.0005]
37. VanGundy A. Brainwriting for new product ideas: an alternative to brainstorming. J Consum Mark 1984;1(2):67-74. [doi: 10.1016/b978-0-12-407157-5.00002-6]
38. Quek, Tam, Tran, Zhang, Zhang, Ho, Ho. The global prevalence of anxiety among medical students: a meta-analysis. Int J Environ Res Public Health 2019 Jul 31;16(15):2735 [FREE Full text] [doi: 10.3390/ijerph16152735] [Medline: 31370266]
39. Goodyear-Smith F, Jackson C, Greenhalgh T. Co-design and implementation research: challenges and solutions for ethics committees. BMC Med Ethics 2015 Nov 16;16(1):78 [FREE Full text] [doi: 10.1186/s12910-015-0072-2] [Medline: 26573410]
40. Jorm AF. Mental health literacy. Public knowledge and beliefs about mental disorders. Br J Psychiatry 2000 Nov 02;177(5):396-401. [doi: 10.1192/bjp.177.5.396] [Medline: 11059991]
41. Lallukka T, Pietiläinen O, Jäppinen S, Laaksonen M, Lahti J, Rahkonen O. Factors associated with health survey response among young employees: a register-based study using online, mailed and telephone interview data collection methods. BMC Public Health 2020 Feb 05;20(1):184 [FREE Full text] [doi: 10.1186/s12889-020-8241-8] [Medline: 32024488]
42. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
43. Kumaraswamy N. Academic Stress, Anxiety and Depression among College Students-A Brief Review. International Review of Social Sciences and Humanities 2013;5(1):135-143.
44. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
45. Marston HR, Kowert R. What role can videogames play in the COVID-19 pandemic? Emerald Open Res 2020 Oct 5;2:34 [FREE Full text] [doi: 10.35241/emeraldopenres.13727.2]
46. Colder Carras M, Kalbarczyk A, Wells K, Banks J, Kowert R, Gillespie C, et al. Connection, meaning, and distraction: a qualitative study of video game play and mental health recovery in veterans treated for mental and/or behavioral health

problems. Soc Sci Med 2018 Nov;216:124-132 [FREE Full text] [doi: 10.1016/j.socscimed.2018.08.044] [Medline: 30257787]
47. Barr M, Copeland-Stewart A. Playing video games during the COVID-19 pandemic and effects on players' well-being. Games and Culture 2021 May 06;17(1):122-139. [doi: 10.1177/15554120211017036]
48. Barber C, Stavroulaki K, Santanello C. Examining student motivation to use a gamified system in an immunology and immunization training course. Innov Pharm 2020 Nov 18;11(4):14 [FREE Full text] [doi: 10.24926/iip.v11i4.3328] [Medline: 34007662]
49. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev 2019 Aug 15;32(1):77-112. [doi: 10.1007/s10648-019-09498-w]
50. Finn J. The promise of participatory research. J Progress Hum Serv 2008;5(2):25-42.
51. Queirós A, Faria D, Almeida F. Strengths and limitations of qualitative and quantitative research methods. European Journal of Education Studies 2017;3(9):369-387 [FREE Full text] [doi: 10.5281/zenodo.887089]
52. Lee HM, van Dolen W. Creative participation: Collective sentiment in online co-creation communities. Information & Management 2015 Dec;52(8):951-964. [doi: 10.1016/j.im.2015.07.002]

Abbreviations

PRIME: Plan, Response, Impulses, Motives, and Evaluation
PRODUCES: Problem, Objective, Design, End Users, Cocreators, Evaluation, Scalability

Edited by T Rashid Soron; submitted 28.04.23; peer-reviewed by C Cederved, A Gavarkovs; comments to author 11.06.23; revised version received 16.07.23; accepted 09.02.24; published 18.03.24.

Please cite as:
Labrosse D, Vié C, Hajjam H, Tisseron C, Thellier D, Montagni I
An Escape Game on University Students' Mental Health During the COVID-19 Pandemic: Cocreation Study
JMIR Serious Games 2024;12:e48545
URL: https://games.jmir.org/2024/1/e48545
doi:10.2196/48545
PMID:38498033

Original Paper

Enhancing Serious Game Design: Expert-Reviewed, Stakeholder-Centered Framework

Lance Bunt[1], PhD; Japie Greeff[2], PhD; Estelle Taylor[1], PhD

[1]Unit for Data Science and Computing, North-West University, Vanderbijlpark, South Africa

[2]Optentia, North-West University, Vanderbijlpark, South Africa

Corresponding Author:
Lance Bunt, PhD
Unit for Data Science and Computing
North-West University
1174 Hendrick Van Eck Boulevard
Vanderbijlpark, 1900
South Africa
Phone: 27 720880188
Email: Lance.Bunt@nwu.ac.za

Abstract

Background: Traditional serious game design methods often overlook stakeholder needs. This study integrates stakeholder theory and enterprise architecture (EA), along with the Architecture Development Method, to propose a novel framework for serious game design. Crafted to aid practitioners, researchers, and specialists in leveraging resources more effectively, the framework is validated through a design science research methodology. Expert reviews have further refined its features, making it a robust tool for enhancing serious game design and implementation.

Objective: This paper introduces a framework for designing serious games, covering stakeholder analysis, requirements gathering, and design implementation planning. It highlights the importance of expert review in validating and refining the framework, ensuring its effectiveness and reliability for use in serious game design. Through critical assessment by experts, the framework is optimized for practical application by practitioners, researchers, and specialists in the field, ensuring its utility in enhancing serious game development. The next step will be to validate the framework empirically by applying it to a serious game development project.

Methods: We developed and validated a conceptual framework for serious game design by synthesizing stakeholder theory and EA through literature review, concept mapping, and theory development by way of a design science research approach. The framework is iteratively refined and validated via expert review, drawing on insights from professionals experienced in serious games, stakeholder theory, and EA. This method ensures the framework's practical relevance and effectiveness in addressing real-world design challenges.

Results: An expert review by 29 serious game practitioners validated the framework's success in stakeholder management, confirming its stakeholder-centered effectiveness. Although the experts praised its structured approach, they suggested clearer guidance for game design elements. In addition, the experts, while acknowledging the framework's complexity, saw its depth as valuable for efficient management. The consensus calls for a refined balance between detailed functionality and user-friendly design, with the framework's impact on stakeholder capabilities revealing a spectrum of professional needs.

Conclusions: This paper presents a framework for creating effective and organizationally aligned serious games. Evaluated across execution, practical, and EA levels, it is logical but varies in ease of understanding, with experts calling for more accessibility at the EA level. It enhances stakeholder efficiency and management but is criticized for rigidity and a need for flexibility. Recommendations include streamlining the framework, enhancing clarity, reducing administrative tasks, and incorporating clear guidelines on technology use, motivational elements, and operational tools. This aims to help stakeholders produce more targeted and adaptable game designs. The next iteration will be developed after application to a project and team feedback.

(JMIR Serious Games 2024;12:e48099) doi:10.2196/48099

KEYWORDS

serious games; stakeholder theory; enterprise architecture; serious game development; design framework

Introduction

Background

This paper articulates a stakeholder-centered framework for serious game design in various stages before, during, and after various methods have been applied to develop it. The research presents the framework—informed by stakeholder theory and enterprise architecture (EA)—as it evolved through various stages of development. It underscores the need for a structured approach that focuses on stakeholders throughout the design process to enhance the success of serious game production efforts. This research reflects 3 cycles of the design science research (DSR) paradigm, aiming to balance domain-specific needs with generalizable solutions. The framework emphasizes alignment with stakeholder needs and effective communication among different groups, recognizing the complexity of serious games and the importance of their relevance to users.

This work also presents findings from an expert review questionnaire that used a qualitative methodology to assess the framework's effectiveness. The review gathers feedback from practitioners and specialists in the field, guiding enhancements to the framework's clarity, structure, and usability. The paper concludes with insights into the framework's current state of development and recommendations for its refinement, emphasizing the need to simplify its complexity and communicate its components more effectively for real-world application.

Framework Requirements

Several key requirements for the initial framework emerged from a comprehensive integrative literature review. A conceptual framework is indispensable for comprehending the complex phenomena explored because it provides a structured and systematic method for organizing, analyzing, and interpreting data. Several essential components of a conceptual framework contribute to its efficacy, including clarity, relevance, coherence, simplicity, testability, and generalizability. As such, the following characteristics of the framework that has been developed serve as guiding principles. First, a conceptual framework should be clear and straightforward for its intended audience to comprehend. It should define its key concepts, variables, and relationships as well as provide a thorough overview of the subject under investigation. Second, the framework should be pertinent to the research problem or question being addressed. It must thus be tailored to the context and objective of the study. Third, the framework must have logical coherence, a clear structure, and internal consistency. The framework's concepts, assumptions, and relationships should be logically connected and consistent with one another. Fourth, a good conceptual framework should be as simple as possible while capturing the essential characteristics of the studied phenomenon. Fifth, the framework must be susceptible to empirical testing, with testable hypotheses and predictions that can be evaluated through observation and data analysis. Sixth and last, such a framework should be applicable to other settings or situations and be generalizable beyond the specific context of the study. It should therefore serve as a foundation for the development of broader theoretical insights and generalizations about the subject of study.

Our framework prioritizes stakeholder engagement and management within serious game design, addressing a gap often overlooked in conventional design literature. While incorporating established design elements—such as learning objectives [1-4], game mechanics [5-9], narrative [10,11], user interface and experience [12], and evaluation [13,14]—the framework's novelty lies in its stakeholder-centered approach. It is tailored to align with educational or training standards, drive engagement, and provide meaningful feedback. However, the primary focus is not solely on game design; instead, the framework is rooted in stakeholder theory and EA, which have been the fulcrum of our extensive literature review. By doing so, we address the intricacies of organizational and stakeholder dynamics, ensuring that serious games are developed within a context that appreciates the diverse roles and impacts of various stakeholders.

Products of the Integrative Literature Review

Overview

The integrative literature review presents a taxonomy of serious games; the phases of serious game production; the stakeholders involved in serious game production; stakeholder identification, analysis, and management procedures; and The Open Group Architecture Framework (TOGAF) Architecture Development Method (ADM). These concepts are briefly outlined in the following subsections because they inform the construction of the conceptual serious game framework.

What? Classifying Serious Games

Serious games are edifying artifacts, tools, and games created by development teams that use ludic activity for a specific purpose, format, genre, interaction style, and application area. Serious game taxonomies classify games by *purpose*. This classification can help identify the functions of a serious game and guide the selection and development of educational and training games. The following are some serious game categories:

- Simulation games simulate real-world situations to give learners practical experience and practice in complex or high-risk situations (eg, flight simulators for pilot training and medical simulators for surgery [15]).
- Educational games teach specific knowledge or skills, such as language, math, or history. Game mechanics such as rewards and feedback encourage learning and participation [16].
- Training games teach practical abilities such as customer service, leadership, and teamwork. To track progress and facilitate learning, they may include simulations or role-playing scenarios as well as feedback and assessment [17].
- Health games promote healthy behaviors such as exercise, healthy eating, and disease management. Game mechanics such as rewards and challenges may encourage behavior modification and participation [18].
- Persuasive games aim to influence players to adopt certain behaviors, such as environmental conservation, social

justice, or political activism. Story elements often emotionally engage and motivate players [19].

By classifying serious games by their intended function, a functional taxonomy can help find the best games for learning or training needs and guide their development. These serious game classifications are relevant to this research and accepted under the serious game banner.

Serious games, gamification, and game-based learning differ greatly. These terms all refer to using games or game elements in learning or training, but serious games are games with a specific goal. They usually teach players a skill [20]. Serious games differ from entertainment-focused commercial games. Gamification, by contrast, uses game elements such as points, badges, and leader boards to motivate and engage nongame users [21]. A fitness app, such as *Strava*, that rewards users for reaching fitness goals uses gamification. Game-based learning uses games to teach or train, but, unlike serious games, its main goal is not to achieve a learning outcome [22]. *Civilization* may be used by a history teacher to teach about historical events and civilizations, but it is not meant to teach history. Game-based learning uses games to teach or train without a specific goal.

Who? Serious Game Production Stakeholders

Different stakeholders from varying fields are involved in serious game production and development. Common stakeholders include the following [16]:

- Game developers are responsible for designing and developing the game. They create captivating game mechanics, visuals, audio, and more.
- Subject matter experts (SMEs) are knowledgeable about the serious game topic. They provide content and knowledge for game accuracy and effectiveness.
- Teachers and trainers can use serious games as a teaching or training tool for students or employees. They demonstrate how the serious game can meet learning goals.
- Game publishers distribute serious games to a wider audience. They market, distribute, and sell the game.
- Players are serious game consumers. They play the game and give feedback to improve it.
- Funders and sponsors are individuals or organizations that provide financial support for the development of the serious game (eg, government agencies, private foundations, or businesses).

When? Phases of Serious Game Production

Serious game production follows a similar process to traditional game development but with educational or training goals. The following are the five main serious game development phases:

1. The first step in serious game development is to identify the game's learning objectives, audience, and context. This usually involves a needs assessment or curriculum analysis to identify gaps or areas where game-based learning could be beneficial [15].
2. A detailed plan for game mechanics, story, user interface, and learning content is created during the design phase. This phase may include storyboarding, prototyping, and playtesting to ensure that the game is engaging and meets learning objectives [16].
3. Development includes creating the serious game, including programming, artwork, audio, and multimedia assets. Game developers, instructional designers, SMEs, and other stakeholders work together to ensure that the game meets educational or training goals [23].
4. After creating the serious game, it is tested and evaluated to ensure that it meets its learning objectives. User testing, focus groups, and other student and stakeholder feedback methods may be used [24].
5. The serious game can be deployed for use in educational and training settings in the deployment and maintenance phase and should be supported and updated to ensure that it remains relevant and effective over time.

Serious game developers use the aforementioned steps to create outcome-aligned educational and training games. Such a framework supports agile game development and various development methods.

Where? TOGAF ADM

Serious games can make use of TOGAF ADM. Serious games with multiple stakeholders and complex systems require a structured EA development process. TOGAF ADM can also be customized for different industries and organizations. Such a broad view ensures that the game aligns with organizational goals, making it a good choice for serious game creators. TOGAF ADM is a nine-phase, sequential process for EA [25]:

1. Architecture vision: the EA team creates a high-level vision of the organization's desired future architecture state. This phase determines the architecture development's business drivers, stakeholders, and scope.
2. Business architecture: phase 2 involves understanding the organization's business processes, objectives, and strategies. This phase produces business architecture artifacts that describe the organization's business capabilities, value streams, and structure.
3. Information systems architecture: this phase focuses on understanding the organization's information systems and technology infrastructure. This phase creates architecture artifacts for the organization's application, data, and technology architecture.
4. Technology architecture: phase 4 focuses on selecting and defining technology components for implementing the organization's architecture. This phase creates architecture artifacts for the organization's technology infrastructure's hardware, software, and networks.
5. Opportunities and solutions: phase 5 evaluates architecture solutions that meet business goals and objectives. This phase creates architecture artifacts that describe proposed solutions and their organizational impact.
6. Migration planning: in this phase, a plan is created to transition the organization's architecture to its desired future state. This phase produces architecture artifacts that describe transition activity sequence and timing.
7. Implementation governance: phase 7 oversees the implementation of architecture solutions and ensures alignment with organizational objectives and goals.

Governance framework architecture artifacts are produced in this phase.

8. Architecture change management: this phase manages ongoing changes to the organization's architecture, aligning them with its goals and objectives. This phase produces architecture artifacts that describe change management.
9. Architecture evaluation: this phase evaluates the effectiveness of the architecture solutions and ensures that they meet the organization's goals and objectives. This phase creates architecture artifacts that describe the evaluation process and results.

Interest groups can create EA solutions that meet their business goals by following TOGAF ADM. Moreover, TOGAF application to serious game development requires several crucial steps. First, an architecture vision is created to describe the game's goals, objectives, target audience, and learning outcomes. Stakeholder analysis is then performed to identify the game development stakeholders and their needs and expectations. This ensures that the game is designed with stakeholders in mind. Architecture requirements describe the game's functional and nonfunctional needs. The architecture development phase designs the game's architecture based on the previous step's requirements. This includes game mechanics, visual and audio assets, user interface, and layout design, that is, build, code, integrate visual and audio assets, and test the game's usability and efficacy. To ensure stakeholder satisfaction, the game is monitored and evaluated over time. This may involve player feedback, game performance data analysis, and adjustments. Specifically, this work references the application of an ADM to serious game development and highlights how it can assist organizations in creating the desired strategic resource game. Doing so emphasizes that the ADM not only assists in the development of the serious game but also identifies organizational capabilities, methods, and processes that can be leveraged in future projects, thereby enhancing the team's effectiveness.

Why? Stakeholder Identification, Analysis, and Management

Serious game design requires a stakeholder-centered conceptual framework for the following reasons:

- A stakeholder-centered approach considers various stakeholders' needs and expectations during design. This may lead to more effective, engaging, and audience-relevant games (Bopp, J, A, unpublished data, December 2020).
- A stakeholder-centered approach ensures that the game is designed for the end user, improving usability and effectiveness. This improves player engagement, learning, and game performance [26].
- A stakeholder-centered approach can involve stakeholders in the design process, facilitating participation and acceptance. By increasing stakeholder confidence and ownership, serious game adoption and implementation can succeed [27].
- To improve serious game sustainability and scalability, a stakeholder-centered approach can be used to design games that meet the evolving needs of stakeholders. This can help the serious game stay relevant and effective as stakeholders' needs change [28].

Serious game development relies on stakeholder identification, analysis, and management. This process begins with stakeholder identification. Stakeholder analysis prioritizes their needs and interests, while surveys, interviews, and focus groups help understand them. Stakeholder management involves planning how stakeholders will be engaged, their needs met, and their feedback incorporated into the serious game. Serious game development can use stakeholder management techniques such as regular meetings, an engagement plan, a registry, the prioritization of needs, feedback, and data analytics. Developers can create more effective, engaging, and audience-relevant games by managing stakeholder needs and expectations. A stakeholder-centered framework is needed for serious game design to ensure that stakeholders' needs are met and to improve game effectiveness, usability, and sustainability.

How? Stakeholder Identification, Analysis, and Management

Stakeholder identification, analysis, and management are crucial to project success, including serious game development. Stakeholders are people or groups who care about the project's outcomes and can influence them. Successful stakeholder identification, analysis, and management follow these four steps:

1. Identify internal, external, primary, and secondary stakeholders. Stakeholder analysis maps stakeholders and identifies key players [29].
2. Analyze stakeholders' interests, needs, expectations, and influence on project outcomes. A matrix that maps stakeholders by power and interest can do this [30].
3. The project team can develop strategies to manage stakeholder relationships based on stakeholder analysis. This involves prioritizing stakeholders by influence and interest and creating stakeholder-specific engagement strategies [31].
4. Implement stakeholder management strategies through ongoing communication and engagement, such as project updates, meetings, and consultations. Stakeholder interests must be monitored and the stakeholder management plan adjusted [32].

These steps for identifying, analyzing, and managing stakeholders can help serious game developers maximize project success and build long-term relationships. When a new serious game project begins, the organization or team will already know the relevant stakeholders from stakeholder management and EA. Thus, each project improves the organization or team.

Methods

Framework Development

Overview

This section discusses how the integrative literature review revealed relevant theories, determined its limits, found relevant sources, collected terminology, defined its theoretical pillars, and provided practical approaches to the stakeholder-centered framework. Moreover, this section assesses the framework's evolution over time; and it also theorizes future representations; reviews design processes; suggests improvements; and states the artifact design's aggregate, iterative, and consistent impacts.

Variant 1: Informed by Literature

The preliminary snapshot of the stakeholder-centered framework is a compilation of ideas for a flexible, general-purpose framework to aid in the design of serious games. Initial concepts included generating, developing, and visually communicating the system's fundamental elements, with a focus on user needs and empathy for the target demographic of serious game design stakeholders. Existing serious game literature, models, and approaches inform the framework variant, and an early exploration of these works provides a knowledge base for further consideration. Understanding the methods used in previous research on the same or similar issues assists in determining which methods will be most beneficial to advancing the topic and can aid in the evaluation of prior studies.

Various sources are represented in this formation because serious games' content, definition, sources, liminal works, methods, and existing frameworks are investigated. As such, a substantial portion of this work is theoretical in nature and largely represents the efforts to seek and collect literature on the nature of serious games.

Textbox 1 shows how theoretical and experiential exploration shaped our initial project impression. First, because serious game projects require people and management, stakeholder theory was added to the framework. Second, early EA readings may help organizations achieve their goals. The framework's third pillar, serious game design theory, positions the research and establishes its context. From this early stage, the framework must be applied and evaluated to determine its value for practitioners in the given milieu. This variant was extensively developer (self) reviewed. These steps close the DSR cycle loop and indicate that each variation is evaluated, even if reflectively.

Textbox 1. Sources that informed the first variant of the framework.

Sources and detail with explanation

- Deterding et al [21]
  - The authors define "gamefulness" and "gamification." This influential work examines the differences between full-fledged games, serious games, pervasive games, extending games, game elements, and playful interaction. Even if not adopted, their definitions of "gamification" and "gamefulness" in contrast to serious games and playful interaction refine discourse and enable researchers to better understand and analyze the phenomena.
- Annetta [33]
  - The author has presented a nested model of educational game design elements. Serious games have 6 elements, ranging from identity to instruction. This paradigm is hierarchical, with identity as the foundation for serious game design.
- Garris et al [23]
  - A model by the authors shows the learning approach used in educational game research and its results. First, the main goal of any instructional content is to create a game-like educational program. Second, these qualities trigger a loop of user perceptions or responses such as interest or delight, user behaviors such as perseverance or concentration, and system input. If designers can match educational content with game elements, this cycle creates repeatable and self-motivated play. Third and last, game participation achieves training goals and learning outcomes.
- Ferdig [34]
  - The authors define the "heart of serious game design" as theory, content, and game design. Serious game success requires emergent theory, content, and game design knowledge. Managing disciplinary conflicts and agreeing on serious game design is a major challenge for serious game teams.
- Marne et al [35]
  - The authors list 6 serious game design aspects. This serious game design methodology shows the importance and distinction of pedagogical and game design expertise and their role in serious game development. This model's main benefit is selecting the right experts for each design area.
- Rooney [36]
  - The author proposes a triadic serious game design framework that considers pedagogy, play, and fidelity to create media.
- Vanden Abeele et al [37]
  - The authors advocate the player-centered, iterative, interdisciplinary, and integrated (P-III) serious game design framework. This prominent framework provides a way for creating serious games that hinges on 4 conceptual pillars: player-centered design (from user testing during development to participatory design workshops during the design phase, projects start with inquiries that are influenced by ethnographic research), iterative development (the team establishes multiple milestones, and user testing culminates in a final prototype that can be evaluated), interdisciplinary teamwork (collaboration between instructional and game designers), and integration of play and learning (seamless blend between the game vision and core mechanics on the one hand and learning principles on the other hand).
- Yussof et al [38]
  - The authors propose a serious game conceptual framework. The suggested outline combines gaming requirements with learning and pedagogy theory to provide a conceptual framework for serious game designers and educators.
- Gee and Hayes [39]
  - The authors adapted the mechanics, dynamics, and aesthetics (MDA) framework into the design, play, and experience (DPE) framework. The extended DPE framework shows serious game layers for storytelling, learning, game play, and user experience. Every layer includes design, play, and experience.
- Roungas and Dalpiaz [40]
  - The authors created a conceptual model of serious games to reduce misconceptions in serious game design teams by specifying a standard terminology that stakeholders can accept. The conceptual model also guides serious game design to address Game Design Document and other record keeping and administrative process inconsistency. Combining educational and game elements is the main challenge. Completed conceptual models are displayed in unified modeling language (UML) class diagrams.
- Breuer and Bente [41]
  - The authors examine how serious games relate to e-learning and game-based learning. Serious games may use different learning strategies than edutainment and e-learning, according to them. According to the authors, many serious game definitions and typologies are limited.
- Ferdig [34], Rooney [36], and Deci and Ryan [42].

- Our novel synthesis combines the DPE framework, the serious game design framework proposed by Rooney [36], and self-determination theory (SDT). The idea emphasizes the importance of theory (pedagogy), content (fidelity), and game design (play) in serious game design. Effective serious game development is said to be central to these elements. In the DPE framework, SDT principles such as relatedness (the desire to connect with others), autonomy (the desire to choose one's own paths), mastery (the desire to develop skills and master them), and purpose (the desire to connect actions with greater reason) are proposed to clarify or distinguish the connections between human psychological patterns and game features, mechanics, and dynamics to argue that gaming approaches and thinking can be successful. All 3 theories are combined to create a new serious game development strategy. The final stakeholder-centered framework partially incorporates these theories, but much of it leads the authors of this study to literature on game design.

Variant 2: Position, Activity, and Specialization

The next stakeholder-centered framework revisits unknowns and defines user problems to generate problem statements for subsequent design phases. Recordkeeping is stressed to avoid future issues. Stakeholder theory is emphasized, and how to identify and analyze serious game stakeholders is a key question. These stakeholders include experts, developers, and consumers, whose power and interest are analyzed using stakeholder analysis methods such as the power-interest grid. In the initial framework visualization, the EA pillar influences responsible, accountable, consulted, and informed matrices; Gantt charts; and business process model and notation swimlanes. In addition, variant 2 introduces 2 phases, idea validation and conceptualization, which continue in subsequent variants. We also discuss the 4 main serious game stakeholders from a previous stakeholder management approach: development team, publishers, context-related staff, and supplemental staff. Consumer stakeholders are consulted during development, but only the 3 (or 4) main categories are relevant to core game production.

As shown in Table 1, serious game production stakeholders often play multiple roles in smaller teams due to constraints. Variant 2 of the framework includes idea validation and conceptualization. The former evaluates the team's serious game development prospects, while the latter starts project ideation. The framework has 3 levels: *execution*, *inquiry*, and *practical*. Serious game design stages include idea validation, conceptualization, development, and iteration in the execution level. Academic research and inquiry on serious game manufacture, participatory design, and more occur at the inquiry level. Stage-specific requirements and outcomes are listed in the practical level checklist.

The 3 levels are necessary due to the complexity of serious game development. The variant 2 framework shown in Figure 1 [43] also includes TOGAF ADM, DSR design, and the agile software development life cycle. Collaboration, adaptability, and rapid prototyping are hallmarks of agile software development. Rapid prototyping, customer focus, flexibility, and serious game development improvement are promoted by this approach. Serious game development levels include TOGAF ADM, DSR design, and the agile software development life cycle.

Table 1. Serious game stakeholder categories, positions, activities, and specializations.

Category	Positions	Activities	Specializations
Development team	Programmer, artist, designer, producer, tester, composer, sound designer, and writer	The different tasks that game designers may perform during game development include coding, developing AI[a] systems and camera systems, drawing characters and environments, designing UI[b] elements, populating levels, managing the development team, managing schedules and resources, testing the game for bugs, creating music and sound effects, writing character dialogue, and setting up game objective prompts. Designers may also distinguish good from bad games and explain why, as well as ensure that the game achieves its goals and maintains its vision.	Over time, professionals specialize in 2D or 3D graphics, physics, mathematics, particle systems, UI, AI, input devices, and computer networking. Storyboard artist, concept artist, 3D modeler, environmental artist, texture artist, visual effects artist, UI artist, animator, technical artist, art director, level designer, game designer, system designer, scripter, combat designer, creative director, executive producer, associate producers, and assistant producers can specialize. Some positions are hired later in the development process and may be considered freelance rather than full time.
Publishing team	Product manager, project manager, creative manager, art director, technical director, marketer, and players and users	Publishing team members set priorities, review project milestones, set and meet targets or deadlines, provide feedback on improvements, collaborate with marketing to develop packaging and other visual assets, and promote the game. They also mediate between the studio and the publisher’s legal department, work with licensors and the ESRB[c] to secure a rating, and provide technical support. They may also cooperate with marketing and PR[d] on press materials, co-design the game, and improve its visual language.	These positions may or may not affect the game’s content and aim to streamline development and maintain quality within budget and time constraints. Game designers or writers in publishing usually fill these positions, which may vary in involvement depending on the publisher. In addition to programming, they may handle management issues and work with media outlets and advertising firms with different needs and capabilities.
Context-related team	Subject matter expert, educational theorist, scholar, and research director	They consult with teams on game content and requirements, provide educational material, maintain educational aspects, investigate and test game features, and manage or supervise scholars and data collection. They also propose, choose methods, supervise, budget, and report.	In all serious game design projects, these stakeholders must provide sufficient materials to address serious issues and express them through gameplay. Although they may not be educators, they focus on game curriculum and syllabus development. They may be specialists in a research field or pursuing specialization.
Supplementary team	Business developer, lawyer, brand manager, PR manager, quality assurance manager, talent recruiter, human resource officer, game reviewer, licensor, and funding entity	They create business opportunities, secure funding, and invest in games. They also provide legal advice, review contracts, and handle licensor negotiations. In addition, they maintain brand representation in the game and work with marketing on packaging, create marketing strategies, contact gaming publications and blogs, and organize press events. Moreover, they manage the test department, send developers bug sheets, and ensure quality, as well as recruit, manage, and train new hires.	Their responsibilities include building relationships with teams, reviewing game demonstrations, negotiating contracts, generating marketing strategies, managing the employment process, playing and reviewing games, and securing funding for serious game projects. Game producers run the test department, organize press events, recruit talent, and invest in serious game projects. These people play games, write reviews, and suggest improvements. In addition, they license IP[e] and may work with licensors to get ratings. Moreover, they financially support serious game projects and crowdfund.

[a]AI: artificial intelligence.

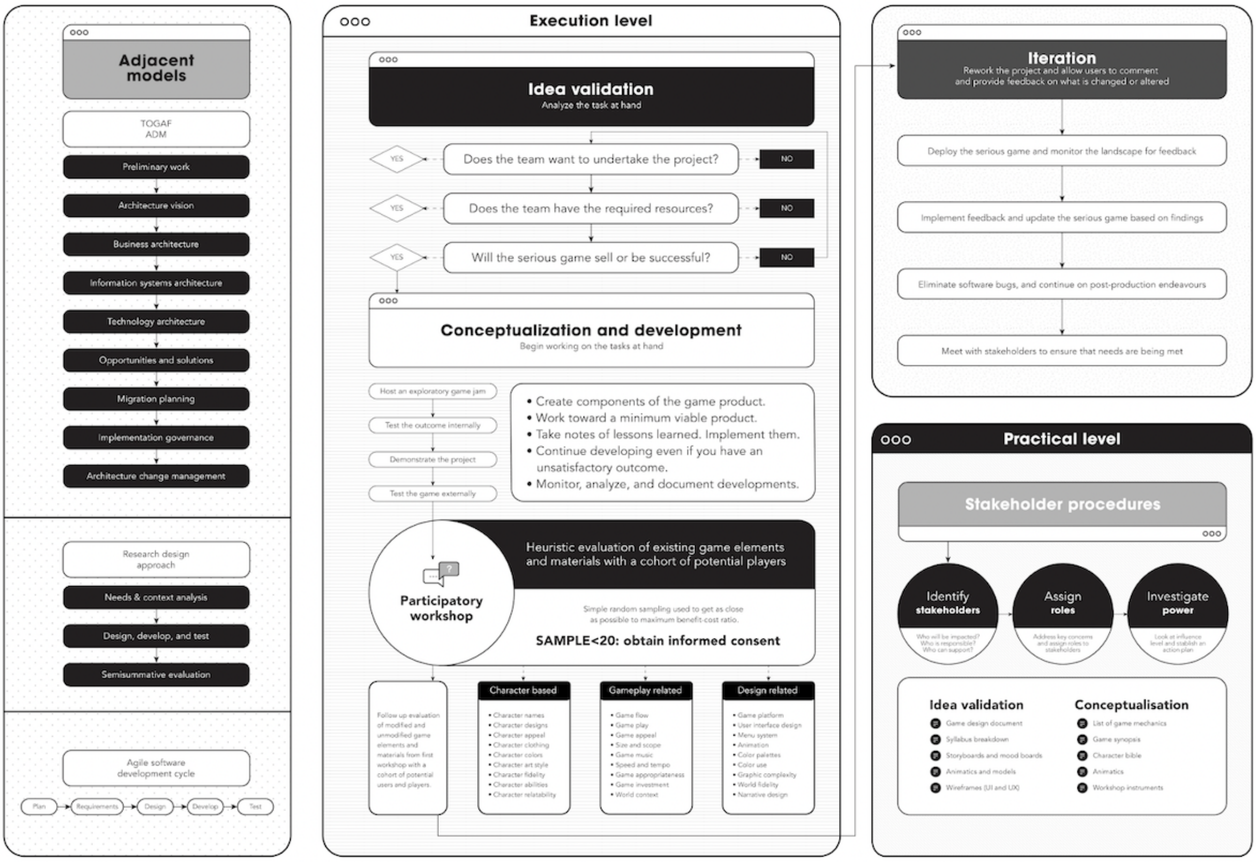
[b]UI: user interface.

[c]ESRB: Entertainment Software Rating Board.

[d]PR: public relations.

[e]IP: intellectual property.

Figure 1. Simplified variant 2 of the stakeholder-centered framework. ADM: Architecture Development Method; TOGAF: The Open Group Architecture Framework; UI: user interface; UX: user experience. For a higher-resolution version of this figure, see Multimedia Appendix 1.



Variant 3: Refinement

Variant 3 of the framework emphasizes human-centeredness and separates idea validation and evaluation (Table 2). The 3 levels are *EA*, *execution*, and *practical*. Each level combines serious game and stakeholder theory literature, practices, and methods, but only the EA level fully represents 1 of the 3 TOGAF ADM research pillars. The levels help practitioners avoid not seeing the forest for the trees and understand the framework's many components, mechanisms, and prescriptions. The composite nature of serious game development makes the framework stratiform, and the levels isolate and aggregate all interaction-based components, connectors, and relations for every aspect of the system's functionality into a single structural model.

Table 2. The third framework variant.

EA[a] level	Execution level	Practical level
This level establishes the strategic framework for serious game design, aligning game objectives with organizational goals and stakeholder needs through a comprehensive stakeholder analysis.	It transitions the process from strategic planning to the tangible design and development phases, detailing the game's mechanics, story, and technical requirements to ensure alignment with the defined objectives.	• This final level focuses on the deployment, testing, and evaluation of the game in real-world scenarios, emphasizing the adjustment of the game design based on user feedback and the effectiveness of achieving intended outcomes.
Core functions		

This level focuses on the strategic aspects of serious game design, aligning game objectives with broader organizational or project goals. It incorporates stakeholder analysis to ensure that the game's objectives meet the needs and expectations of all relevant parties.	At this level, the framework transitions from strategic planning to operational design and development. It involves the detailed design of the game, including gameplay mechanics, narrative elements, and technical specifications. This level ensures that the game's design is feasible and aligns with the strategic objectives outlined at the EA level.	• The practical level is where the game is deployed and assessed in real-world settings. This involves testing, gathering feedback from end users and stakeholders, and iterating on the design based on this feedback. The focus here is on practical application and the effectiveness of the game in achieving its intended outcomes. The level offers options for game developers:
		• Prediscovery stage
		  • Basic: stakeholder team selection, assembled team, and game design document
		  • Standard: basic outputs, selection of game mechanics, and storyboards
		  • Advanced: standard outputs and detailed curriculum itemization
		• Production
		  • Basic: stakeholder prioritization, game synopsis, and character bible
		  • Standard: basic outputs and mood boards
		  • Advanced: standard outputs, wireframes, and animatics
		• Periphery
		  • Basic: Ongoing stakeholder prioritization, game art development, and level design
		  • Standard: basic outputs and prototype development
		  • Advanced: standard outputs, deeper design practice, and quality assurance

[a]EA: enterprise architecture.

Ethical Considerations

This study was approved by the Institutional Review Board of North-West University, ensuring adherence to ethical standards in research involving human participants (approval number: NWU-01775-20-A9). Informed consent was obtained from all participants prior to their inclusion in the study (in the expert review questionnaire). Participants were informed of their right to opt out at any time without any consequences. Data collected during this study were anonymized to protect participant confidentiality. Identifiable information was removed, and data were stored securely in a password-protected database. Participants were not compensated for their time and effort in participating in the study. They were, however, promised a copy of the academic work once published.

Results

Expert Review Analysis

Overview

The expert review questionnaire regarding the stakeholder-centered framework was distributed to 220 serious game practitioners and experts internationally, of whom 29 (13.2%) completed it. On average, questionnaire completion took 57.2 (range 14.5-252) minutes. Considering the in-depth nature of the research, a completion time of approximately 24 minutes (excluding the outlier) is acceptable, despite the recommended 15-minute length for questionnaires. This study's niche focus on serious games results in a smaller expert pool; thus, the response rate and data volume are considered satisfactory. The questionnaire, designed for comprehensive data collection on the stakeholder-centered framework, uses both qualitative and inferential statistical analyses.

The research accounts for web-based survey challenges by ensuring content validity and question clarity, balancing open-ended and closed-ended questions, and maintaining reliability. Despite initial plans, level-specific explainer videos were excluded to prevent extending the questionnaire's length. An introductory explainer video was provided [44], and participants had full access to the framework for a thorough review. To ensure depth and accuracy, participants were granted access to all aspects of the framework to ensure that they could perform a multifaceted expert review.

Section A: General Information

The questionnaire respondents predominantly skewed younger, with 45% (13/29) aged 26 to 35 years and 38% (11/29) aged 36 to 45 years. Those aged 46 to 55 years constituted 10% (3/29) of the sample, while those aged 18 to 25 years and ≥66 years each represented 3% (1/29). White individuals made up 69% (20/29) of the respondents, followed by 14% (4/29) of individuals of other ethnic backgrounds including Hispanic individuals, Latinx individuals, people of color, and others. Asian respondents accounted for 7% (2/29) of the sample; and African, Indian, and undisclosed categories each accounted for 3% (1/29). Gender distribution among the respondents was fairly even, with 56% (16/29) identifying as man and 41% (12/29) as woman; of the 29 respondents, 1 (3%) preferred not to disclose their gender. In terms of geography, 45% (14/29) of the participants were from South Africa, reflecting the study's origin and local interest. Thailand and Australia each contributed 7% (2/29) of the respondents, while the remaining countries (11/13, 85%) each contributed 3% (1/29) of the respondents, broadening the international representation.

Section B: Game Development Experience

Overview

The survey section on game design experience collected data on qualifications, occupations, and development experience, including roles and satisfaction in game development. The respondents had high qualifications, with 55% (16/29) holding doctoral degrees, 24% (7/29) master's degrees, and 14% (4/29) honors degrees. The occupational profile was academic-centric, with 24% (7/29) being lecturers and 17% (5/29) senior lecturers. Others (17/29, 59%) included professors, researchers, and various roles in private industry. A significant proportion of the respondents (16/29, 56%) had >5 years of game development experience, showcasing their expertise in the field. Most (22/29, 76%) had also been involved in serious game development, although a few (3/29, 10%) had not, and some (2/29, 7%) were unsure or had projects in development. In terms of team experience, the majority (22/29, 76%) affirmed involvement, with a small percentage (3/29, 10%) either working independently or not at all in serious game development. Satisfaction across 11 development factors (DFs) was measured, with the highest scores (out of 4) being for collaboration (3.54), skills (3.56), vision (3.56), and educational aspects (3.7) and the lowest for management (2.85). The DFs and their resulting scores were as follows:

- DF1: collaboration, average score=3.54
- DF2: communication, average score=3.22
- DF3: resources, average score=3.19
- DF4: team composition, average score=3.3
- DF5: skills, average score=3.56
- DF6: management, average score=2.85 (this is the lowest average score among the 11 DFs studied)
- DF7: vision, average score=3.56
- DF8: procedures and processes, average score=2.96
- DF9: outcomes, average score=3.12
- DF10: conflict, average score=3.31
- DF11: educational and edifying aspects, average score=3.7 (this is the highest average score among the 11 DFs investigated)

Respondents expressed their views on various aspects of serious game development in this section:

- Collaboration (DF1): respondents were largely satisfied with their collaborative efforts in developing games.
- Communication (DF2): although rated slightly lower than collaboration, communication during serious game design was still positively regarded.
- Resources (DF3): the resources available for serious game development, including educational materials, software tools, and marketing aids, were deemed satisfactory.
- Team composition (DF4): the composition of serious game teams was viewed favorably, with the right mix of skills and expertise viewed as to the team's goals and performance.
- Skills (DF5): team members' skills were rated as fitting for serious game development tasks.
- Management (DF6): satisfaction with management was moderate, indicating that some areas may require improvement.
- Vision (DF7): respondents were content with the guiding visions for serious game projects, which help in goal setting and decision-making.
- Procedures and processes (DF8): there was some dissatisfaction with the processes involved in transforming ideas into final products.
- Outcomes (DF9): the outcomes of serious game projects were generally met with approval, suggesting satisfaction with the services or interventions provided.
- Conflict (DF10): opinions on conflict were mixed but leaned toward satisfaction with handling disagreements during serious game projects.
- Educational and edifying aspects (DF11): given the respondents' backgrounds in education and research, they highly rated the educational value of the games produced.

The section B responses indicated that the experts were well-versed in game development, with a specific focus on serious game development. Their moderate satisfaction across key production factors attested to their practical experience, reinforcing the study's credibility and reliability. Predominantly researchers, these individuals engage deeply with the field, often acting as SMEs in serious game projects. The most frequently reported challenges were resource-related: time, budget, and skills. Acknowledging these common hurdles faced by serious game professionals helps refine the framework to address and mitigate such issues more effectively.

Serious Game Development Roles

Researcher emerged as the most common role among serious game professionals, accounting for 14% (4/29) of the respondents, highlighting their involvement in data collection, analysis, and contribution to scholarly literature. The role of *educational theorist* followed at 10% (3/29), underscoring expertise in teaching methods. *Content expert*, *designer*, and *tester* each constituted 9% (3/29) of the respondents. Additional roles such as *project manager*, *CEO* (chief executive officer), and *UX* (user experience)/*UI* (user interface) designer were specified under *Other*. With education-related roles being predominant, this reflects the survey's findings on respondent occupations. A total of 176 roles were reported, averaging 6 roles per person, indicating the multifaceted nature of serious game stakeholder involvement. The diversity of roles suggests that stakeholders often wear multiple hats in their projects. Notably, *lawyer* and *licensor* were the only roles not represented among the respondents.

Serious Game Development Activities

Respondents reported a broad spectrum of activities within serious game development, categorized into *preactivity stage*, *development*, *postactivity stage*, *continuous*, and *unknown*:

- *Preactivities* are preparatory steps such as topic research, fundraising, context analysis, problem definition, game scope determination, learning content creation, and initial consultations.
- *Development* activities encompass the actual creation process, including game design, iteration, implementation, programming, artwork, and character design.
- *Postactivities* might consist of usability testing and game evaluations, depending on the project's goals.
- *Continuous* activities are ongoing tasks such as management, research, education, administration, and marketing that span the project's life cycle.
- *Unknown* captures any unclear or undefined responses.

The bulk of the feedback pertained to the hands-on development tasks—programming, art, writing, and design—aligning with the framework's emphasis on development processes. Game research and evaluation were equally represented, each with 9 mentions, while learning content development received 7 mentions, reflecting the educational aspect of serious game projects.

Serious Game Development Issues

Respondents were asked about common issues encountered during game development, with the question focused on *resources* and *game-specific* challenges.

Resource-related issues highlighted included the following:

- Time management, with 7 (24%) of the 29 respondents noting the extensive duration needed for serious game projects, often described as time consuming and unrealistic
- Budget constraints, also mentioned by 7 (24%) of the 29 respondents, indicating that limited funding, especially within educational environments, affects the scope of development
- Skills shortage, with responses pointing to a lack of necessary expertise and experience among serious game stakeholders
- Team-related factors, with, for example, size and composition, tools for development, intellectual property concerns, and marketing resources highlighted as challenges

Game-specific issues centered on the following aspects:

- The balance between educational content and entertainment value, with respondents expressing difficulty in finding the right mix
- Validation of serious game effectiveness, including measuring the impact of serious games on players, which was mentioned as a key concern

End-user considerations include player demographics, abilities, and gaming background, along with their engagement levels and ability to reach states of flow during gameplay.

Sections C, D, and E: Framework-Level Impressions

The 3 levels of the framework are *EA*, *execution* (process oriented), and *practical* (outcomes). The following aspects of the conceptual framework levels were examined in sections C, D, and E of the expert review questionnaire:

- Comprehensibility: the degree to which the framework, including its overall structure and key components, can be understood and comprehended by its intended audience
- Fluency: the ease with which the framework can be applied or implemented by its users, considering the clarity of instructions and the usability of any associated tools and resources
- Length: the appropriate duration or scope of the framework to ensure that it is neither too long nor too short and provides adequate guidance to achieve the desired results
- Accessibility: the extent to which the framework is accessible to all potential users, including those with physical or cognitive limitations, and the availability of the resources required to implement the framework
- Applicability: the relevance and utility of the framework in addressing the challenges or opportunities it is intended to address
- Utility: the effectiveness of the framework in achieving its intended outcomes, including its capacity to produce measurable and quantifiable outcomes
- Contextuality: the extent to which the framework is tailored to the context or situation in which it will be applied, including cultural and social considerations
- Outputs: the tangible and measurable results or outcomes produced by the application of the framework, such as changes in behavior and performance enhancements, as well as other demonstrable effects

The results from sections C, D, and E are presented in Table 3.

Table 3. Questionnaire results for sections C, D, and E (n=29).

Framework aspect	5-point scale ratings					Overall comments
	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly disagree, n (%)	
Comprehensibility (clarity of framework structure and components)						
EA[a] level	—[b]	9 (33)	11 (41)	—	7 (26)	Needs more clarity
Execution level	—	21 (75)	6 (21)	—	1 (4)	Clear to most respondents
Practical level	23 (81)	—	—	—	—	Rated highest for clarity
Fluency (ease of applying or implementing the framework)						
EA level	19 (68)	—	8 (29)	—	1 (4)	Logical progression noted
Execution level	22 (79)	—	4 (14)	—	2 (7)	Highly logical flow
Practical level	23 (81)	—	—	—	—	Streamlined progression
Length (adequacy of framework duration and scope)						
EA level	—	7 (27)	13 (46)	7 (27)	—	Divided opinions
Execution level	13 (46)	—	11 (39)	4 (14)	—	Some found it lengthy
Practical level	17 (62)	—	—	—	—	Most agreeable length
Accessibility (ease of access for all users, including those with limitations)						
EA level	—	5 (20)	12 (44)	10 (36)	—	Complex for laypersons
Execution level	—	13 (48)	11 (41)	3 (11)	—	Accessible to general users
Practical level	13 (48)	—	—	—	—	Most accessible level
Applicability (relevance and adaptability of the framework to different scenarios)						
EA level	—	11 (39)	13 (48)	3 (11)	—	Requires more relevance
Execution level	19 (68)	—	7 (25)	2 (7)	—	Pertinent to game design
Practical level	21 (74)	—	—	—	—	Highly relevant to practice
Utility (effectiveness in producing intended outcomes)						
EA level	—	10 (36)	15 (54)	3 (11)	—	Utility is acknowledged
Execution level	—	10 (36)	15 (54)	3 (11)	—	Comparable to EA level
Practical level	13 (46)	—	13 (46)	—	—	Divided on effectiveness
Contextuality (suitability of the framework for various cultural and social settings)						
EA level	—	10 (36)	15 (52)	3 (11)	—	Needs more adaptability
Execution level	15 (52)	—	7 (26)	3 (11)	—	Flexible across settings
Practical level	13 (48)	—	—	—	—	Well-tailored to contexts
Outputs (tangible results or benefits from using the framework)						
EA level	—	10 (37)	13 (48)	4 (15)	—	Essential outputs produced
Execution level	16 (57)	—	9 (32)	3 (11)	—	Effective in generating outcomes
Practical level	19 (67)	—	—	—	—	Highest positive review

[a]EA: enterprise architecture.

[b]Not applicable.

Section F: Overall Impressions

Expert feedback on the stakeholder-centered framework revealed several key themes. Most of the participants agreed that the framework is useful for facilitating serious game development, highlighting its organized approach and detailed guidance. However, worries about its intricacy indicate that it could be overwhelming for smaller teams or individuals inexperienced in serious game development. Some of the respondents proposed that the framework should be used primarily as a diagnostic tool rather than a prescriptive one, hinting at a possible adjustment in its application to better suit serious game developers' varying levels of expertise.

Feedback on the framework's features suggested the necessity for additional refinement to improve its structure and comprehensibility. Respondents requested a clearer definition of serious game design elements and a greater emphasis on the gameplay experience. The relationship between the different levels of the framework and how well they work together with other frameworks were recognized as areas that require focus. There was a clear desire for participatory processes, emphasizing a preference for a stakeholder-centered approach that maximizes production and team engagement.

The framework's implementation elicited varied responses, with a significant portion finding it easy to use, while a noteworthy percentage encountered challenges. These observations emphasize the significance of customized training and emphasize the intricate nature of the framework. Providing stakeholders with necessary tools and ensuring that the framework is easily accessible can help alleviate these challenges.

Opinions on the provision of necessary development instruments were mostly positive, although some of the respondents noted that the framework does not fully address all aspects of information systems development or early serious game design requirements. This indicates potential for growth and a more detailed incorporation of serious game design mechanisms. Many of the respondents viewed the motivation to excel while using the framework positively because it offers clarity on roles and progression through the serious game development stages. However, some of the respondents doubted its impact on motivation, citing the possibility of heightened demands because of the framework's procedural intricacy. Respondents had a positive outlook on how the framework would affect stakeholder efficiency and management, expecting enhancements in planning and stakeholder engagement. However, some of them believed that it might lead to increased resource demands and project delays, highlighting the importance of finding a balance between specific instructions and managing work efficiently.

Summary

Overall, the stakeholder-centered framework was acknowledged as a valuable tool for serious game development, but it requires simplification and more user-friendly adjustments. The feedback is crucial for future improvements, guaranteeing that the framework stays pertinent and efficient for various serious game development scenarios.

Discussion

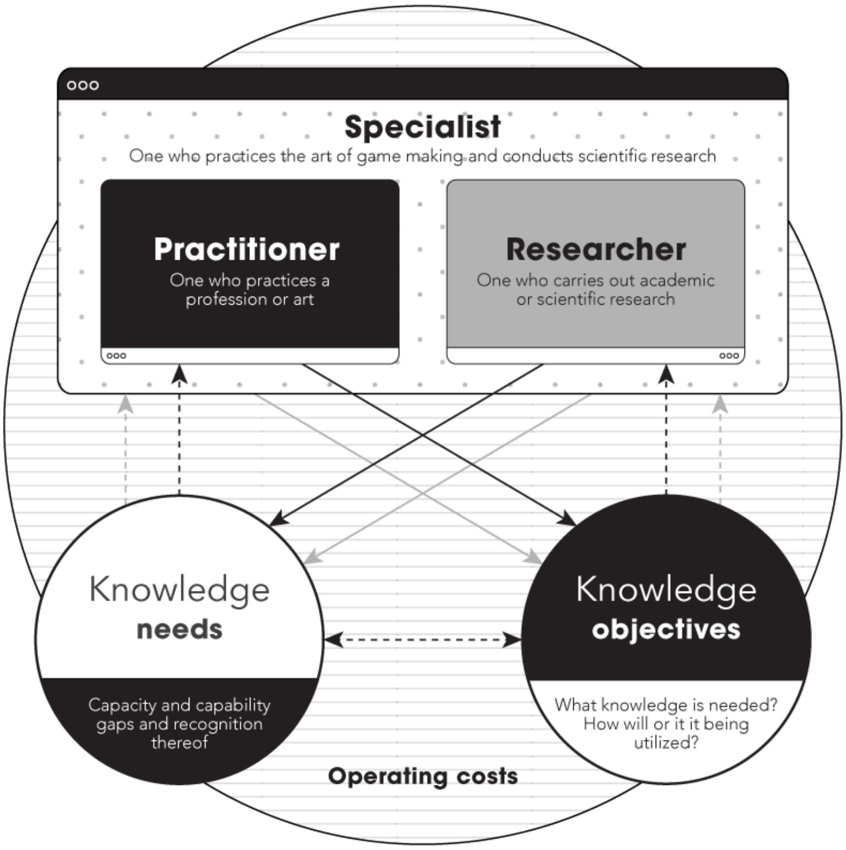
Principal Findings

The study highlights that serious game practitioners, researchers, and specialists have varying knowledge needs and objectives; for instance, a serious game practitioner in private industry may seek financial information related to serious games for profitability, while a researcher may focus on evaluating the effectiveness of serious game media. Serious game specialists may require a combination of knowledge needs to fulfil their role. As indicated in Figure 2, it is crucial for managerial staff to be aware of these differences and understand their team members' knowledge needs and objectives to effectively manage operating costs and stakeholder needs in a serious game project.

The neglect of elements such as threat assessment in serious game practice can lead to increased risk, which can impact revenue and project success. It is important for administrative and managerial staff to consider different types of risks, such as integrated, behavioral, strategic, financial, compliance, legal, and operational risks before, during, and after serious game production. Risk assessment has significant implications for stakeholder management, uncertainty management, hazard evaluation, control measures, and workplace safety and should be considered in any framework aimed at supporting stakeholders in creating effective serious game media.

The stakeholder-centered framework is mostly prescriptive and lacks personalization, and stakeholder input is necessary for improvement. The framework also lacks emphasis on game design principles and evaluation. Future research can explore ways to facilitate stakeholder participation and integrate *serious* considerations such as learning analytics, knowledge management systems, evaluation frameworks, and more.

Figure 2. Overcoming variance in knowledge needs and objectives for serious game practitioners, researchers, and specialists.



Limitations

One limitation is the challenge of incorporating 3 different theoretical pillars, which increases the complexity and can make it difficult to achieve brevity in practical research. In addition, the qualitative nature of the investigation means that the focus is on a specific sample group with distinct demographic, psychological, social, and cultural traits, making it difficult to generalize the findings to all comparable groups or circumstances. As a result, transferability is more relevant for this qualitative research.

Time constraints are another limitation imposed on the study because it is cross-sectional in nature, giving us a limited time frame to deliver our work for examination. However, the DSR approach allows the framework to undergo imminent development and iteration.

The modest sample size of the study (n=29) could also be considered a limitation because serious games are a niche field that often require expertise in education, health, or public policy, which may limit their developers. However, we contacted 220 people to take part in the study after extensive market research, and additional data would have reduced random variation and increased statistical power, making the research more accurate and reliable.

Another limitation of the research is that the framework could only be evaluated on a particular level of response assessment. Future studies should be conducted with teams to see how the framework functions in practice, according to all 4 levels of artifact assessment: response (participant feelings), learning (knowledge transfer), behavior (work performance), and outcomes (effect over time).

Finally, the reliability of questionnaire data analysis is highly dependent on several factors, such as the quality and depth of the responses, the structure of the questionnaire, and the lack of observations regarding alterations in the respondents' states of mind, feelings, and behaviors. Therefore, these factors should be taken into consideration when interpreting the results of the study.

Recommendations

We offer the following recommendations for future work:

- Forthcoming work on this topic should isolate each theoretical domain, examine them discretely, and combine, compare, and synthesize the results.
- A positivistic study that gathers quantitative data would intensify the generalizability of the findings relating to the stakeholder-centered framework. Quantitative research, such as experimental studies, offers a good basis for developing wide generalizability, given that generalizability requires data from large populations.
- Longitudinal research over an extended period of time could better assess the affects and effects of the framework. The analysis could also be richer if the inquiry extends beyond a single moment in time. DSR is typically carried out in iterative cycles of design, implementation, and evaluation, which enables researchers to refine and improve their solutions over time, allowing for strong longitudinal studies. This iterative approach also allows for data collection at multiple points in time, which can provide insights into the

effectiveness and long-term viability of the solutions being developed. The framework becomes a living artifact in this way.

- A larger sample population can feature in impending studies. The greater the sample size, the more precise the calculated mean values will be. Error margins are also reduced if a bigger sample is used.
- More participants enable the facts to speak for themselves, rather than depending on assumptions and the researcher's subjective relationships with the data. Additional data also lead to more accurate and precise units of analysis.
- The stakeholder-centered framework can be assessed according to all 4 levels of artifact assessment proposed by Petri and von Wangenheim [45]. A longitudinal study of this nature would be equipped to establish the effectiveness of the framework regarding its learning potential, behavioral impacts, and outcomes.
- The methods used to appraise the framework could be expanded in future work. This would improve the reliability of the data analysis carried out; for example, structured interviews, semistructured interviews, in-depth interviews, focus groups, field research, ethnography, and observation could be used to strengthen analysis efforts.

Features for the Framework Going Forward

Now that that main recommendations for future research have been presented, we need to consider the requisite features for the framework going forward (Table 4).

Variant 4 of the stakeholder-centered framework reflects the academic and investigative nature of serious games and their development (Figure 3). It includes instructions, demonstrations, descriptions, and definitions that facilitate game design and development and focuses on procedural information rather than technology, expertise, or resources. The length of the tool has been reduced, and the framework is less prescriptive, providing flexibility in project assumptions, goals, and processes. The progressive web application version of the framework enables users to take part in conversations with one another, categorizes procedural information into subdivisions, and includes built-in support mechanisms. The application affords the researcher added control over the transmission, presentation, structure, and extent of the intelligent system and can adapt to any changing needs or patterns of its user base.

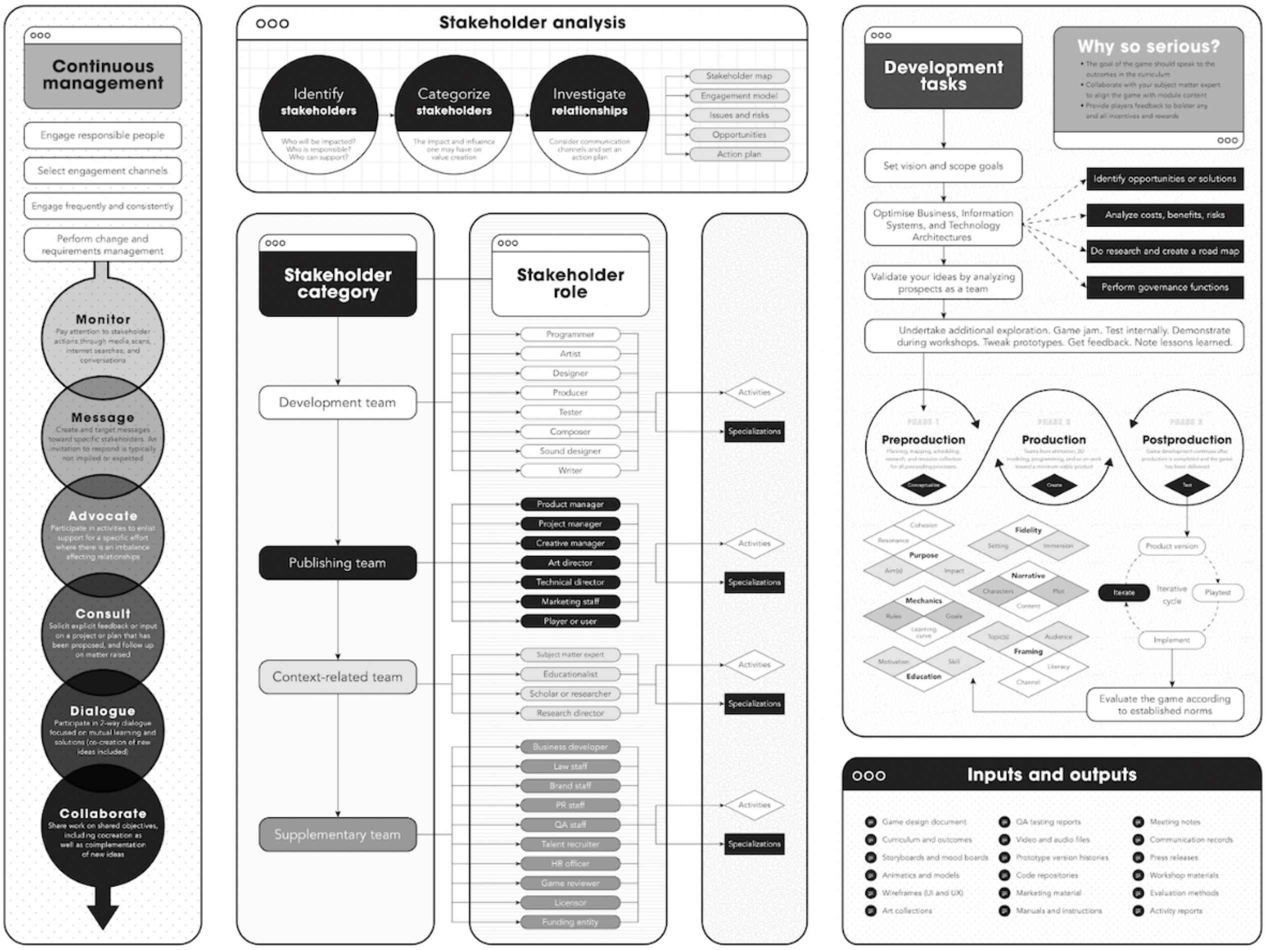
Variant 4 also integrates the 3 theoretical domains in a more subtle manner than previous versions. It presents a terser technical diagram for serious game design that omits some information to improve accessibility and usability. The procedures still begin with stakeholder analysis, categorizing stakeholders by their impact and influence on value creation in the development endeavors. The resulting stakeholder categories are development, publishing, context-related, and supplementary teams, which consist of stakeholder roles with their own activities and specializations.

Table 4. How desired framework traits correlate to improvement areas, as well as evidence for the intersection thereof in the progressive web artifact.

Desired framework trait	Improvement areas (from expert review)										Artifact execution (within the progressive web application)
	Abbreviation	Concentration	Configuration	Guidance	Idealism	Milestones	Prescription	Semantics	Simplification	Transmission	
Concerned with serious aspects		✓		✓			✓	✓			Serious game conventions built into the artifact: principles, designs, and evaluations
Concise	✓		✓					✓	✓	✓	Artifact uses chunking to boost content processing
Diagnostic		✓		✓		✓				✓	Artifact facilitates the achievement of development objectives through measurement
Flexible	✓		✓	✓			✓		✓	✓	Artifact is open to changes in assumptions, goals, and process
Informative		✓		✓	✓	✓		✓			Artifact includes instructions, demonstrations, descriptions, and definitions on serious game design
Invested in work, not technology		✓	✓	✓				✓		✓	Artifact makes provision for technology but focuses on procedural information
Participatory		✓	✓				✓			✓	Discussion is facilitated in the artifact by way of private and community chat functions
Procedural			✓	✓	✓	✓	✓			✓	Various activities are divided into practical sections and directed at various stakeholders

Desired framework trait	Improvement areas (from expert review)										Artifact execution (within the progressive web application)
	Abbreviation	Concentration	Configuration	Guidance	Idealism	Milestones	Prescription	Semantics	Simplification	Transmission	
Relevant		✓		✓				✓			Extraneous information is removed using filters, and streams of material are categorized
Repeatable	✓			✓		✓			✓	✓	Sets of actions provided in the artifact are reusable and easily duplicated
Stakeholder centered		✓			✓		✓			✓	Sections of the artifact are targeted toward specific stakeholders
Supportive			✓	✓		✓	✓		✓		Frequently asked questions, data protection, and self or continual support
Sustainable	✓			✓		✓				✓	Scalable design is incorporated into the artifact
Usable	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	User control, consistency standards, minimalist design, and platform compatibility
Vehicle for good design		✓	✓			✓	✓			✓	Artifact encourages good design practice by encouraging users to assess their own practice

Figure 3. Variant 4 (version 2.0) of the stakeholder-centered framework. HR: human resources; MVP: minimum viable product; PR: public relations; QA: quality assurance; SME: subject matter expert; UI: user interface; UX: user experience. For a higher-resolution version of this figure, see Multimedia Appendix 2.



Conclusions

The proliferation of serious games and game-assisted learning in education and beyond requires keen awareness, careful contemplation, and steady advancement [46-48]. As serious games become more common in contexts aiming to edify in innovative ways, scholars should not only consider methods to improve the efficacy thereof but also think about how to realistically and efficiently fabricate them as well. Serious game project stakeholders need practical ways to align their interests with those of the enterprise. Stakeholder roles, activities, specializations, potential, competence, and capabilities all impact these experts' productive capacity and labor prospects. However, serious game initiatives vary significantly. To help future game makers, this research inspects serious game design stakeholders and techniques to produce a system capable of supporting these individuals in a range of environments. A stakeholder-centered framework, in this view, may help serious game developers manage their teams and drive practice in beneficial and sustainable ways. In the future, we hope that this investigation will aid in a decrease in serious game project failure, communication breakdown, and apathy regarding the genre of games intending to do more than purely entertain. However, additional research and innovation is needed in fields adjacent to, and embedded in, serious games to support the growing need for novel approaches to demonstrate, educate, simulate, and inform.

Acknowledgments

This work was undertaken as part of a doctoral study at the North-West University, Vanderbijlpark, South Africa. The authors are grateful to Professor Roelien Goede, Professor Japie Greeff, and Professor Estelle Taylor for their contribution and support. The financial support of the Unit for Data Science and Computing at North-West University is gratefully acknowledged. We used the generative AI tools QuillBot [49] to help refine text into manageable portions for the manuscript and ChatGPT [50] to condense the results presented in Table 3 into a more suitable format.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

The study was conceptualized by LB along with JG and ET. All authors were involved in the interpretation of results, manuscript write-up, and revisions. LB was involved in the design of the framework, data collection, data analysis, and the writing of methods and results. Moreover, LB was involved in data collection and logistics. LB conducted statistical analysis under the supervision of JG and ET. JG and ET supervised LB on all tasks pertaining to this project.

Conflicts of Interest

LB is supported by the Unit for Data Science and Computing, North-West University. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Simplified variant 2 of the stakeholder-centered framework. ADM: Architecture Development Method; TOGAF: The Open Group Architecture Framework; UI: user interface; UX: user experience.
[PNG File , 2321 KB - games_v12i1e48099_app1.png]

Multimedia Appendix 2
Variant 4 (version 2.0) of the stakeholder-centered framework. HR: human resources; MVP: minimum viable product; PR: public relations; QA: quality assurance; SME: subject matter expert; UI: user interface; UX: user experience.
[PNG File , 4533 KB - games_v12i1e48099_app2.png]

References

1. Girard C, Ecalle J, Magnan A. Serious games as new educational tools: how effective are they? A meta‐analysis of recent studies. Comput Assist Learn 2012 Jun 13;29(3):207-219 [FREE Full text] [doi: 10.1111/j.1365-2729.2012.00489.x]
2. Spil TA, Romijnders V, Sundaram D, Wickramasinghe N, Kijl B. Are serious games too serious? Diffusion of wearable technologies and the creation of a diffusion of serious games model. Int J Inf Manage 2021 Jun;58:102202 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102202] [Medline: 32836650]
3. Clapper TC. Serious games are not all serious. Simul Gaming 2018 Jul 17;49(4):375-377 [FREE Full text] [doi: 10.1177/1046878118789763]
4. Suryapranata LK, Soewito B, Kusuma GP, Gaol FL, Warnars HL. Quality measurement for serious games. In: Proceedings of the International Conference on Applied Computer and Communication Technologies (ComCom). 2017 Presented at: ComCom 2017; May 17-18, 2017; Jakarta, Indonesia URL: https://doi.org/10.1109/comcom.2017.8167098 [doi: 10.1109/comcom.2017.8167098]
5. Pacheco-Velazquez E, Bester A, Rabago-Mayer L, Ro V. What do we evaluate in serious games? A systematic review. In: Proceedings of the 17th European Conference on Games Based Learning. 2023 Presented at: ECGBL 2023; October 5-6, 2023; Enschede, Netherlands URL: https://doi.org/10.34190/ecgbl.17.1.1627 [doi: 10.34190/ecgbl.17.1.1627]
6. Koscianski A, Zanotto D. A design model for educational multimedia software. Creat Educ 2014 Dec;5(23) [FREE Full text] [doi: 10.4236/ce.2014.523224]
7. Taylor L. Video game internal turfs and turfs of play. M/C J 2004;7(2) [FREE Full text] [doi: 10.5204/mcj.2346]
8. Gaspar JD, Lage EM, Silva FJ, Mineiro É, Oliveira IJ, Oliveira I, et al. A mobile serious game about the pandemic (COVID-19 - did you know?): design and evaluation study. JMIR Serious Games 2020 Dec 22;8(4):e25226 [FREE Full text] [doi: 10.2196/25226] [Medline: 33301416]
9. Blakey H. Designing player intent through “playful” interaction: a case study of techniques in transistor and journey. M/C J 2021;24(4) [FREE Full text] [doi: 10.5204/mcj.2802]
10. Landers RN. Developing a theory of gamified learning: linking serious games and gamification of learning. Simul Gaming 2015 Jan 15;45(6):752-768 [FREE Full text] [doi: 10.1177/1046878114563660]
11. Haoran G, Bazakidi E, Zary N. Serious games in health professions education: review of trends and learning efficacy. Yearb Med Inform 2019 Aug 25;28(1):240-248 [FREE Full text] [doi: 10.1055/s-0039-1677904] [Medline: 31022747]
12. Carlier S, Van der Paelt S, Ongenae F, De Backere F, De Turck F. Empowering children with ASD and their parents: design of a serious game for anxiety and stress reduction. Sensors (Basel) 2020 Feb 11;20(4):966 [FREE Full text] [doi: 10.3390/s20040966] [Medline: 32054025]
13. Tan JW, Zary N. Diagnostic markers of user experience, play, and learning for digital serious games: a conceptual framework study. JMIR Serious Games 2019 Jul 16;7(3):e14620 [FREE Full text] [doi: 10.2196/14620] [Medline: 31313660]
14. Scurati GW, Nylander JW, Ferrise F, Bertoni M. Sustainability awareness in engineering design through serious gaming. Des Sci 2022 Mar 18;8:e12 [FREE Full text] [doi: 10.1017/dsj.2022.9]

15. Alvarez J, Michaud L. Serious Games : Advergaming, Edugaming, Training. Montpellier, France: IDATE Consulting & Research; 2008.
16. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. Comput Educ 2012 Sep;59(2):661-686. [doi: 10.1016/j.compedu.2012.03.004]
17. Sitzmann T. A meta‐analytic examination of the instructional effectiveness of computer‐based simulation games. Pers Psychol 2011 May 27;64(2):489-528. [doi: 10.1111/j.1744-6570.2011.01190.x]
18. Baranowski T, Blumberg F, Buday R, DeSmet A, Fiellin LE, Green CS, et al. Games for health for children-current status and needed research. Games Health J 2016 Feb;5(1):1-12 [FREE Full text] [doi: 10.1089/g4h.2015.0026] [Medline: 26262772]
19. Bogost I. Persuasive Games: The Expressive Power of Videogames. Cambridge, MA: MIT Press; 2007.
20. Michael DR, Chen · S. Serious Games: Games that Educate, Train and Inform. Boston, MA: Thomson Course Technology; 2006.
21. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
22. Plass JL, Homer BD, Kinzer CK. Foundations of game-based learning. Educ Psychol 2016 Feb 06;50(4):258-283. [doi: 10.1080/00461520.2015.1122533]
23. Garris R, Ahlers R, Driskell JE. Games, motivation, and learning: a research and practice model. Simul Gaming 2016 Aug 18;33(4):441-467. [doi: 10.1177/1046878102238607]
24. Kato PM. Video games in health care: closing the gap. Rev Gen Psychol 2010 Jun 01;14(2):113-121. [doi: 10.1037/a0019441]
25. The TOGAF ® Standard, Version 9.2. 's-Hertogenbosch, The Netherlands: Van Haren Publishing; 2018.
26. Kato PM. Evaluating efficacy and validating games for health. Games Health J 2012 Feb;1(1):74-76. [doi: 10.1089/g4h.2012.1017] [Medline: 26196436]
27. De Freitas S, Rebolledo‐Mendez G, Liarokapis F, Magoulas GD, Poulovassilis A. Learning as immersive experiences: using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world. Brit J Educ Technol 2010;41(1):69-85. [doi: 10.1111/j.1467-8535.2009.01024.x]
28. Arnab S, Brown K, Clarke S, Dunwell I, Lim T, Suttie N, et al. The development approach of a pedagogically-driven serious game to support Relationship and Sex Education (RSE) within a classroom setting. Comput Educ 2013 Nov;69:15-30. [doi: 10.1016/j.compedu.2013.06.013]
29. Freeman RE. Strategic Management: A Stakeholder Approach. Lanham, MD: Pitman Publishing; 1984.
30. Mitchell RK, Agle BR, Wood DJ. Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. Acad Manage Rev 1997 Oct;22(4):853-886. [doi: 10.2307/259247]
31. Larson EW, Gray CF. Project Management: The Managerial Process. New York, NY: McGraw-Hill Education; 2014.
32. Harrison JS, St. John CH. Foundations in Strategic Management. Boston, MA: Cengage Learning; 2013.
33. Annetta LA. The “I's” have it: a framework for serious educational game design. Rev Gen Psychol 2010 Jun 01;14(2):105-113. [doi: 10.1037/a0018985]
34. Ferdig R. Learning and teaching with electronic games. J Educ Multimed Hypermedia 2007 Jan;16(3):217-223 [FREE Full text]
35. Marne B, Wisdom J, Huynh-Kim-Bang B, Labat JM. The six facets of serious game design: a methodology enhanced by our design pattern library. In: Proceedings of the 7th European Conference on Technology Enhanced Learning. 2012 Presented at: EC-TEL 2012; September 18-21, 2012; Saarbrücken, Germany. [doi: 10.1007/978-3-642-33263-0_17]
36. Rooney P. A theoretical framework for serious game design. Int J Game Based Learn 2014 Oct;2(4):41-60. [doi: 10.4018/ijgbl.2012100103]
37. Vanden AV, De Schutter B, Geurts L, Desmet S, Wauters J, Husson J, et al. P-III: a player-centered, iterative, interdisciplinary and integrated framework for serious game design and development. In: Proceedings of the Joint Conference of the Interdisciplinary Research Group of Technology, Education, Communication, and the Scientific Network on Critical and Flexible Thinking. 2011 Presented at: ITEC/CIP/T 2011; October 19-21, 2011; Ghent, Belgium. [doi: 10.1007/978-3-642-33814-4_14]
38. Yusoff A, Crowder RM, Gilbert L, Wills G. A conceptual framework for serious games. In: Proceedings of the Ninth IEEE International Conference on Advanced Learning Technologies. 2009 Presented at: ICALT 2009; July 15-17, 2009; Riga, Latvia. [doi: 10.1109/icalt.2009.19]
39. Gee JP, Hayes E. “No quitting without saving after bad events”: gaming paradigms and learning in the sims. Int J Learn Media 2009;1(3):49-65. [doi: 10.1162/ijlm_a_00024]
40. Roungas B, Dalpiaz F. A model-driven framework for educational game design. In: Proceedings of the Games and Learning Alliance. 2015 Presented at: GALA 2015; December 9-11, 2015; Rome, Italy. [doi: 10.1007/978-3-319-40216-1]
41. Breuer JS, Bente G. Why so serious? On the relation of serious games and learning. Eludamos J Comput Game Culture 2010 Apr 26;4(1):7-24. [doi: 10.7557/23.6111]
42. Deci EL, Ryan RM. The general causality orientations scale: self-determination in personality. J Res Pers 1985 Jun;19(2):109-134. [doi: 10.1016/0092-6566(85)90023-6]

43. McKenney S, Voogt J, Kirschner PA. Learning by design: nourishing expertise and interventions. In: Superfine AC, Goldman SR, Ko ML, editors. Teacher Learning in Changing Contexts. Milton Park, UK: Routledge; 2022.
44. L Bunt - PhD Questionnaire Video. YouTube. URL: https://www.youtube.com/watch?v=S_yO1Bs2Ju0 [accessed 2024-05-23]
45. Petri G, von Wangenheim CG. A method for the evaluation of the quality of games for computing education. In: Proceedings of the Congresso Brasileiro de Informática na Educação. 2019 Presented at: CBIE 2019; November 11-14, 2019; Brasília, Brazil.
46. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020 Jul 24;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
47. Tsekleves E, Cosmas J, Aggoun A. Benefits, barriers and guideline recommendations for the implementation of serious games in education for stakeholders and policymakers. Brit J Educ Technol 2014 Oct 24;47(1):164-183. [doi: 10.1111/bjet.12223]
48. Zhonggen Y. A meta-analysis of use of serious games in education over a decade. Int J Comput Games Technol 2019 Feb 03;2019:1-8. [doi: 10.1155/2019/4797032]
49. QuillBot homepage. QuillBot. URL: https://quillbot.com/ [accessed 2024-05-23]
50. ChatGPT login. ChatGPT . URL: https://chatgpt.com/auth/login?next=%2F%3Foai-dm%3D1 [accessed 2024-05-23]

Abbreviations

ADM: Architecture Development Method
CEO: chief executive officer
DF: development factor
DSR: design science research
EA: enterprise architecture
SME: subject matter expert
TOGAF: The Open Group Architecture Framework
UI: user interface
UX: user experience

Edited by G Eysenbach, T Leung; submitted 11.04.23; peer-reviewed by V Pelser-Carstens, C Tziraki, T Baranowski; comments to author 20.01.24; revised version received 23.02.24; accepted 17.03.24; published 31.05.24.

Please cite as:
Bunt L, Greeff J, Taylor E
Enhancing Serious Game Design: Expert-Reviewed, Stakeholder-Centered Framework
JMIR Serious Games 2024;12:e48099
URL: https://games.jmir.org/2024/1/e48099
doi:10.2196/48099
PMID:38820585

Original Paper

Effects of Virtual Reality Therapy Combined With Conventional Rehabilitation on Pain, Kinematic Function, and Disability in Patients With Chronic Neck Pain: Randomized Controlled Trial

Qifan Guo[1,2,3*], MD; Liming Zhang[1,2*], BSc; Leo Lianyi Han[4], PhD; Chenfan Gui[1], MD, DPT; Guanghui Chen[5], MD; Chunyan Ling[6], MD; Wei Wang[6], BSc; Qiang Gao[1,2], MD, PhD

[1]West China Hospital, Sichuan University, Chengdu, China

[2]Department of Rehabilitation Medicine, West China Hospital, Sichuan University, Chengdu, China

[3]Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, China

[4]Biostatistics Group, State Key Laboratory of Genetic Engineering, Greater Bay Area Institute of Precision Medicine (Guangzhou), Fudan University, Guangzhou, China

[5]Department of Traumatology and Orthopedics of Traditional Chinese Medicine, The First Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, China

[6]Department of Acupuncture and Tuina, The First Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, China

[*]these authors contributed equally

Corresponding Author:
Qiang Gao, MD, PhD
West China Hospital
Sichuan University
No 37 Guoxuexiang
Chengdu, 610041
China
Phone: 86 18980605992
Email: gaoqiang_hxkf@163.com

Abstract

Background: Neck pain is a common condition that leads to neck motor dysfunction and subsequent disability, with a significant global health care burden. As a newly emerging tool, virtual reality (VR) technology has been employed to address pain and reduce disability among patients with neck pain. However, there is still a lack of high-quality studies evaluating the efficacy of VR therapy combined with conventional rehabilitation for patients with chronic neck pain, particularly in terms of kinematic function.

Objective: This study aims to investigate the effect of VR therapy combined with conventional rehabilitation on pain, kinematic function, and disability in patients with chronic neck pain.

Methods: We conducted an assessor-blinded, allocation-concealed randomized controlled trial. Sixty-four participants experiencing chronic neck pain were randomly allocated into the experimental group that underwent VR rehabilitation plus conventional rehabilitation or the control group receiving the same amount of conventional rehabilitation alone for 10 sessions over 4 weeks. Pain intensity, disability, kinematic function (cervical range of motion, proprioception, and mean and peak velocity), degree of satisfaction, and relief of symptoms were evaluated at 3 timepoints (baseline, postintervention, and at 3 months follow-up). A 2*3 mixed repeated measures analysis of variance was utilized for analyzing the difference across indicators, with a significant difference level of .05.

Results: Both groups demonstrated significant improvements in pain, disability, and kinematic functions (*P*<.05) at postintervention and at 3-month follow-up. The experimental group showed superior therapeutic outcomes compared to the control group in pain reduction (mean difference from the baseline: 5.50 vs 1.81 at posttreatment; 5.21 vs 1.91 at the 3-month follow-up, respectively; *P*<.001), disability improvement (mean difference from baseline: 3.04 vs 0.50 at posttreatment; 3.20 vs 0.85 at the 3-month follow-up, respectively; *P*<.001), and enhanced kinematic functions (*P*<.05). Moreover, participants in the experimental group reported better satisfaction and relief of symptoms than the control group (*P*<.05), with better initiative for exercising during the follow-up period. However, there was no between-group difference of improvement in proprioception. No adverse events were reported or observed in our research.

Conclusions: The findings of our study support the efficacy of combining VR therapy with conventional rehabilitation in alleviating pain, enhancing kinematic function, and reducing disability of patients with chronic neck pain. Future research should focus on refining the therapeutic protocols and dosages for VR therapy as well as on optimizing its application in clinical settings for improved convenience and effectiveness.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000040132; http://www.chictr.org.cn/showproj.aspx?proj=64346

(JMIR Serious Games 2024;12:e42829) doi:10.2196/42829

KEYWORDS

virtual reality; neck pain; disability; kinematic function; rehabilitation; physiotherapy; neck; pain; chronic; therapy; kinematic; efficacy; patient; effect

Introduction

Chronic neck pain is a prevalent global health issue, with reported prevalence rates ranging from 10% to 24% [1,2]. This condition is closely associated with motor dysfunction in the cervical region, characterized by deficits in various kinematic functions of the neck [3-5]. Cervical kinematic functions can be operationally defined as the capacity of the neck muscles to generate and regulate movement of the head and neck, including range of motion (ROM), which is the degree of movement that can be achieved in various directions of the cervical spine, velocity, coordination, strength, and endurance. These parameters can be quantified through specific evaluations and measurements such as ROM assessments, manual muscle testing, and functional movement tests. Prior studies have proven that motor dysfunction occurs commonly in patients with neck pain [6-8], and these dysfunctions are highly correlated with the level of pain and disability. That is because weakness in neck muscle strength and coordination will provide more unstable support of the neck segments and additional stress on the neck structure, which restricts the patient's neck movement and results in pain [9]. These diminished motor dysfunctions as well as worse pain and disability undoubtedly impair a patient's work performance and quality of life, leading to large economic losses [10]. Given the abovementioned pivotal role of cervical kinematics in neck pain, interventions aimed at improving motor function hold promise in managing this condition.

To date, active exercise is recommended to be a valid therapy for patients with chronic neck pain based on the current clinical guidelines [11,12]. Virtual reality (VR) is a unique form of exercise established by Morton Heiling in 1962 and has been evolving over the past 60 years [13,14]. VR technology commonly generates virtual environments resembling the real world through devices such as computers or head-mounted displays and interacts with patients to enable them to accomplish the targeted therapeutic goals [15,16]. Regarding the economics of VR treatment, studies [17,18] have reported low costs in VR-based treatments. The hardware devices involved in VR therapy are readily available and inexpensive. Additionally, the one-time cost of patient-specific VR software allows for repeated use, making VR applications relatively less expensive in medical settings. VR serves as a valuable assessment and intervention tool in rehabilitation due to its clinical benefits supported by ongoing research [19], and orthopedic and neurological rehabilitation are the common areas where VR therapy is utilized in clinical practice [20,21]. The potential therapeutic mechanisms of VR include task-oriented repetition, positive feedback, and embodied simulation [22].

As a noninvasive method of pain reduction, VR therapy, both independently and in combination with other interventions, has been investigated in numerous studies. Prior research [23-25] has demonstrated the potential of VR therapy to alleviate pain and disability in patients with orthopedic conditions such as rheumatoid arthritis, shoulder impingement syndrome, and low back pain. However, to date, there is still less research exploring the effects of VR therapy or combined treatment on individuals with chronic neck pain, particularly in terms of improving the cervical motor function [26,27]. Mukherjee et al [28] investigated the efficacy of VR therapy in the treatment of cervical spondylosis. Their findings revealed that patients who underwent VR therapy along with conventional physiotherapy demonstrated notable improvements in pain intensity and active cervical ROM (CROM) compared to those who underwent conventional therapy alone in the short-term period ($P<.05$). However, another study [29] reported that after receiving 4 weeks of VR training, patients with neck pain exhibited significant improvement in mean and maximal velocity, with no observed improvement in CROM indicators compared with the control group. A recent meta-analysis [30] consisting of 2 randomized controlled trials (RCTs) suggested that VR therapy combined with kinematic training could enhance the global perceived effect, patient satisfaction, and general health of patients with neck pain compared to treatment with kinematic training alone. However, evidence supporting the efficacy of VR therapy in strengthening cervical kinematic function remains inconclusive. Given the current gaps in research and the conflicting findings, further high-quality studies are warranted to ascertain the therapeutic effectiveness of VR therapy or combined treatments for individuals with chronic neck pain. Therefore, this RCT aims to evaluate the effects of VR therapy combined with conventional rehabilitation on pain, kinematic functions, and disability in patients with chronic neck pain.

Methods

Study Design and Ethics Approval

This study was designed as an assessor-blinded, allocation-concealed RCT (Multimedia Appendix 1). Ethics approval for this study was obtained from the West China Hospital Clinical Trials and Biomedical Ethics Committee of Sichuan University (approval: HX-IRB-AF-18-2021-1102).

This trial was registered in the Chinese Clinical Trial Registry (ChiCTR2000040132) on November 22, 2020. This study conformed to the Declaration of Helsinki, and all patients provided written consent after recruitment.

Participants

This study was conducted at the Department of Rehabilitation Medicine in West China Hospital. Patients were recruited through various channels such as social networks, posters, and brochures from October to December 2021. Inclusion criteria included age of 18 years and older, a diagnosis of chronic neck pain (>3 months) by a physician, reported pain intensity ≥3 on the numeric rating scale (NRS), and disability ≥6 on the neck disability index. Exclusion criteria included existing vestibular pathology, cervical fracture/dislocation, whiplash injuries, neurological/cardiovascular/respiratory disorders affecting patients' physical performance, inability to provide informed consent, and pregnancy.

Randomization

Randomization was performed using a computer-generated sequence from Randomization.com, with a researcher not involved in treatment overseeing the process. Patients were allocated to either the experimental group or control group based on the odd or even nature of the assigned number within sealed opaque envelopes to ensure blind allocation. Although a blinded researcher assessed the patients during the trial, blinding was not feasible for participants or therapists due to the layout of the VR therapy.

Intervention

VR Treatment

The VR equipment that we used included several hardware and software (Chengdu Feiming Technology Co, Ltd). Hardware included a Pico G2 4k head-mounted VR glass, monitor screen, and optical motion capture camera. Patients wearing VR glasses sat at a distance of 100 cm from the front of the monitor screen. The monitor will display the real-time virtual images that patients see during the experimental process. Therapists can assess the patient's real-time treatment stage by looking at the monitor screen and provide corresponding assistance. During treatment, the optical motion capture camera and customized software collected and analyzed the cervical movement trajectories. Meanwhile, considering the requirement of fully immersive VR therapy, a specific shoulder strap was designed for patient safety during treatment.

In VR therapy, 3 modules (ROM, proprioception, and velocity modules) were designed to enhance the specific kinematic functions of individuals experiencing chronic neck pain. These modules involved patients engaging in targeted cervical movements to attain therapeutic objectives through visual cues. Prior to the beginning of the treatment, each patient underwent a practice trial to mitigate any potential learning biases. Throughout the VR session, patients were fully immersed in a virtual setting resembling a living room, where they could manipulate virtual objects to interact with designated targets. The VR equipment incorporated visual and auditory feedback to augment the interactive and engaging nature of the therapy. The detailed descriptions of the 3 modules are provided below.

In the ROM module, a virtual flying bird is manipulated by the patients' cervical movement. Patients could move birds by neck flexion, extension, lateral flexion, and rotation movement to catch gold or avoid the fire rings appearing in the scene. The placement of the gold items and fire rings was based on baseline kinematic data, with the game's difficulty adjusted continuously to facilitate gradual improvements in CROM across all movement directions (Figure 1).

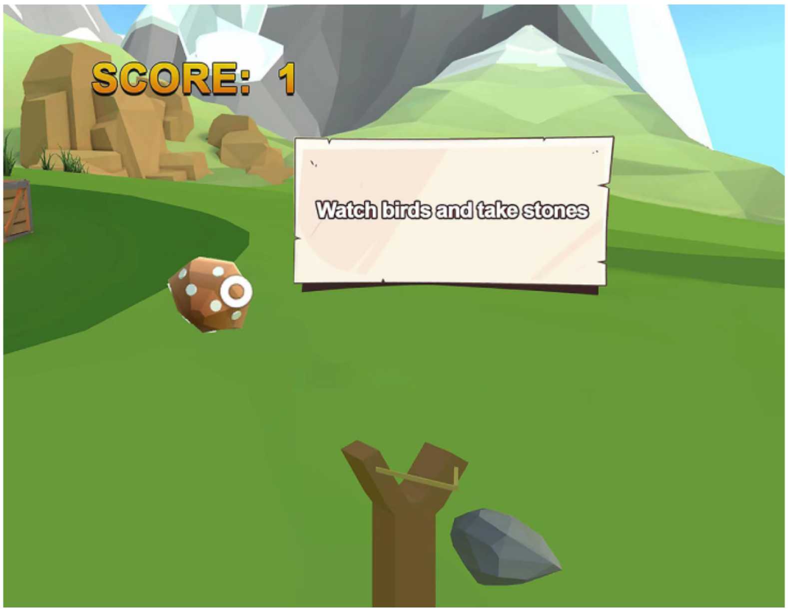
Figure 1. Range of motion module. The bird in the picture is manipulated by the patient's cervical movement (flexion, extension, lateral flexion, and rotation) to catch the gold and avoid obstacles appearing in all directions.

In the proprioception module, patients engage in tasks requiring them to control a virtual bow and arrow by using cervical movements to aim and shoot at a bull's-eye target with closed eyes. Initially, patients face the screen to align the arrow with the bull's-eye, memorizing this starting position. Subsequently, patients close their eyes and follow instructions from the VR system to move their necks to a specific position. They then have to return their neck to the initial position (representing the bull's-eye location) and shoot the arrow. The relocation error, indicating the angular deviation between the shot point and the bull's-eye, serves as a measure of patients' proprioceptive abilities (Figure 2).

Figure 2. Proprioception module. The patient was asked to remember the initial bull's-eye position and then close the eyes. Thereafter, the computer would guide the neck of the patient with the eyes closed to a specific position. The patient needs to move the neck back to the original position based on memory and manipulate the bow and arrow to shoot the bull's-eye.

In the velocity module, participants were tasked with hitting randomly appearing mushrooms within the virtual scene by manipulating virtual stones with cervical movements. Upon mushroom appearance, patients adjusted the slingshot position by moving their neck and launched a stone to hit the mushroom before it disappeared after 5 seconds. Subsequent mushrooms would appear sequentially, prompting patients to swiftly target and strike them. Patient performance was scored based on the success rate of hitting the mushrooms, thereby fostering patient engagement and compliance with the virtual therapy protocol (Figure 3).

Figure 3. Velocity module. Patients were asked to manipulate stones at maximum neck movement speed to hit randomly appearing virtual mushrooms and obtain the corresponding scores.



Conventional Rehabilitation

Patients in the experimental group received a 10-minute conventional rehabilitation session consisting of 5 minutes of active exercise (eg, muscle stretching exercise, strengthening exercise, sling exercise therapy), supported by established guidelines [11,31]. Additionally, patients underwent a 5-minute transcutaneous electrical nerve stimulation for analgesia, which was validated for efficacy in prior studies [2]. Patients in the control group were treated with 30 minutes of conventional rehabilitation, including 15 minutes of active exercise modalities same as the experimental group and an additional 15 minutes of transcutaneous electrical nerve stimulation therapy. The prescription of conventional rehabilitation was tailored based on each patient's motor dysfunction and the clinical expertise of the rehabilitation therapists.

Procedure

Sixty-four patients were randomly allocated into the experimental group or control group, and they underwent 10 treatment sessions over 4 weeks. In the experimental group, interventions included 20 minutes of VR therapy and 10 minutes of conventional rehabilitation per session, while the control group received 30 minutes of only conventional rehabilitation per session. Throughout the treatment, patient safety was closely monitored, and sessions were halted if any adverse events (eg, motion sickness, headache, falls) occurred. Patients were allowed to resume training once the symptoms subsided; otherwise, they were advised to discontinue participation in the study. Following the intervention, both groups were encouraged to continue neck exercises at home for 3 months after treatment. On completion of the follow-up period, each patient was asked to rate their frequency of neck exercise within the 3-month interval to represent their adherence to continued neck exercise. The rating scale ranged from 0 to 4 (0 = no training; 1 = 0-1 hours of training per week; 2 = 1-2 hours of training per week; 3 = 2-3 hours of training per week, ≥4 = >3 hours of training per week). A comparison of the data sets from both groups was conducted to observe the patients' initiative in training at unsupervised situations.

Outcome Measures

All outcome measures were evaluated at 3 timepoints: preintervention, immediately postintervention, and 3-month follow-up. The primary outcomes focused on pain and disability (key concerns for individuals seeking medical help for neck pain). These outcomes were evaluated using offline scales. Secondary outcomes included kinematic indicators (eg, CROM, proprioception, mean and peak velocity), patient satisfaction, and relief of symptoms, which are all crucial aspects in the rehabilitation process for chronic neck pain. These secondary outcomes were assessed using a combination of web-based VR equipment and offline scales.

Primary Outcomes

NRS

The NRS was used to measure the current neck pain intensity. NRS graded the pain intensity from 0 (no pain) to 10 (worst pain imaginable), with higher scores indicating worse pain. Pain levels were categorized as mild (1-3), moderate (4-6), and severe (≥7) based on the score range [32-34]. The NRS has shown validity and reliability, with a minimum clinically important difference (MCID) of 2.7 established in previous studies [35,36].

Neck Disability Index

The neck disability index was employed as a self-reported questionnaire to measure neck pain–related disability. It consisted of 10 items about activities of daily living, with each item scored from 0 (absence of disability) to 5 (complete disability). The neck disability index is recognized for its validity and reliability, with an MCID of 3.5 points considered significant [37,38].

Secondary Outcomes

CROM

CROM was measured using a VR device in 6 directions: flexion, extension, left and right rotation, and left and right lateral flexion. The results were calculated by taking the average of 3 measuring values. This VR equipment evaluation approach demonstrates high repeatability and sensitivity on these cervical kinematics parameters (ie, CROM, proprioception, mean and peak velocity). The reliability and validity of VR devices to measure CROM have been validated. The minimal detectable change (MDC) of CROM in different directions has been previously reported, while the value changed across the 6 movements ranging from 3.6° to 6.5° [39,40].

Proprioception

Proprioception was defined as the perception of change in direction, position, or speed produced by motor organs (eg, muscles, tendons, joints) in 6 directions. It was calculated as the mean of the relocation difference in 3 tests. Prior studies have reported the psychometric properties of VR equipment evaluating proprioception [4,41] but not provided the MCID.

Mean and Peak Velocity

Mean and peak velocity are crucial indicators reflecting cervical kinematic functions. The mean and peak velocity in 4 directions (flexion, extension, left and right rotation) were obtained by calculating the average values of 3 assessed data on angular velocity during the trial. VR devices have shown good repeatability in measuring cervical motion velocity. Although the MDC for average speed is 14.31°/s, that for maximum speed is 34.95°/s [4,42].

Global Perceived Effect

Global perceived effect is a self-administered questionnaire applied to evaluate patient satisfaction and the relief of symptoms in this study [43]. The satisfaction level ranges from –5 (totally dissatisfied) to 5 (totally satisfied). Similarly, patients could report their relief of symptoms by using the Global Perceived Effect scale, with lower scores representing worse therapeutic effects. These 2 indicators were only measured immediately postintervention and at 3 months after intervention.

Sample Size Calculation

The NRS was chosen as the primary outcome measure in this study. With reference to a previous study [44], the effect size estimate for the NRS was medium (SE 0.25). The correlation between repeated measures was assumed to be 0.5. Three measurements were presumed to be performed (baseline, postintervention, and 3-month follow-up) with a sphericity correction of 0.5. Based on the statistical power of 0.85 and an α level of .05, a total sample size of 50 patients was initially estimated. To account for potential dropout rates that have been observed to exceed 15% in similar studies [27,29,45], a conservative dropout rate of 25% was chosen to ensure sufficient patients for statistical analysis, resulting in a final inclusion of

64 patients. The sample size calculation was conducted using the G*Power software (version 3.1.7; University of Düsseldorf).

Statistical Analysis

Statistical analysis was conducted using the SPSS statistical software (version 25.0; IBM Corp) by a blinded researcher. Data analysis followed the intention-to-treat principle, while the Shapiro-Wilk test was applied to check the normality of various data. Descriptive statistics were used to reflect the different types of results such as mean and standard deviation for the parametric variables and median and quartiles for the nonparametric variables. Group equivalence was assessed via the 2-sided independent-sample *t* test or Pearson chi-squared test by comparing the baseline data between the groups. For most variables (all outcomes except the relief of symptoms), which showed normal distribution, a 2*3 mixed repeated measures analysis of variance (ANOVA) with 1 between-subject factor (treatment) and 1 within-subject factor (time) was performed to compare all variables. Post hoc comparisons were conducted using the Bonferroni test, with *P* values for multiple comparisons adjusted using SPSS software. To address violations of the sphericity assumption, the Greenhouse-Geisser correction was applied. With regard to the analysis of the relief of symptoms, nonparametric statistics were used due to the skewed distribution of data. To account for the dropouts, multiple imputations were used to fill the missing data. To show the effect sizes of observed between- or within-group change, partial eta squared and rank correlation were calculated for the parametric and nonparametric variables, respectively. Based on the previous study [46], effect sizes were classified into small (0.2-0.5), medium (0.5-0.8), and large effect sizes (≥0.8). *P* values less than .05 were indicated to be statistically significant.

Results

Baseline Measures

A total of 120 patients underwent the initial screening for eligibility, of which 56 participants were excluded. Following screening, 64 participants were randomly allocated to either the experimental group or the control group. In the 3-month follow-up period, 3 (5%) participants dropped out of the study due to time conflicts or personal reasons. The flow diagram of participant recruitment and research is shown in Figure 4. The baseline characteristics of the participants in both groups are detailed in Table 1. As shown, there was no between-group difference in age, gender, etiology, disability, pain, and other kinematic indicators. No adverse events were reported during treatment, except some discomfort (eg, complaints of heavy helmets, slightly aggravated pain). No differences existed between the 2 groups over the compliance of the patients continuing neck exercise during the 3-month follow-up period (experimental group 2.31, SD 1.25 vs control group 1.96, SD 1.19; *P*=.22). However, a higher proportion of experimental group participants (16 out of 31) engaged in neck exercises for an average of at least 2 hours per week during the follow-up period compared to control group participants, where only 30% (9/30) achieved this level of compliance, indicating the actual differences between the 2 groups.

Figure 4. Flow diagram showing participants' flow and follow-up evaluation.

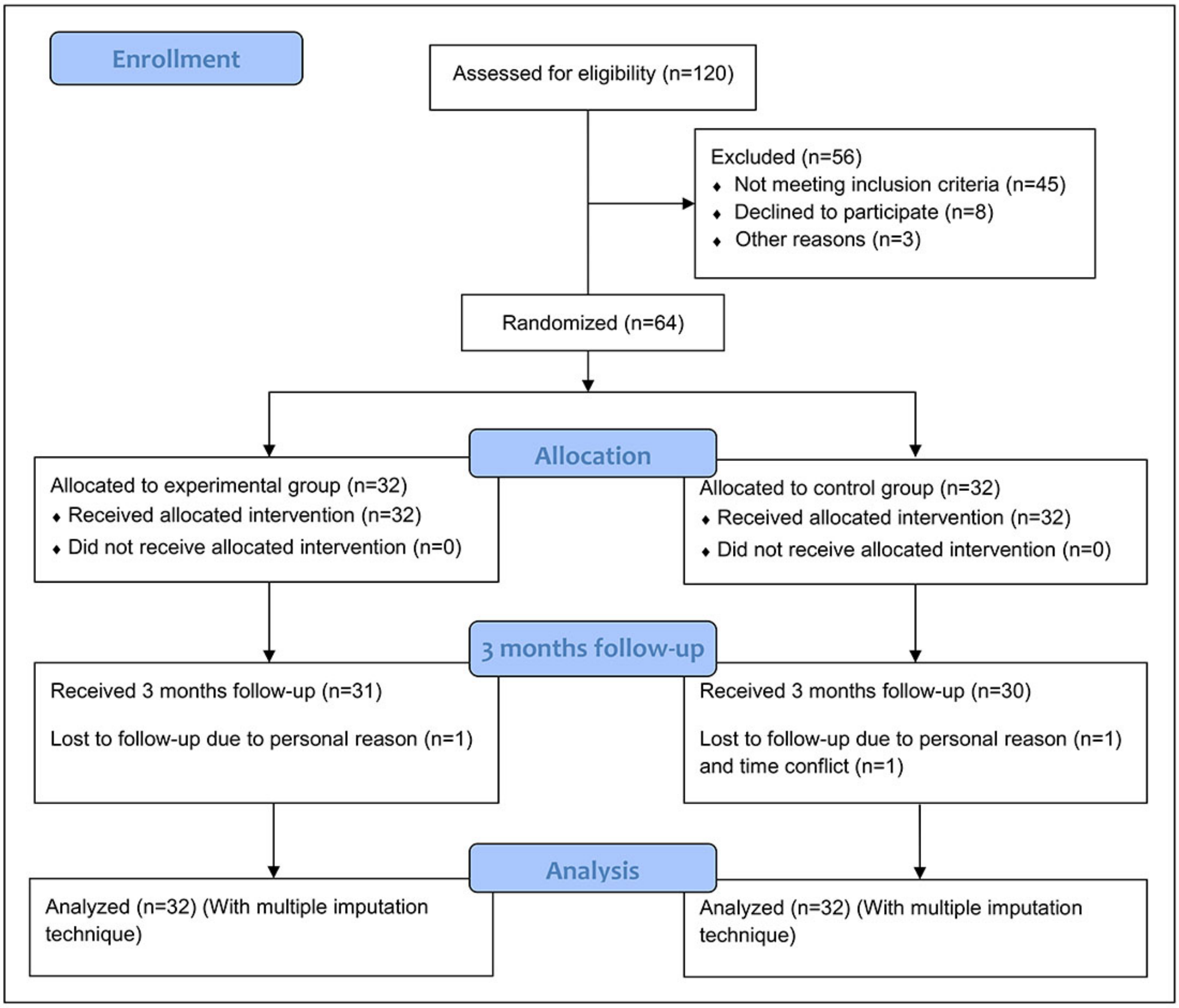


Table 1. Baseline characteristics of the participants in the experimental and the control groups.

Variables	Experimental group (n=32)	Control group (n=32)	*P* value
Age (years), mean (SD)	35.94 (11.02)	40.09 (11.97)	.15[a]
Gender (female), n (%)	25 (78)	22 (69)	.40[b]
Etiology (idiopathic neck pain), n (%)	16 (50)	21 (66)	.43[b]
Disability (Neck Disability Index), mean (SD)	11.34 (2.74)	10.72 (2.32)	.33[a]
Pain intensity (Numeric Rating Scale), mean (SD)	5.38 (1.39)	5.16 (1.42)	.54[a]
Range of motion, mean (SD)			
Flexion	53.96 (10.69)	51.60 (9.99)	.37[a]
Extension	63.55 (8.88)	58.42 (12.85)	.07[a]
Left flexion	37.62 (7.78)	38.42 (8.90)	.71[a]
Right flexion	38.16 (8.67)	40.91 (8.47)	.20[a]
Left rotation	71.92 (9.8)	70.21 (8.96)	.47[a]
Right rotation	71.86 (8.16)	69.43 (9.55)	.28[a]
Proprioception, mean (SD)			
Flexion	2.96 (1.31)	3.33 (1.80)	.36[a]
Extension	3.03 (1.25)	3.16 (1.65)	.73[a]
Left flexion	2.85 (1.45)	3.00 (1.29)	.67[a]
Right flexion	2.73 (1.07)	2.86 (1.57)	.71[a]
Left rotation	1.96 (0.70)	2.53 (1.36)	.08[a]
Right rotation	2.85 (1.35)	2.98 (1.64)	.72[a]
Mean velocity, mean (SD)			
Flexion	12.54 (2.80)	11.02 (2.68)	.07[a]
Extension	14.24 (2.52)	14.90 (2.38)	.28[a]
Left rotation	15.64 (3.47)	15.15 (3.80)	.59[a]
Right rotation	17.27 (2.51)	16.59 (2.58)	.29[a]
Peak velocity, mean (SD)			
Flexion	68.63 (17.18)	76.77 (26.13)	.15[a]
Extension	77.62 (21.63)	77.71 (17.05)	.99[a]
Left rotation	88.71 (18.46)	88.97 (18.72)	.96[a]
Right rotation	94.15 (14.24)	100.66 (22.47)	.17[a]

[a]Independent sample *t* test.
[b]Pearson chi-squared test.

Primary Variables Measure

Neck Disability

As presented in Table 2 and Figure 5, a repeated measures ANOVA showed a main effect of group ($F_{1,3}$=12.738; *P*=.001; η_p^2=0.291), time ($F_{2,62}$=124.140; *P*<.001; η_p^2=0.800), and the group*time interaction ($F_{2,62}$=31.620; *P*<.001; η_p^2=0.505) on neck disability. Compared with those in the control group, participants in the experimental group showed a significant alleviation in neck disability at postintervention (*P*<.001; η_p^2=0.517) and 3-month follow-up (*P*<.001; η_p^2=0.438). Furthermore, therapies in both groups were shown to improve disability in patients with chronic neck pain after intervention or 3-month follow-up in comparison with the baseline (*P*<.01). Further, the extent of disability alleviation in the experimental group exceeded the MCID at both measurement timepoints (5.50 at posttreatment; 5.21 at the 3-month follow-up), while the controls showed a reduction in the disability score by 1.81 and 1.91 points compared to baseline. A higher percentage of experimental group participants experienced disability score reductions exceeding the MCID compared to the control group at both timepoints (experimental group: 29/32, 91% vs control group: 9/32, 28% at posttreatment; experimental group: 25/31, 81% vs control group 6/30, 20% at the 3-month follow-up).

Figure 5. Rehabilitation effect of virtual reality therapy on disability and pain intensity. NDI: neck disability index; NRS: numeric rating scale; *: a statistically significant difference (*P*<0.05) between the two groups at that timepoint (postintervention or 3-month follow-up).

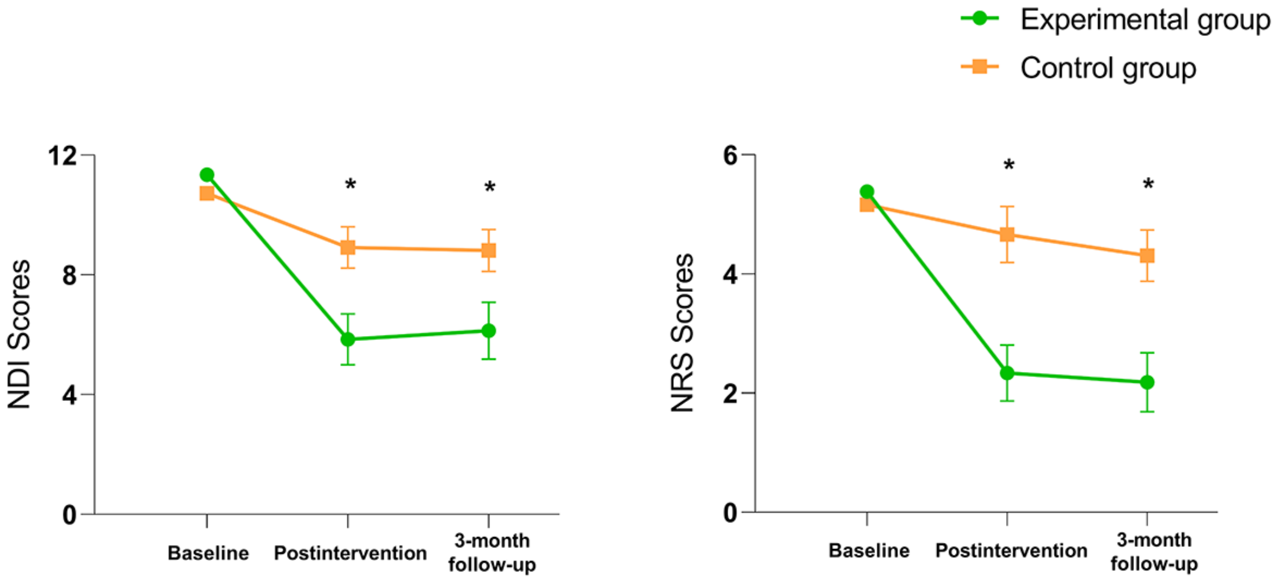


Table 2. Within- and between-group differences in outcome measures.

Variables	Preintervention, mean (SD)		Postintervention, mean (SD)		Cohen d[a]	3-month follow-up, mean (SD)		Cohen d	Group*time, F test (*df*)
	Experimental group	Control group	Experimental group	Control group		Experimental group	Control group		
Disability	11.34 (2.74)	10.72 (2.32)	5.84 (2.38)[b,c]	8.91 (1.92)[c]	0.517	6.13 (2.66)[b,c]	8.81 (1.96)[b]	0.438	31.62 (2, 62)[d]
Pain intensity	5.38 (1.39)	5.16 (1.42)	2.34 (1.31)[b,c]	4.66 (1.31)	0.582	2.18 (1.38)[b,c]	4.31 (1.20)	0.587	27.28 (2, 62)[d]
Range of motion									
Flexion	53.96 (10.69)	51.60 (9.99)	64.78 (9.01)[b,c]	54.25 (8.25)	0.582	67.96 (6.03)[b,c]	53.76 (8.82)	0.680	11.31 (1.592, 49.363)[d]
Extension	63.55 (8.88)	58.42 (12.85)	70.89 (7.07)[b,c]	59.78 (9.93)	0.395	69.09 (8.97)[b,c]	54.29 (8.48)	0.603	8.74 (2, 62)[d]
Left flexion	37.62 (7.78)	38.42 (8.90)	44.23 (7.37)[b,c]	38.11 (8.22)	0.277	47.02 (4.02)[b,c]	35.83 (8.69)	0.578	22.40 (2, 62)[d]
Right flexion	38.16 (8.67)	40.91 (8.47)	45.78 (7.43)[b,c]	40.18 (7.55)	0.192	45.42 (5.73)[b,c]	35.17 (6.50)[b]	0.514	22.48 (2, 62)[d]
Left rotation	71.92 (9.8)	70.21 (8.96)	79.85 (6.47)[b,c]	70.85 (10.09)	0.305	84.92 (5.39)[b,c]	67.86 (11.13)	0.660	21.97 (2, 62)[b]
Right rotation	71.86 (8.16)	69.43 (9.55)	77.19 (7.34)[b,c]	70.62 (8.10)	0.239	82.23 (7.52)[b,c]	65.77 (6.83)	0.682	26.72 (2, 62)[d]
Proprioception									
Flexion	2.96 (1.31)	3.33 (1.80)	2.55 (1.28)	2.71 (1.38)	0.007	1.66 (1.02)	1.55 (0.81)	0.007	0.58 (2, 62)
Extension	3.03 (1.25)	3.16 (1.65)	2.43 (0.86)	2.67 (1.03)	0.035	2.53 (1.40)	2.24 (0.70)	0.026	0.96 (2, 62)
Left flexion	2.85 (1.45)	3.00 (1.29)	2.57 (1.13)	2.52 (1.38)	0.001	2.02 (1.28)	2.41 (1.02)	0.064	0.62 (2, 62)
Right flexion	2.73 (1.07)	2.86 (1.57)	2.28 (0.82)	2.98 (1.12)	0.182	2.56 (1.36)	2.66 (1.08)	0.003	1.33 (2, 62)
Left rotation	1.96 (0.70)	2.53 (1.36)	2.61 (1.38)	2.21 (0.92)	0.053	2.04 (1.31)	2.79 (0.97)	0.209	6.28 (1.687, 52.289)[d]
Right rotation	2.85 (1.35)	2.98 (1.64)	2.36 (1.19)	2.29 (1.43)	0.002	1.81 (0.95)	2.47 (1.45)	0.120	1.25 (2, 62)
Mean velocity									
Flexion	12.54 (2.80)	11.02 (2.68)	15.71 (2.74)	12.93 (3.12)	0.270	14.30 (2.59)	12.58 (2.27)	0.218	1.28 (2, 62)
Extension	14.24 (2.52)	14.90 (2.38)	14.85 (2.00)	13.68 (2.97)	0.085	15.22 (1.66)[b,c]	13.14 (2.39)	0.337	6.53 (2, 62)[d]
Left rotation	15.64 (3.47)	15.15 (3.80)	18.13 (3.76)[b]	16.44 (3.28)	0.105	20.27 (3.46)[b,c]	16.99 (2.13)[b]	0.452	3.80 (1.609, 19.872)[e]
Right rotation	17.27 (2.51)	16.59 (2.58)	18.75 (2.12)	17.56 (1.93)	0.156	17.50 (2.03)	16.63 (2.02)	0.079	0.24 (2, 62)
Peak velocity									
Flexion	68.63 (17.18)	76.77 (26.13)	81.74 (20.69)[b]	72.61 (21.01)	0.092	117.76 (31.10)[b,c]	91.55 (21.72)[b]	0.376	14.69 (2, 62)[d]
Extension	77.62 (21.63)	77.71 (17.05)	76.17 (16.90)	71.71 (18.45)	0.049	99.55 (25.10)[b,c]	70.74 (14.55)	0.530	13.31 (2, 62)[d]
Left rotation	88.71 (18.46)	88.97 (18.72)	93.97 (12.52)	98.71 (27.52)	0.022	173.00 (51.51)[b,c]	132.43 (29.69)[b]	0.365	13.75 (1.522, 47.174)[d]

Variables	Preintervention, mean (SD)		Postintervention, mean (SD)		Cohen d[a]	3-month follow-up, mean (SD)		Cohen d	Group*time, F test (df)
	Experimental group	Control group	Experimental group	Control group		Experimental group	Control group		
Right rotation	94.15 (14.24)	100.66 (22.47)	102.32 (24.37)	89.02 (13.23)[b]	0.167	124.33 (29.63)[b,c]	93.26 (10.63)	0.534	13.29 (2, 62)[d]
Satisfaction	N/A[f]	N/A	3.03 (1.33)	1.66 (1.82)	0.860	2.97 (1.28)	1.81 (1.99)	0.693	N/A
Relief of symptoms	N/A	N/A	3.00 (2.00)	2.00 (3.00)	N/A	4.00 (1.00)	2.00 (2.00)	N/A	N/A

[a]Cohen d was calculated for the differences between postintervention or 3-month follow-up and preintervention in the experimental group compared to the control group.

[b]Indicates P<.02 (0.05/3) for within-group comparisons by Bonferroni correction compared to the baseline.

[c]P<.05 statistically significant differences were found compared to the control group at the same measuring timepoint.

[d]P<0.01 significant main effects were revealed on the group*time interaction.

[e]P<.05 significant main effects were revealed on the group*time interaction.

[f]N/A: not applicable.

Neck Pain Intensity

For pain, ANOVA results revealed significant differences over time ($F_{1.744,54.077}$=87.369; P<.001; η_p^2=0.738), group ($F_{1,31}$=28.138; P<.001; η_p^2=0.476), and the group*time interaction ($F_{2,62}$=27.277; P<.001; η_p^2=0.468). The post hoc analysis indicated a significant between-group difference at postintervention (P<.001; η_p^2=0.582) and 3 months postintervention (P<.001; η_p^2=0.587), with the experimental group representing better enhancement. Compared with baseline, patients in the experimental group experienced pain relief immediately postintervention and at 3-month follow-up, while control group participants did not exhibit significant pain reduction throughout the study. Besides, pain intensity scores decreased in both groups compared to baseline (experimental group 3.04 vs control group 0.50 at posttreatment; experimental group 3.20 vs control group 0.85 at the 3-month follow-up), with patients in the experimental group exceeding the MCID at 2 timepoints. The percentage of data exceeding the MCID significantly differed between the 2 groups (experimental group: 21/32, 66% vs control group: 3/32, 9% at posttreatment; experimental group: 20/31, 65% vs control group 4/30, 13% at the 3-month follow-up).

Secondary Variables Measure

CROM

The results of ANOVA on CROM revealed a significant effect of the group, time, and group*time interaction (P<.05). Participants in the experimental group obtained greater ROM improvement in 6 directions at postintervention and at 3-month follow-up (P<.05) compared to the control group participants. Notably, significant changes were observed after intervention and follow-up in the experimental group from those at baseline (P<.05), while no differences were observed in the control group. The experimental group participants achieved ROM improvements exceeding the MDC in all directions at both timepoints, except for extension ROM at the 3-month follow-up, highlighting the clinical effectiveness of the intervention in the experimental group compared to the control group. Specific data on this indicator can be found in Multimedia Appendix 2.

Proprioception

Regarding proprioception, the ANOVA results revealed no significant difference in the group*time interaction for 6 directions, except for left rotation. For the proprioception of left rotation, significant effects occurred in the group*time interaction. The post hoc analysis showed that patients receiving VR intervention attained lesser improvement after the 3-month follow-up than the control group. However, no within-group differences were reported. Upon further analysis of other directions, proprioception in flexion, extension, and right rotation directions was found to achieve improvement in both groups after treatment and follow-up versus the baseline, and proprioception of left flexion showed a noticeable improvement after the follow-up in comparison with the baseline. However, the between-group analysis showed no marked difference in all directions. The detailed data regarding this parameter are presented in Multimedia Appendix 2.

Mean and Peak Velocity

There were significant main effects for the interaction between time and group (P<.05) for the mean velocity of extension and left rotation. The post hoc analysis revealed significant gains after the 3-month follow-up in the experimental group compared to that in the control group or baseline for these 2 directions. No significant effects for the interaction between time and group were found in the mean velocity of flexion and right rotation. For intergroup comparisons, patients receiving VR training showed better improvement in the mean velocity of flexion than the control group, which was not found in the right rotation direction. As for the within-group comparison, both groups showed superiority over baseline in the mean velocity of flexion at posttreatment and the 3-month follow-up, and the mean velocity of right rotation remained negative. However, the magnitude of improvement in the mean velocity in both the groups did not surpass the corresponding MDC in any direction.

A repeated measures ANOVA showed a main effect for the group*time interaction for the peak velocity in all 4 directions. Compared with the baseline, the experimental group participants gained significant enhancement (*P*<.05) at the 3-month follow-up. Furthermore, the between-group comparisons supported better therapeutic effects with VR devices than the control after intervention and 3-month follow-up period (Table 2). Increased maximal velocity of flexion and left rotation directions was significantly higher in the experimental group over MDC after the follow-up, which did not occur in the other directions or the control group. For specific data on the mean and peak velocity, please refer to Multimedia Appendix 2.

Global Perceived Effect

Considering patient satisfaction, the results revealed a significant between-group difference at postintervention and 3-month follow-up (experimental group 3.03, SD 1.33 vs control group 1.66, SD 1.82) with an advantage to the experimental group. The within-group analysis showed no significant effects among different timepoints for the 2 groups. In the Mann-Whitney *U* test, obvious between-group differences were found in the relief of symptoms at postintervention (experimental group 3.00, SD 2.00 vs control group 2.00, SD 3.00) and 3-month follow-up (experimental group 4.00, SD 1.00 vs control group 2.00, SD 2.00). Furthermore, significant improvements were observed in the experimental group after the 3-month follow-up compared to those at postintervention, while no differences were observed in the control group (Table 2). The specific data related to these indicators can be obtained in Multimedia Appendix 2.

Discussion

Overview

This RCT was intended to compare the benefits of combining VR therapy and conventional rehabilitation with those of conventional rehabilitation alone for treating chronic neck pain. Overall, our results show that patients in both groups reported reduced pain and disability and demonstrated improved kinematic functions. Direct comparisons between the 2 groups revealed that VR treatment in addition to conventional rehabilitation was superior to conventional rehabilitation alone for improvement in the pain, disability, and kinematic indicators, and the effects of combined therapy could be maintained over the 3-month follow-up period. Additionally, participants in the VR therapy group reported higher satisfaction levels, better symptom improvement, and greater willingness to engage in exercises during the follow-up period.

Effects of VR Therapy Combined With Conventional Rehabilitation on Pain and Disability

Although reduced pain and disability were found in both treatment groups at 2 timepoints, these indicators were decreased nearly 3 times more in the experimental group than in the control group. Furthermore, the improvement of pain and disability observed at 2 measuring timepoints in the experimental group was higher than the MCID, which has been previously reported as 2.7 and 3.5 points, respectively [37,38]. This finding indicates the significant and clinical effectiveness of VR therapy in addition to conventional rehabilitation in alleviating pain and disability. The corresponding size effects were medium, highlighting the notable differences between the 2 groups.

Multiple studies have shown that patients with neck pain experienced significant improvements in pain intensity with VR treatment compared to baseline [45,47] and markedly superior to control groups receiving laser training [27] or conventional rehabilitation [28]. Further, a recent meta-analysis [48] incorporating 8 RCTs revealed that better analgesic effects were found in the multimodal intervention (VR technique in combination with other therapies) than in the other interventions and in the patients treated in the clinic or research unit than the controls. This also provides a new perspective on VR analgesia research. However, some studies have reported conflicting results. For instance, a study [44] investigating the efficacy of a 120-minute VR therapy session for patients with chronic neck pain indicated remarkable improvement in pain intensity at rest or during motion compared to baseline as well as alleviation in the disability level. However, no significant between-group differences were observed in these metrics in the VR intervention group as compared to the 2 control groups undergoing conventional rehabilitation alone or general sensorimotor training plus conventional rehabilitation. This discrepancy may be attributed to the smaller sample size in that study (17 individuals per group) [44], which lowered the statistical power representation of between-group differences. Similarly, the VR gaming scenario utilized in that study [44] lacked sufficient visual and auditory feedback compared to the VR design in our research, and this might have limited the analgesic effect of VR treatment.

The potential efficacy of VR therapy in reducing neck pain and disability may be attributed to its ability to enhance coordination between the deep and superficial cervical muscles [49]. Poor sensorimotor control by cervical muscles in patients with neck pain has been indicated in previous research [5,9] and is considered to trigger associated disability and kinematic disorders. Although muscle activation was not evaluated in this study, VR therapy appears to promote the function and coordination of cervical muscles, thereby reducing the stress on cervical segments and alleviating neck pain and disability. Another possible reason could be the deep engagement required by the virtual environment, blocking the transmission of sensory information related to pain and achieving analgesic effects.

Effects of VR Therapy Combined With Conventional Rehabilitation on Cervical Kinematic Function

The secondary outcomes yielded interesting findings that VR therapy could increase the ROM, mean velocity, and peak velocity at 2 timepoints compared to those in the baseline or control group. This conclusion was consistent with that reported in previous research [28,44,50]. Tejera et al [50] in 2021 reported the positive results of VR therapy on increasing CROM in patients with chronic neck pain, which can be attributed to the sufficient feedback provided by VR devices. The visualization of images was widely perceived as useful in activating the corticospinal system and enhancing the intensity of muscle recruitment, thereby improving the overall neck kinematic functions. Fowler et al [51] showed that VR might encourage patients to turn their heads farther and faster by its effect on reducing fear of movement, which has been reported in other studies [52-54]. Besides these, continuous progressive VR treatment dosage based on real-time assessment data on motor function assisted patients in restoring their motor function.

As can be observed from the mean and maximum velocity data in various directions, the experimental group always showed no between-group differences after training but showed between-group differences after the 3 months follow-up compared to the control group, suggesting that this may be attributed to the insufficient training time during the intervention. Upon analyzing the training length after the intervention, we could see that the training frequency of patients in the experimental group was higher than that of the control group during the follow-up period, reaching an average of 1-2 hours of training per week, and more training time outside of the experiment would probably promote further improvement of motor function. These findings indicate that researchers as well as clinical specialists should pay more attention to the supervision and education of home-based active exercise in the future.

Regarding proprioception, both groups showed significant improvement in several directions after treatment or 3-month follow-up; however, no between-group differences were found. Prior studies [49,55] have confirmed that multiple exercise programs, including head relocation practice, gaze stability, eye-follow, and eye/head coordination, are effective in improving proprioception. In this study, however, only head relocation practice was used (the participant was instructed to memorize the head-neck position and try to find the initial position with eyes closed after moving), and satisfactory results were obtained. Moreover, other studies [27,44] have utilized alternative proprioceptive training with similarly favorable outcomes. Some investigators noticed that eye-follow and eye/head coordination training greatly enhanced patients' accuracy, which was likely attributed to improved motor control and coordination of the neck [47]. This suggests that more consideration should be given to focus on all forms of proprioceptive training in the clinical management of patients with chronic neck pain.

Effects of VR Therapy Combined With Conventional Rehabilitation on Satisfaction and Relief of Symptoms

Besides the indicators mentioned above, the marked between-group difference was observed in patient satisfaction and relief of symptoms at both timepoints, with some advantages in the combined treatment. These 2 self-rating indicators are considered important for recovering from chronic neck pain. The enjoyment derived from VR equipment, multiple visual and auditory feedback, personalized tasks, and adjustable difficulty levels likely contributed to the higher satisfaction levels and greater therapeutic efficacy in the experimental group. Notably, no adverse events such as motion sickness were reported during the research period.

Limitations

Several limitations of this study warrant consideration. The absence of a placebo group receiving sham VR therapy raises concerns about the potential overestimation of VR therapy's therapeutic effects. However, the substantial improvements in pain, disability, and CROM surpassing the MCID suggest that the effect may be due to the treatment itself other than the placebo effect. Moreover, the impact of varying durations of conventional rehabilitation (30 minutes vs 10 minutes) on therapeutic outcomes remains uncertain, potentially influencing the perceived efficacy of VR treatment. Additionally, the inability to blind patients due to the experimental nature of this study introduces a risk of bias. The lack of long-term follow-up data further limits the generalizability of the findings. Lastly, the absence of assessment indicators for mental function and quality of life hinders the comprehensive evaluation of VR therapy's overall therapeutic efficacy.

Implications

This study provides support for the effectiveness of a combined approach involving VR therapy and conventional rehabilitation in managing chronic neck pain. However, uncertainties persist regarding the optimal dosage, underlying mechanisms of VR therapy, and the comparative effectiveness of different VR equipment types (eg, semi-immersive, nonimmersive). Future investigations should design specific trials to address these knowledge gaps. Furthermore, exploring the synergistic benefits of integrating VR training with other evidence-based interventions such as manipulation and sensorimotor training is warranted.

Conclusions

In conclusion, the integration of VR intervention with conventional rehabilitation demonstrates significant improvements in pain, disability, and kinematic function among patients with chronic neck pain at both postintervention and 3-month follow-up assessments. Although patients can benefit from conventional rehabilitation alone, the combination of VR therapy and conventional rehabilitation is more effective for improvement in the abovementioned indicators. Considering the higher satisfaction as well as greater training initiative in the experimental group and the absence of adverse events, this feasible and effective intervention could be integrated into the standard rehabilitation treatment plan for patients with chronic neck pain. Future research endeavors should focus on refining therapeutic regimens, determining optimal dosages for VR therapy, and streamlining the implementation of this intervention in clinical settings to enhance convenience and efficacy.

Acknowledgments

The authors would like to thank all of the patients who took part in this study. This work was supported by the Key Research Project of the Science and Technology Department of Sichuan Province (2021YFS0069).

Authors' Contributions

Q Guo, LZ, LLH, CG, and Q Gao contributed to the conception and design, acquisition of data, drafting of the paper, and critical revision of important intellectual content. GC, CL, and WW were responsible for the analysis and interpretation of the data. All authors discussed the results, commented on the manuscript, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth) checklist (v 1.6.1).
[PDF File (Adobe PDF File), 453 KB - games_v12i1e42829_app1.pdf]

Multimedia Appendix 2
The main and simple effect data sets.
[DOCX File , 18 KB - games_v12i1e42829_app2.docx]

References

1. Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. Mayo Clin Proc 2015 Feb;90(2):284-299. [doi: 10.1016/j.mayocp.2014.09.008] [Medline: 25659245]
2. Martimbianco ALC, Porfírio GJM, Pacheco RL, Torloni MR, Riera R. Transcutaneous electrical nerve stimulation (TENS) for chronic neck pain. Cochrane Database Syst Rev 2019 Dec 12;12(12):CD011927 [FREE Full text] [doi: 10.1002/14651858.CD011927.pub2] [Medline: 31830313]
3. Stanton TR, Leake HB, Chalmers KJ, Moseley GL. Evidence of impaired proprioception in chronic, idiopathic neck pain: systematic review and meta-analysis. Phys Ther 2016 Jun;96(6):876-887 [FREE Full text] [doi: 10.2522/ptj.20150241] [Medline: 26472296]
4. Sarig Bahat H, Chen X, Reznik D, Kodesh E, Treleaven J. Interactive cervical motion kinematics: sensitivity, specificity and clinically significant values for identifying kinematic impairments in patients with chronic neck pain. Man Ther 2015 Apr;20(2):295-302. [doi: 10.1016/j.math.2014.10.002] [Medline: 25456272]
5. Röijezon U, Björklund M, Bergenheim M, Djupsjöbacka M. A novel method for neck coordination exercise--a pilot study on persons with chronic non-specific neck pain. J Neuroeng Rehabil 2008 Dec 23;5:36 [FREE Full text] [doi: 10.1186/1743-0003-5-36] [Medline: 19105826]
6. Rudolfsson T, Björklund M, Djupsjöbacka M. Range of motion in the upper and lower cervical spine in people with chronic neck pain. Man Ther 2012 Feb;17(1):53-59. [doi: 10.1016/j.math.2011.08.007] [Medline: 21945295]
7. Lee H, Nicholson LL, Adams RD. Neck muscle endurance, self-report, and range of motion data from subjects with treated and untreated neck pain. J Manipulative Physiol Ther 2005 Jan;28(1):25-32. [doi: 10.1016/j.jmpt.2004.12.005] [Medline: 15726032]
8. Beltran-Alacreu H, López-de-Uralde-Villanueva I, Calvo-Lobo C, Fernández-Carnero J, La Touche R. Clinical features of patients with chronic non-specific neck pain per disability level: A novel observational study. Rev Assoc Med Bras (1992) 2018 Aug;64(8):700-709 [FREE Full text] [doi: 10.1590/1806-9282.64.08.700] [Medline: 30673040]
9. Harel NY, Song K, Tang X, Strittmatter SM. Nogo receptor deletion and multimodal exercise improve distinct aspects of recovery in cervical spinal cord injury. J Neurotrauma 2010 Nov;27(11):2055-2066 [FREE Full text] [doi: 10.1089/neu.2010.1491] [Medline: 20809785]
10. Takasaki H, Treleaven J, Johnston V, Jull G. Contributions of physical and cognitive impairments to self-reported driving difficulty in chronic whiplash-associated disorders. Spine (Phila Pa 1976) 2013 Aug 15;38(18):1554-1560. [doi: 10.1097/BRS.0b013e31829adb54] [Medline: 23698571]
11. Blanpied PR, Gross AR, Elliott JM, Devaney LL, Clewley D, Walton DM, et al. Neck Pain: Revision 2017. J Orthop Sports Phys Ther 2017 Jul;47(7):A1-A83. [doi: 10.2519/jospt.2017.0302] [Medline: 28666405]
12. Wang X, Wang Y, Zhang Z. Exercise therapy for neck pain: Chinese expert consensus. Journal of Shanghai University of Sport 2020 Mar 19;44(01):59-69. [doi: 10.16099/j.sus.2020.01.007]
13. Sarig Bahat H, Weiss PL, Laufer Y. The effect of neck pain on cervical kinematics, as assessed in a virtual environment. Arch Phys Med Rehabil 2010 Dec;91(12):1884-1890. [doi: 10.1016/j.apmr.2010.09.007] [Medline: 21112430]
14. Iskander M, Ogunsola T, Ramachandran R, McGowan R, Al-Aswad LA. Virtual reality and augmented reality in ophthalmology: a contemporary prospective. Asia Pac J Ophthalmol (Phila) 2021;10(3):244-252 [FREE Full text] [doi: 10.1097/APO.0000000000000409] [Medline: 34383716]
15. Lopez-de-Uralde-Villanueva I, Beltran-Alacreu H, Fernandez-Carnero J, Kindelan-Calvo P, La Touche R. Widespread pressure pain hyperalgesia in chronic nonspecific neck pain with neuropathic features: a descriptive cross-sectional study. Pain Physician 2016 Feb;19(2):77-88 [FREE Full text] [Medline: 26815252]

16. Muñoz-García D, López-de-Uralde-Villanueva I, Beltrán-Alacreu H, La Touche R, Fernández-Carnero J. Patients with concomitant chronic neck pain and myofascial pain in masticatory muscles have more widespread pain and distal hyperalgesia than patients with only chronic neck pain. Pain Med 2017 Mar 01;18(3):526-537. [doi: 10.1093/pm/pnw274] [Medline: 28034980]
17. Cipresso P, Giglioli IAC, Raya MA, Riva G. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. Front Psychol 2018;9:2086 [FREE Full text] [doi: 10.3389/fpsyg.2018.02086] [Medline: 30459681]
18. Brown A, Green T. Virtua reality: low-cost tools and resources for the classroom. TechTrends 2016 Jun 30;60(5):517-519. [doi: 10.1007/s11528-016-0102-z]
19. Pawassar CM, Tiberius V. Virtual reality in health care: bibliometric analysis. JMIR Serious Games 2021 Dec 01;9(4):e32721 [FREE Full text] [doi: 10.2196/32721] [Medline: 34855606]
20. Wu J, Zeng A, Chen Z, Wei Y, Huang K, Chen J, et al. Effects of virtual reality training on upper limb function and balance in stroke patients: systematic review and meta-meta-analysis. J Med Internet Res 2021 Oct 12;23(10):e31051 [FREE Full text] [doi: 10.2196/31051] [Medline: 34636735]
21. Gumaa M, Rehan Youssef A. Is virtual reality effective in orthopedic rehabilitation? a systematic review and meta-analysis. Phys Ther 2019 Oct 28;99(10):1304-1325. [doi: 10.1093/ptj/pzz093] [Medline: 31343702]
22. Gray ML, Goldrich DY, McKee S, Schaberg M, Del Signore A, Govindaraj S, et al. Virtual reality as distraction analgesia for office-based procedures: a randomized crossover-controlled trial. Otolaryngol Head Neck Surg 2021 Mar;164(3):580-588. [doi: 10.1177/0194599820942215] [Medline: 32746734]
23. Lin D, Lin Y, Chai H, Han Y, Jan M. Comparison of proprioceptive functions between computerized proprioception facilitation exercise and closed kinetic chain exercise in patients with knee osteoarthritis. Clin Rheumatol 2007 Apr;26(4):520-528. [doi: 10.1007/s10067-006-0324-0] [Medline: 16786252]
24. Pekyavas NO, Ergun N. Comparison of virtual reality exergaming and home exercise programs in patients with subacromial impingement syndrome and scapular dyskinesis: Short term effect. Acta Orthop Traumatol Turc 2017 May;51(3):238-242 [FREE Full text] [doi: 10.1016/j.aott.2017.03.008] [Medline: 28446376]
25. Matheve T, Bogaerts K, Timmermans A. Virtual reality distraction induces hypoalgesia in patients with chronic low back pain: a randomized controlled trial. J Neuroeng Rehabil 2020 Apr 22;17(1):55 [FREE Full text] [doi: 10.1186/s12984-020-00688-0] [Medline: 32321516]
26. Sarig Bahat H, Takasaki H, Chen X, Bet-Or Y, Treleaven J. Cervical kinematic training with and without interactive VR training for chronic neck pain - a randomized clinical trial. Man Ther 2015 Feb;20(1):68-78. [doi: 10.1016/j.math.2014.06.008] [Medline: 25066503]
27. Sarig Bahat H, Croft K, Carter C, Hoddinott A, Sprecher E, Treleaven J. Remote kinematic training for patients with chronic neck pain: a randomised controlled trial. Eur Spine J 2018 Jun;27(6):1309-1323. [doi: 10.1007/s00586-017-5323-0] [Medline: 29018956]
28. Mukherjee M, Bedekar N, Sancheti P, Shyam A. Immediate and short-term effect of virtual reality training on pain, range of motion, and kinesiophobia in patients with cervical spondylosis. Indian J Phys Ther Res 2020;2(1):55. [doi: 10.4103/ijptr.ijptr_73_19]
29. Bahat HS, German D, Palomo G, Gold H, Nir YF. Self-kinematic training for flight-associated neck pain: a randomized controlled trial. Aerosp Med Hum Perform 2020 Oct 01;91(10):790-797. [doi: 10.3357/AMHP.5546.2020] [Medline: 33187565]
30. Ahern MM, Dean LV, Stoddard CC, Agrawal A, Kim K, Cook CE, et al. The effectiveness of virtual reality in patients with spinal pain: a systematic review and meta-analysis. Pain Pract 2020 Jul;20(6):656-675. [doi: 10.1111/papr.12885] [Medline: 32196892]
31. Zronek M, Sanker H, Newcomb J, Donaldson M. The influence of home exercise programs for patients with non-specific or specific neck pain: a systematic review of the literature. J Man Manip Ther 2016 May;24(2):62-73 [FREE Full text] [doi: 10.1179/2042618613Y.0000000047] [Medline: 27559275]
32. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005 Aug;14(7):798-804. [doi: 10.1111/j.1365-2702.2005.01121.x] [Medline: 16000093]
33. Wan L, Zhao Q, Chen J. Expert consensus on the application of pain evaluation questionnaires in China(2020). Chinese Journal Of Painology 2020 Jun 25;16(3):177-187. [doi: 10.3760/cma.j.cn101379-20190915-00075]
34. No authors listed. Pain: clinical manual for nursing practice. Nurs Stand 1994 Dec 07;9(11):55. [doi: 10.7748/ns.9.11.55.s69] [Medline: 27527475]
35. Juul T, Søgaard K, Davis AM, Roos EM. Psychometric properties of the Neck OutcOme Score, Neck Disability Index, and Short Form-36 were evaluated in patients with neck pain. J Clin Epidemiol 2016 Nov;79:31-40. [doi: 10.1016/j.jclinepi.2016.03.015] [Medline: 27038852]
36. Young IA, Dunning J, Butts R, Mourad F, Cleland JA. Reliability, construct validity, and responsiveness of the neck disability index and numeric pain rating scale in patients with mechanical neck pain without upper extremity symptoms. Physiother Theory Pract 2019 Dec;35(12):1328-1335. [doi: 10.1080/09593985.2018.1471763] [Medline: 29856244]

37. Cleland JA, Childs JD, Fritz JM, Whitman JM. Interrater reliability of the history and physical examination in patients with mechanical neck pain. Arch Phys Med Rehabil 2006 Oct;87(10):1388-1395. [doi: 10.1016/j.apmr.2006.06.011] [Medline: 17023251]
38. Pool JJM, Ostelo RWJG, Hoving JL, Bouter LM, de Vet HCW. Minimal clinically important change of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. Spine (Phila Pa 1976) 2007 Dec 15;32(26):3047-3051. [doi: 10.1097/BRS.0b013e31815cf75b] [Medline: 18091500]
39. Sarig-Bahat H, Weiss PL, Laufer Y. Cervical motion assessment using virtual reality. Spine (Phila Pa 1976) 2009 May 01;34(10):1018-1024. [doi: 10.1097/BRS.0b013e31819b3254] [Medline: 19404177]
40. Audette I, Dumas J, Côté JN, De Serres SJ. Validity and between-day reliability of the cervical range of motion (CROM) device. J Orthop Sports Phys Ther 2010 May;40(5):318-323. [doi: 10.2519/jospt.2010.3180] [Medline: 20436238]
41. Sarig Bahat H, Watt P, Rhodes M, Hadar D, Treleaven J. High-vs. low-tech cervical movement sense measurement in individuals with neck pain. Musculoskelet Sci Pract 2020 Feb;45:102097. [doi: 10.1016/j.msksp.2019.102097] [Medline: 32056822]
42. Sarig Bahat H, Sprecher E, Sela I, Treleaven J. Neck motion kinematics: an inter-tester reliability study using an interactive neck VR assessment in asymptomatic individuals. Eur Spine J 2016 Jul;25(7):2139-2148. [doi: 10.1007/s00586-016-4388-5] [Medline: 26831536]
43. Evans R, Bronfort G, Maiers M, Schulz C, Hartvigsen J. "I know it's changed": a mixed-methods study of the meaning of Global Perceived Effect in chronic neck pain patients. Eur Spine J 2014 Apr;23(4):888-897 [FREE Full text] [doi: 10.1007/s00586-013-3149-y] [Medline: 24408636]
44. Nusser M, Knapp S, Kramer M, Krischak G. Effects of virtual reality-based neck-specific sensorimotor training in patients with chronic neck pain: A randomized controlled pilot trial. J Rehabil Med 2021 Feb 10;53(2):jrm00151 [FREE Full text] [doi: 10.2340/16501977-2786] [Medline: 33369684]
45. Cetin H, Kose N, Oge HK. Virtual reality and motor control exercises to treat chronic neck pain: A randomized controlled trial. Musculoskelet Sci Pract 2022 Dec;62:102636. [doi: 10.1016/j.msksp.2022.102636] [Medline: 35952621]
46. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Routledge; Jun 30, 1988.
47. Rezaei I, Razeghi M, Ebrahimi S, et al. A novel virtual reality technique (Cervigame®) compared to conventional proprioceptive training to treat neck pain: a randomized controlled trial. J Biomed Phys Eng 2019 Jun;9(3):355-366 [FREE Full text] [doi: 10.31661/jbpe.v0i0.556] [Medline: 31341881]
48. Guo Q, Zhang L, Gui C, Chen G, Chen Y, Tan H, et al. Virtual reality intervention for patients with neck pain: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023 Apr 03;25:e38256 [FREE Full text] [doi: 10.2196/38256] [Medline: 37010891]
49. Jull G, Falla D, Treleaven J, Hodges P, Vicenzino B. Retraining cervical joint position sense: the effect of two exercise regimes. J Orthop Res 2007 Mar;25(3):404-412 [FREE Full text] [doi: 10.1002/jor.20220] [Medline: 17143898]
50. Tejera DM, Beltran-Alacreu H, Cano-de-la-Cuerda R, Leon Hernández JV, Martín-Pintado-Zugasti A, Calvo-Lobo C, et al. Effects of virtual reality versus exercise on pain, functional, somatosensory and psychosocial outcomes in patients with non-specific chronic neck pain: a randomized clinical trial. Int J Environ Res Public Health 2020 Aug 16;17(16):5950 [FREE Full text] [doi: 10.3390/ijerph17165950] [Medline: 32824394]
51. Fowler CA, Ballistrea LM, Mazzone KE, Martin AM, Kaplan H, Kip KE, et al. A virtual reality intervention for fear of movement for veterans with chronic pain: protocol for a feasibility study. Pilot Feasibility Stud 2019;5:146 [FREE Full text] [doi: 10.1186/s40814-019-0501-y] [Medline: 31890259]
52. López-de-Uralde-Villanueva I, Notario-Pérez R, Del Corral T, Ramos-Díaz B, Acuyo-Osorio M, La Touche R. Functional limitations and associated psychological factors in military personnel with chronic nonspecific neck pain with higher levels of kinesiophobia. Work 2017;58(3):287-297. [doi: 10.3233/WOR-172634] [Medline: 29154308]
53. Harvie DS, Smith RT, Moseley GL, Meulders A, Michiels B, Sterling M. Illusion-enhanced virtual reality exercise for neck pain: a replicated single case series. Clin J Pain 2020 Feb;36(2):101-109. [doi: 10.1097/AJP.0000000000000780] [Medline: 31714324]
54. Yilmaz Yelvar GD, Çırak Y, Dalkılınç M, Parlak Demir Y, Guner Z, Boydak A. Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial. Eur Spine J 2017 Feb;26(2):538-545. [doi: 10.1007/s00586-016-4892-7] [Medline: 27981455]
55. Revel M, Minguet M, Gregoy P, Vaillant J, Manuel JL. Changes in cervicocephalic kinesthesia after a proprioceptive rehabilitation program in patients with neck pain: a randomized controlled study. Arch Phys Med Rehabil 1994 Aug;75(8):895-899. [doi: 10.1016/0003-9993(94)90115-5] [Medline: 8053797]

Abbreviations

ANOVA: analysis of variance
CROM: cervical range of motion
MCID: minimum clinically important difference
MDC: minimal detectable change

NRS: numeric rating scale
RCT: randomized controlled trial
ROM: range of motion
VR: virtual reality

Edited by G Eysenbach; submitted 20.09.22; peer-reviewed by H Liu, CQ He, C Fowler; comments to author 14.04.23; revised version received 16.10.23; accepted 17.03.24; published 24.04.24.

Please cite as:
Guo Q, Zhang L, Han LL, Gui C, Chen G, Ling C, Wang W, Gao Q
Effects of Virtual Reality Therapy Combined With Conventional Rehabilitation on Pain, Kinematic Function, and Disability in Patients With Chronic Neck Pain: Randomized Controlled Trial
JMIR Serious Games 2024;12:e42829
URL: https://games.jmir.org/2024/1/e42829
doi:10.2196/42829
PMID:38656775

Original Paper

Gamification of Behavior Change: Mathematical Principle and Proof-of-Concept Study

Falk Lieder[1,2*], PhD; Pin-Zhen Chen[2*], BSc; Mike Prentice[2], PhD; Victoria Amo[2], BSc, MSc; Mateo Tošić[2], BSc

[1]Department of Psychology, University of California, Los Angeles, Los Angeles, CA, United States

[2]Max Planck Institute for Intelligent Systems, Tübingen, Germany

[*]these authors contributed equally

Corresponding Author:
Falk Lieder, PhD
Department of Psychology
University of California, Los Angeles
502 Portola Plaza, Los Angeles
Los Angeles, CA, 90095
United States
Phone: 1 424 259 5300
Email: falk.lieder@psych.ucla.edu

Abstract

Background: Many people want to build good habits to become healthier, live longer, or become happier but struggle to change their behavior. Gamification can make behavior change easier by awarding points for the desired behavior and deducting points for its omission.

Objective: In this study, we introduced a principled mathematical method for determining how many points should be awarded or deducted for the enactment or omission of the desired behavior, depending on when and how often the person has succeeded versus failed to enact it in the past. We called this approach *optimized gamification of behavior change*.

Methods: As a proof of concept, we designed a chatbot that applies our optimized gamification method to help people build healthy water-drinking habits. We evaluated the effectiveness of this gamified intervention in a 40-day field experiment with 1 experimental group (n=43) that used the chatbot with optimized gamification and 2 active control groups for which the chatbot's optimized gamification feature was disabled. For the first control group (n=48), all other features were available, including verbal feedback. The second control group (n=51) received no feedback or reminders. We measured the strength of all participants' water-drinking habits before, during, and after the intervention using the Self-Report Habit Index and by asking participants on how many days of the previous week they enacted the desired habit. In addition, all participants provided daily reports on whether they enacted their water-drinking intention that day.

Results: A Poisson regression analysis revealed that, during the intervention, users who received feedback based on optimized gamification enacted the desired behavior more often (mean 14.71, SD 6.57 times) than the active (mean 11.64, SD 6.38 times; $P<.001$; incidence rate ratio=0.80, 95% CI 0.71-0.91) or passive (mean 11.64, SD 5.43 times; $P=.001$; incidence rate ratio=0.78, 95% CI 0.69-0.89) control groups. The Self-Report Habit Index score significantly increased in all conditions ($P<.001$ in all cases) but did not differ between the experimental and control conditions ($P>.11$ in all cases). After the intervention, the experimental group performed the desired behavior as often as the 2 control groups ($P\geq.17$ in all cases).

Conclusions: Our findings suggest that optimized gamification can be used to make digital behavior change interventions more effective.

Trial Registration: Open Science Framework (OSF) H7JN8; https://osf.io/h7jn8

(JMIR Serious Games 2024;12:e43078) doi:10.2196/43078

KEYWORDS

gamification; points; feedback; behavior change; habit formation; chatbot; digital interventions; mobile phone; artificial intelligence

Introduction

Background

Overview

People often struggle to change their behavior in ways that would benefit them in the long run. For instance, many people could improve their health and life expectancy by building healthy habits such as exercising [1], healthy eating [2], or drinking a glass of water before every meal [3,4]. People who want to adopt healthy habits because they know about their benefits nevertheless struggle to change their behavior accordingly.

Building a good habit is especially difficult when the benefits of the desired behavior cannot be felt immediately. One approach to alleviate this challenge is to create apps that encourage the desired behavior using incentives or immediate positive feedback [5-7] or discourage unwanted behavior using immediate negative feedback [8]. Doing so in a gameful way can be a promising approach to improving people's health behaviors [5-7,9-12]. Using game elements to improve people's behavior in the real world is known as gamification [13]. One of the most commonly used gamification methods is awarding people points for specific behaviors [14,15]. The awarded points are often used to provide feedback to the user, define levels, award badges, or create leaderboards.

Despite the widespread use of points, levels, badges, and leaderboards, there is currently no principled way to choose precisely how many points a person should be awarded and when. This is a problem as making those decisions based on intuition can lead to point systems that inadvertently incentivize counterproductive behaviors or undermine the users' motivation [16-18]. For instance, a recent study found that the point system of the popular gamified habit formation app Habitica is actively harmful [17].

To help practitioners master the challenge of designing effective point systems that reliably foster positive behavior change, we introduced a mathematical principle for computing the number of points a person should receive for engaging in the desired behavior depending on their history of either engaging or not engaging in this behavior and how many points they should lose when they fail to do so. We called this principle *optimized gamification of behavior change*. As a proof of concept, we applied optimized gamification to design a chatbot that helps people develop the healthy habit of drinking water before every meal [3,4]. Our chatbot awards points for the desired behavior and deducts points for its omission. Critically, the number of points that the user gains or loses is computed using optimized gamification. Our chatbot combines optimized gamification with three established principles of behavior change: it (1) guides the user to set an implementation intention, (2) reminds them of their good intentions, and (3) supports self-monitoring.

We evaluated our chatbot in a longitudinal field experiment and found that optimized gamification can make digital behavior change interventions more effective. Our findings provide a proof of concept for a very general and principled approach to improving human behavior in the real world. In addition to this theoretical contribution, we introduced a chatbot for helping people develop a specific healthy habit (water drinking): the Good Habit Bot. This chatbot can easily be adapted to other health behaviors, including more critical health behaviors such as exercising, healthy eating, and other good habits that people want to establish.

The plan for this paper is as follows. We first introduce relevant background information about behavior change and gamification. We then present our theory of optimized gamification for behavior change. Next, we present the chatbot we designed as a proof of principle. After that, we present the methods and results of our field experiment and discuss its findings and implications.

Behavior Change Goals Versus Automatic Behavior

Human behavior is controlled by a combination of goal-directed decision-making (eg, *I will buy a gym membership because I want to lose weight*) and more automatic reactions to certain stimuli (eg, always stopping by the gym on the way home from work) [19]. Goal-directed decision-making derives choices from the outcomes that people value (eg, health or money) or want to avoid (eg, pain) according to their mental models of how those outcomes can be obtained. In contrast, automatic reactions do not consider the behavior's likely consequences in the current situation.

Obstacles to Behavior Change: Automaticity and Temporal Discounting

Most of our behavior is not primarily controlled by goal-directed decision-making but determined by people's automatic reactions. Therefore, automatic behavioral responses can interfere with people's ability to act in accordance with their behavior change goals. This is the proverbial conflict between bad habits and good intentions. People can inhibit and override their automatic behavioral responses (*bad habits*), but their capacity to do so is limited [20]. Therefore, the automaticity of human behavior is a crucial obstacle to intentional behavior change.

A second obstacle to successful behavior change is that the mechanisms of goal-directed decision-making are demonstrably biased in favor of immediate outcomes [21]. People give too much weight to their decision's immediate consequences and too little weight to its long-term consequences. This phenomenon is known as present bias [22]. It has been proposed that present bias occurs because the brain discounts delayed benefits as if they become less valuable the later they occur [23]. This phenomenon, which is known as temporal discounting, is well established in research on economic decisions and animal behavior.

Moreover, according to temporal motivation theory [24], temporal discounting is one of the main reasons why people fail to enact good intentions. Such failures of self-regulation are a critical obstacle to health behavior change [25,26]. Consistent with this explanation, people who discount delayed outcomes more strongly are likelier to engage in unhealthy behaviors and experience poor health [26].

Reinforcement Learning as a Mechanism of Behavior Change

As automatic responses are powerful drivers of human behavior, successful behavior change typically involves translating behavior change goals into automatic behavioral responses [27].

Automatic behavioral responses, including exercise habits, can be acquired through learning from experience. Model-free reinforcement learning is a well-established mechanism of learning automatic behavioral responses from experience [19]. This mechanism increases or decreases a person's propensity to engage in a specific behavior in a particular situation (eg, going to the gym after work) according to whether they experience the behavior's overall consequences as positive or negative. The vast literature on operant conditioning in animals and humans underscores that learning from reward and punishment is a powerful mechanism of behavior change [28,29]. Another complementary learning mechanism involves strengthening habits through mere repetition [30].

Supporting Behavior Change Through Incentives and Reinforcement

The literature surveyed previously demonstrates that goal-directed decision-making and automatic behavioral responses are responsive to rewards and punishments. Goal-directed decision-making is sensitive to anticipated future rewards, and automatic behavioral responses are shaped by the rewards or punishments that those behaviors have generated in the past. These effects can be leveraged to support behavior change. To foster behavior change via goal-directed decision-making, behavior change interventions can create and announce incentives for engaging in the desired behavior. To foster behavior change via reinforcement learning, behavior change interventions can reinforce the desired behavior with rewards or positive feedback.

A highly effective behavior change intervention that leverages both effects is contingency management [31]. Contingency management incentivizes a desired behavior change and rewards people when they enact it. Voucher-based reinforcement therapy for treating addiction is a highly successful example of contingency management [32]. This behavior change intervention awards the patient a voucher every time they submit a negative drug test. More recently, it has also been applied to foster other types of behavior change, including physical exercise [7,33] and treatment attendance [34]. Contingency management appears to be more effective when the desired behavior is reinforced more promptly, more frequently, and with rewards that are larger or increase throughout the intervention [35-37].

Digital Behavior Change Interventions

Developing digital behavior change interventions is a young and booming field [38]. Mobile apps have shown potential for fostering positive behavior change in domains such as physical exercise and healthy eating [38,39]. However, the average effect size of such interventions is still relatively small [40,41]. Most behavior change interventions are not derived from any theory, model, or framework [38]. Therefore, we suspect that there is still room for improvement and that at least some of this potential can be realized by adopting a more theory-driven approach.

Goal setting and self-monitoring are the most commonly used behavior change techniques [38]. A meta-analysis of studies on digital interventions for promoting physical exercise found that these 2 techniques are also the most effective ingredients of current digital behavior change interventions [41]. Goal setting entails guiding people to articulate their intent to perform certain behaviors in certain situations (eg, drinking a glass of water before every meal). Supporting self-monitoring often takes the form of helping people check or record whether they have enacted those intentions or track related outcomes (eg, their weight). Goal setting is especially effective when people formulate simple plans that specify the intended behavior and the situation in which they want to perform it as concretely as possible [42,43]. This approach is known as *implementation intentions*. Moreover, reminding people of their intentions via SMS text messages [44] and presenting them with positive reinforcement when they enact their intentions [5,7,33] have been found to be highly effective in promoting physical exercise.

Gamification

A recent meta-analysis found that approximately 1 in 5 digital behavior change interventions are designed within the gamification framework [38]. Gamification entails applying principles from game design and game elements, such as storytelling and rules for earning points and winning the game, to address real-world problems [13]. The basic idea is to motivate people to do things that benefit them or others, such as exercising and studying, in a gameful way. Previous research has found that gamification can improve people's behavior, achieve desired outcomes, and improve people's subjective experiences [45].

Gamification is already widely used in designing digital behavior change interventions [5], and previous studies have suggested that it can improve people's health behaviors [7,9-12,33]. One gamification strategy that is effective in digital behavior change interventions is awarding the user points as positive feedback for the desired behavior [5,7,33]. Such extrinsic incentives can increase the frequency of the targeted behaviors without affecting people's intrinsic motivation [46,47].

However, when gamification is not correctly designed, it can backfire and have adverse effects [16,17,48]. This has also been observed in the behavior change literature [12,49] and in gamified habit formation apps [17]. Getting the incentives exactly right can be crucial as points, levels, badges, and leaderboards do not foster the user's intrinsic motivation [46,47] and might sometimes even undermine it when they are not embedded in a compelling narrative [18]. Motivated by these problems, many authors have called for a more theory-driven approach to gamification in general [50] and gamifying digital behavior change interventions in particular [12].

Optimized Gamification

Previous work has investigated how many points should be awarded for which behavior to maximally benefit the user in the context of to-do list apps that help the user achieve their own goals [51,52]. Building on temporal motivation theory [24], this work assumed that people's motivation is insufficiently sensitive to long-term benefits such as good health in old age and overly sensitive to immediate costs (eg, the effort of exercising) and short-term pleasure (eg, from receiving immediate positive feedback). To help people overcome the resulting motivational problems (eg, procrastination) [53], Lieder et al [51] developed a mathematical theory for designing point systems that provide immediate positive feedback for activities that are beneficial in the long run and immediate negative feedback for activities that are not. The basic idea is to align each action's immediate and long-term consequences. The action that is best in the long run should be made most appealing in the short run, and actions with undesirable long-term consequences should be made unappealing in the short run.

Therefore, optimized gamification strives to incentivize each of the available actions through a number of points proportional to how much that action increases or decreases the sum total of future happiness. This idea is implemented by modeling the activities to be incentivized as steps that lead toward a valuable goal or away from it. Actions that lead toward the goal increase the time the person will spend in the more valuable state in which the goal has been achieved. In contrast, actions that lead away from the goal decrease the time the person will spend in the more valuable state in which the goal has been achieved and increase the effort required to achieve it afterward. On the basis of this mathematical model, dynamic programming and reinforcement learning methods can estimate how much a given action improves or worsens the person's situation. These estimates are then translated into incentives that encourage good choices and discourage bad ones. The resulting point values are optimal in that they would enable even a purely myopic decision maker who only cares about immediate outcomes to choose the actions that are best for them in the long run [51].

Although optimized gamification construes points as incentives and uses mathematical and computational methods from the field of reinforcement learning, using it does not constitute a commitment to behaviorism and is fully compatible with cognitive theories of motivation [54].

Optimized gamification has been used to encourage users to tackle the tasks on their to-do lists [32] and encourage students to select the most valuable learning activities [55,56]. Optimized gamification has also been applied to give people feedback on how they think about what to do [57] and on whether they succeeded in staying focused on a chosen task or got distracted [58]. However, to date, this approach has never been applied to support habit formation.

Objectives

The first goal of this study was to introduce a principled method for computing feedback on the enactment or omission of the desired behavior and experimentally test whether it can be used to enhance digital behavior change interventions. The second goal of this study was to introduce a chatbot that uses this method to help people develop healthy water-drinking habits and evaluate it in a longitudinal field experiment.

Methods

Optimized Gamification of Behavior Change

Overview

We conceptualized behavior change as a special case of repeatedly choosing and learning when to do what. As reviewed in the *Background* section, optimized gamification can encourage desired behaviors and accelerate learning [57]. To apply this method to promote the desired behavior and accelerate the formation of healthy habits, we first have to model habit formation as a Markov decision process (MDP) [59].

Modeling Habit Formation

An MDP is a scenario in which an agent faces a series of choices. Each choice (*a*) has 2 effects. First, it yields an immediate reward (*r*) that may be positive, negative, or zero. Second, it may change the state (*s*) the agent finds itself in. In an MDP, the agent's goal is to maximize the sum of the rewards it accumulates from its first decision to its last one.

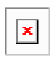
We model behavior change problems as a straightforward MDP, in which a person repeatedly chooses between 2 possible actions when they find themselves in a particular situation: enacting the desired behavior (a=1) or not enacting it (a=0). Our model assumes that a given behavior change intervention aims to turn the desired behavior into a habit. Therefore, we define the state as the strength of the person's healthy habit, measured using a single number, s_{habit}, which can range from 0 to 1. Following standard habit formation models [30], we assume that enacting the habit increases its strength from

$$s_{habit} \text{ to } s_{habit} + \alpha \times (1 - s_{habit}),$$

where α is a free parameter that describes how quickly habits form. Conversely, our model assumes that failure to enact the desired behavior decreases the strength of the habit to $s_{habit} \times (1 - \alpha)$. We assume that the habit has been cultivated when its strength exceeds some threshold θ (eg, θ=0.9) and model the health benefits conferred by achieving this goal as a reward (r_{goal}) that is attained when the habit strength crosses this threshold. Enacting the desired behavior is assumed to incur a cost that decreases with the strength of the habit ($r[s_{habit}, 1]=–[1 – s_{habit}]$), whereas not performing the behavior is assumed to be effortless ($r[s_{habit}, 0]=0$).

Computing Optimal Feedback

The basic idea of optimized gamification is to reward each action using a number of points that are proportional to its long-term benefits. These long-term benefits are measured via the decrease in future costs and the increase in future rewards brought about by transitioning to a state in which the habit is stronger. In situations in which the benefits of developing the good habit outweigh its costs, the value of having a habit of strength s_{habit} and then following through with the process of building the habit is as follows:



where the number $n(s_{habit}; \theta)$ specifies how often the behavior must be enacted until the habit strength reaches its target value θ. Therefore, for someone who will follow through with building the habit, the long-term benefits of enacting the habit one more time when its current strength is s_{habit} are

$$f(s_{habit}, 1) = V^*(s_{habit} + \alpha \times [1 - s_{habit}]) - V^*(s_{habit}) = 1 - s_{habit}.$$

Conversely, the long-term costs of failing to enact the desired behavior in the situation in which it is supposed to become a habit are

$$f(s_{habit}, 0) = V^*(s_{habit} \times [1 - \alpha]) - V^*(s_{habit}).$$

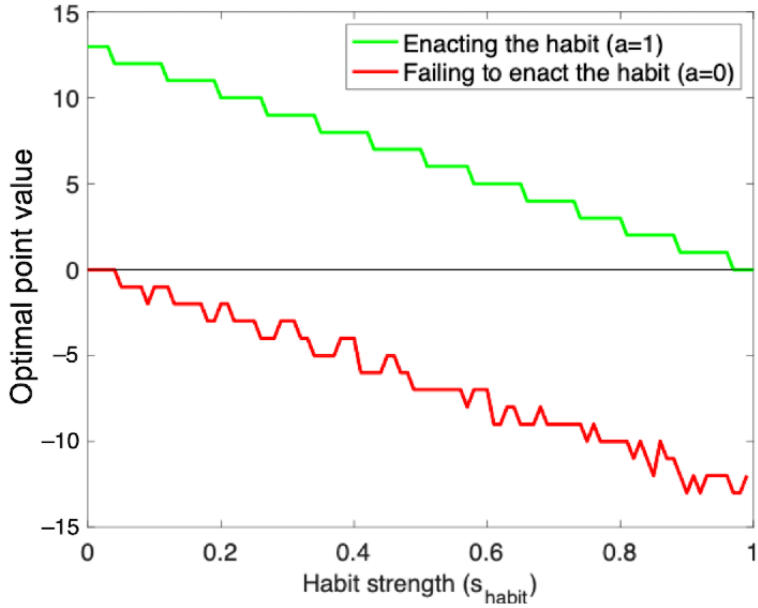
Please note that, even though we are talking about a situation in which it is rational for people to build the habit, this does not mean that we assume people to be rational. For our method, it does not matter why people follow through with building the habit. In fact, we assume that some people will follow through with building the habit only because they are (irrationally strongly) motivated by the immediate rewards conferred by feedback.

The lowest possible negative value of f is $f(\theta, 0)$, and the largest possible positive value is $f(0, 1)$. Although the exact values depend on the model parameters, they are typically approximately –1 and 1, respectively. Therefore, to transform those values into points, it is desirable to scale them by the desired maximum point value (M) that the application should award to the user and then round the scaled values to the nearest integer. This yields the following equation for the number of points that the application should award when a user reports that they have enacted their intention (a=1) or have not enacted their intention (a=0):

$$\text{points}(s_{habit}, a) = \text{round}(M \times f(s_{habit}, a)) \text{ (Equation 1).}$$

Figure 1 illustrates the point values as a function of the user's habit strength in an example application with a learning rate of α=0.1, a target habit strength of θ=0.9, and a maximal point value of M=13. Multimedia Appendix 1 provides more details on our mathematical model and the resulting optimal point system.

Figure 1. Point values for enacting versus failing to enact the habit at different habit strengths for a learning rate of α=0.1 and a threshold of θ=0.9.



As illustrated in Figure 1, the number of points for enacting the desired behavior is largest when the habit is weakest and gradually decreases toward 0 as the habit strengthens. This is intuitive as performing the desired behavior becomes easier the more often one has already performed it. Conversely, the number of points that should be deducted when the user fails to enact the habit is smallest when the strength of the habit is 0 and then increases as the habit becomes stronger. This is an intuitive consequence of our model's assumption that failing to enact the desired behavior in the specified situation weakens the habit strength by approximately 10%. The stronger the habit, the more is lost when its strength is reduced by 10%. The number of points awarded for enacting the desired behavior is a monotonically decreasing function of the habit strength. In contrast, the point value for failing to enact the desired behavior changes more erratically. This is because the number of steps required to reach the desired habit strength changes abruptly with the current habit strength. For instance, failing to enact the behavior at a habit strength of 0.09 increases the number of times the behavior needs to be enacted to achieve the desired habit strength from 21 to 22 times. However, if the habit strength is 0.08 or 0.10, the number of times the desired behavior has to be enacted remains 22 and 21 times, respectively. Individual users rarely experience such irregular changes as the change in their habit strength typically skips across those small areas in which the point value changes nonmonotonically. Moreover, our simulations suggest that the penalty for failing to enact the desired behavior can be approximated using a linear function with the same slope as the number of points for performing the desired behavior.

Application to Supporting Positive Behavior Change

The optimized gamification method described previously can be applied to help people form good habits. The equations for computing the number of points are easy to implement within digital behavior change interventions such as chatbots and habit trackers. All that is needed is to ask the user which habit they want to develop and estimate its initial strength, set the learning rate parameter and the habit's target strength to reasonable values (eg, α=0.1 and θ=0.9), and record when the user did versus did not act in accordance with the desired habit. To define the desired habit, the user has to specify the desired behavior and the situation in which they want to perform it. The user's initial habit strength can be estimated through the desired behavior's relative frequency in that situation in the previous weeks (eg, at 2 of the previous 7 lunches). The habit strength can then be initialized using that proportion. Alternatively, when it makes sense to assume that the user wants to build an entirely new habit, the strength can be initialized as 0. Then, whenever the user reports having or not having enacted their intention, the optimized gamification equation can be applied to compute how many points the user should gain or lose. Whenever the user reports having performed the desired behavior, the estimate of the habit strength should be increased to

$$s_{habit} + \alpha \times (1 - s_{habit}),$$

and whenever the user reports having missed an opportunity to enact their intention, the estimate of their habit strength decreases to $s_{habit} \times (1 - \alpha)$. Then, the same procedure repeats when the user reports on their next opportunity to enact the behavior.

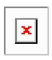
This approach can be used to support many different types of positive behavior change. It can help people build good habits in areas such as exercise, sleep, taking medication, nutrition, work, chores, and leisure activities. It can be used in interventions focusing on specific habits and, in general, in habit formation tools that let users choose any habit they want to develop. Another possible application is helping people overcome bad habits (eg, smoking cessation) [60]. It can support applications that run on practically any device, from smartwatches and wristbands to mobile phones, desktop applications, web applications, and smart glasses.

Proof of Concept: A Chatbot for Building a Healthy Water-Drinking Habit

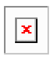
As a proof of concept for the application of optimized gamification to support behavior change, we implemented this idea as a Telegram (Telegram FZ LLC) chatbot called the *Good Habit Bot*. This chatbot helps the user develop a healthy water-drinking habit by combining 4 behavior change techniques: goal setting, reminders, support for self-monitoring, and feedback. Concretely, the Good Habit Bot guides the user to formulate an implementation intention that links a specific desired behavior to a concrete daily event, reminds the user of their intention on a daily basis, checks in with them on whether they followed through on their intention every day, and then gives them positive or negative feedback depending on whether they did or did not follow through on their intention.

When the user starts their first conversation with the Good Habit Bot, the chatbot introduces itself and says that its purpose is to help the user form a healthy water-drinking habit. The Good Habit Bot then asks the user to choose which of 8 concrete, recurring moments in their day they want to use as the trigger for their water-drinking habit (eg, *when my wake-up alarm rings* or *when I have the first bite of my lunch*; for the complete list, see Multimedia Appendix 2). Next, the Good Habit Bot asks the user how much water they want to drink in that situation (eg, 1 glass or 0.5 glasses) and how often they did so in the previous 7 days. The chatbot then uses the number of days *n* of the previous week in which the user acted in accordance with the habit (eg, n=2) to initialize their habit strength by n/7 (eg, s_{habit} = 2/7). Then, in the evening of the first day (ie, at 9 PM), the chatbot reminds the user of their intention to drink a specific amount of water at a particular moment of the following day (eg, *Remember your intention*: *When I have the first bite of my lunch, I will drink 1.5 glasses of water*). Then, sometime after the moment in which the user wanted to enact their intention, the chatbot asks them whether they did so (ie, *Did you accomplish your goal today to drink 1.5 glasses of water*). If the user affirms that they followed through on their intention, the chatbot gives them positive feedback (Figure 2). This feedback comprises a congratulatory message (eg, *That's wonderful!)* whose text alternates among 5 possible phrases (Multimedia Appendix 2) and a second message that awards the user the number of points computed by our optimized gamification method (eg, *I am glad to grant you 5 points for keeping a good habit! Your total score is 49 points*). In contrast, if the user responds that they missed their chance to enact their water-drinking intention, the Good Habit Bot tells them *Okay. Keep going tomorrow!* and informs them how many points they lost and how many they have left (Multimedia Appendix 2). Afterward, the chatbot updates the user's habit strength. Later that day, the Good Habit Bot reminds the user of their intention for the next day, and then the cycle repeats.

Critically, the chatbot computes how many points to award or take away from the user according to the optimized gamification method described previously (equation 1). One can read the number of points the chatbot we used in our experiment awarded in different situations in Figure 1 as it used the same set of parameters (ie, α=0.1; θ=0.9; *M*=13). For instance, if a user who reported having performed the desired behavior twice in the previous week enacted their intention on the first day, they earned 9 points, and their habit strength increased from



Conversely, if they failed to enact their intention, they would lose 4 points, and their habit strength would decrease to



The Good Habit Bot is freely available on the Telegram messenger app, where it can be found by searching for its alias, @learn_good_habits_bot.

Figure 2. Screenshot of the feedback method with point values computed by our optimized gamification method.

Study Design

To evaluate the effectiveness of our intervention and assess the relative contributions of reminders versus optimized gamification, we ran a longitudinal between-subject experiment with 1 experimental condition with optimized gamification (*optimized gamification condition*), 1 *baseline condition*, and 1 *control condition with reminders and feedback* (Table 1). In the *optimized gamification condition*, the chatbot delivered all 4 techniques described previously, including feedback messages based on optimized gamification (Figure 1). Participants in the other 2 conditions could not gain or lose any points (no optimized gamification). They differed in whether they received feedback messages for enacting versus failing to enact the intended behavior. In the *control condition with reminders and feedback*, participants received a positive feedback message when they reported having enacted their intention (eg, *That's wonderful!)* and a more neutral message when they reported not having done so (eg, *Okay. Keep going tomorrow!).* To create the *baseline condition*, we replaced the first control group's positive and negative feedback message with a neutral message (*OK*) and removed the daily reminders. Participants completed the self-report measures of habit strength described in the following sections before the intervention (pretest time point), immediately after the intervention (posttest time point), and approximately 3 weeks later (follow-up). Moreover, participants reported how often they engaged in the desired behavior the week before the intervention (pretest time point) and the week after the intervention (posttest time point). Finally, participants also completed daily reports of whether they enacted their intention on each day of the intervention.

Table 1. Experimental conditions.

Experimental condition	Implementation intentions	Support for self-monitoring	Reminders	Feedback
Optimized gamification condition	Yes	Yes	Yes	Optimized gamification and positive vs negative text
Control condition with reminders and feedback	Yes	Yes	Yes	Positive vs negative text
Baseline condition	Yes	Yes	No	None

Ethical Consideration

This experiment was conducted according to study protocol 401/2020BO2 approved by the Independent Ethics Commission at the Faculty of Medicine of the University of Tübingen. All data was collected and handled according to the General Data Protection Regulation of the European Union. All data has been de-identified.

Recruitment and Reimbursement

We recruited 132 participants (n=41, 31.1% for the baseline condition; n=43, 32.6% for the optimized gamification condition; n=48, 36.4% for the control condition; n=93, 70.5% female) on the web-based research platform Prolific. Based on considerations about the cost of the study, the sample size was determined a priori so that we would retain 3x40=120 participants after an anticipated 10% of participants dropped out of the study. The requirements for participation were being a native English speaker, not having interacted with our chatbot before, and having previously completed at least 10 prolific assignments with an approval rate of at least 95%. Moreover, participants had to be aged ≥18 years. The average age of the participants was 31.5 (SD 9.9; range 19-79) years.

The study description informed participants about the study's duration, activities, time commitment, and pay. Participants were paid £1.95 (US $2.45) for completing the onboarding survey. We informed them that the base pay for completing the remainder of the 40-day study would be £7.80 (US $9.81) and that they could earn an additional bonus of £8 (US $10.06). The description strongly recommended that only people who were already using Telegram on their smartphones should participate. Moreover, the study description informed potential participants about the potential health benefits of regular water drinking. Participants were then shown the consent form. Upon providing informed consent, participants who already had the Telegram app clicked on a link that started a conversation with the Telegram chatbot for their corresponding experimental condition. Participants who had not installed the Telegram app yet were directed to download it first.

At the end of the 40-day study, the chatbot directed participants to a second Prolific HIT where they received the announced payments contingent on their sustained active participation. All participants who completed the pretest, posttest, and follow-up measures and continued to report their daily intention enactment after the 10th day received a second payment of £15.80 (US $19.86). Participants who did not meet these criteria did not receive the second payment.

Outcome Measures and Procedure

Outcome Measures

We measured the outcome variables described in the following sections.

Dropout

We measured whether a participant dropped out of our study using a binary variable indicating whether the participant stopped responding to all daily reports at least 3 days before the end of the study.

Engagement

We measured each participant's *engagement* with our digital intervention based on the number of days on which they interacted with the Good Habit Bot.

Self-Report Habit Index

The Self-Report Habit Index (SRHI) [61] is a 12-item self-report measure of habit strength on a 7-point Likert scale. It comprises 3 subscales measuring the behavior's *history of repetition*, its *automaticity*, and the extent to which it is part of the person's *identity*. In this study, we administered the first 2 subscales. The SRHI has been found to be a 1D construct. Therefore, we averaged the scores of all items. The SRHI has been found to have high validity and very high reliability (Cronbach α of approximately 0.90; test-retest reliability: r=0.91).

Daily Intention Enactment

To measure how often each participant enacted their water-drinking intention during the intervention, we asked them the following question—*Did you accomplish your goal to drink 1 glass of water?*—on each day of the intervention. The question was asked between 30 minutes and 2.5 hours after the time at which the participant intended to drink water. Participants responded by selecting between the responses *Yes, I did!* and *No, I didn't.* After the study, we calculated each participant's *daily intention enactment* score by counting on how many of the 21 days of the intervention period they enacted their intention. Therefore, the daily intention enactment score could range from 0 to 21.

Retrospective Intention Enactment

To measure how regularly participants engaged in the intended behavior (eg, drinking a glass of water before lunch) in the weeks before and after the intervention, we asked them to answer the following question—*On how many days of the previous week did you keep the habit of drinking water?*—by selecting one of the answer choices (0, 1, 2, 3, 4, 5, or 6 days). We referred to the resulting number of days as the *retroactive intention enactment* score.

Procedure

We created 3 separate Prolific HITs for each of the 3 conditions of the experiment (Table 1), and each person was allowed to participate in at most one of these HITs. The experiment ran from November 11, 2021 to December 19, 2021. As illustrated in Table 2, the experiment was divided into 3 phases: the preintervention period (day 0), intervention period (days 1-21), and postintervention period (days 22-40). In the preintervention period, participants provided informed consent, completed the onboarding process, and completed the pretest. We blinded participants to the experimental manipulation by giving all participants the same information about the chatbot they were interacting with and the anticipated benefits of interacting with it. During onboarding, participants were directed to start the first conversation with the Good Habit Bot in the Telegram app on their mobile phones. In this initial conversation, the chatbot asked the participants to select a concrete daily situation in which they wanted to drink water and how much water they wanted to drink, as described previously. The pretest comprised 2 self-report measures: the *SRHI* and the *retrospective intention enactment* measure for the week before the study.

During the intervention period, each participant interacted with 1 of the 3 versions of our chatbot according to the condition they were in (Table 1). On each day of the intervention period, all 3 groups reported whether they successfully enacted their intention to drink water in the specific situation they had selected (*daily intention enactment*). At the end of the intervention period, all 3 groups completed the *SRHI* for the second time (posttest time point).

The postintervention period started with a 1-week break during which the chatbot did not communicate with the participants. Then, on day 28 (follow-up 1) and day 35 (follow-up 2), the chatbot asked participants from all 3 groups to report on how many days of the previous week (ie, the first and second week of the postintervention period, respectively) they had acted in accordance with the desired habit (*retroactive intention enactment*). Finally, on day 40, the chatbot asked all participants to complete the *SRHI* questionnaire for the third time (follow-up). Participants received 3 email reminders to resume interacting with the chatbot on the day of the first follow-up survey (day 28), the day of the second follow-up survey (day 35), and the day of the final follow-up survey (day 40).

Table 2. Experimental procedure.

Experimental phase and day	Optimized gamification condition	Control condition with reminders and feedback	Baseline condition
Before the intervention			
Day 0	• Onboarding • Retrospective intention enactment • SRHI[a]	• Onboarding • Retrospective intention enactment • SRHI	• Onboarding • Retrospective intention enactment • SRHI
Intervention			
Days 1-21	• Reminder, report, and feedback (optimal points+text) • Daily intention enactment	• Reminder, report, and feedback (text only) • Daily intention enactment	• Daily intention enactment
Day 21	• SRHI	• SRHI	• SRHI
After the intervention			
Days 22-27	• No reports	• No reports	• No reports
Day 28	• Retrospective intention enactment	• Retrospective intention enactment	• Retrospective intention enactment
Days 29-34	• No reports	• No reports	• No reports
Day 35	• Retrospective intention enactment	• Retrospective intention enactment	• Retrospective intention enactment
Days 36-39	• No reports	• No reports	• No reports
Day 40	• SRHI	• SRHI	• SRHI

[a]SRHI: Self-Report Habit Index.

Data Analysis

The hypotheses and statistical analysis plan were preregistered on the internet [62]. Participants who completed 0 daily water consumption reports were excluded from all analyses apart from the dropout analysis. Other than that, all analyses were conducted on all available data from all participants who completed at least one daily water consumption report. We retained 37/41 participants from the baseline condition, 39/43 participants from the control condition with reminders and feedback, and 42/48 participants from the optimized gamification condition. All comparisons between conditions were based on the originally assigned groups. We used Poisson regression analyses for binary outcome variables. For continuous outcome variables, we used linear multilevel modeling.

Results

There was no indication of failure of random assignment for initial habit strength, automaticity, or history of repetition (pairwise *P*>.14 in all cases).

Dropout

As an initial step, we examined whether there was a differential dropout among the 3 conditions using a chi-square test of independence. There was no effect of condition on dropout (χ^2_2=0.7; Cramer *V*=0.07; *P*=.72).

Engagement

The *engagement* variable was entered into a Poisson regression model with 2 dummy variables for the effects of optimized gamification and feedback and reminders. The *automaticity* and *history of repetition* scores from the SRHI before the intervention and the preintervention *retrospective intention enactment* score were entered as control variables. As shown in Table 3, optimized gamification did not increase engagement compared with the *baseline condition*. However, it appears that being in the control condition with reminders and feedback without optimal points reduced *engagement* compared with the *optimized gamification* condition (Table 3) and the *baseline* condition (incidence rate ratio=0.84, 95% CI 0.75-0.94; *P*=.002).

Table 3. Predicting the number of days on which participants engaged with the app (engagement) from their condition and time-1 habit-related control variables (N=126)[a].

Predictor	Incidence rate ratio (95% CI)	*P* value
Intercept	18.88 (16.61-21.43)	<.001
Baseline vs optimized gamification	1.02 (0.91-1.14)	.72
Reminders and feedback vs optimized gamification	0.86 (0.76-0.96)	.008
Preintervention habit strength	1.03 (1.00-1.07)	.07
History of repetition	0.85 (0.79-0.92)	<.001
Automaticity	1.08 (1.01-1.16)	.03

[a]Nagelkerke R^2=0.33. The treatment group with optimal points is the reference group.

Daily Intention Enactment

The outcome variable measuring participants' daily enactment of the desired behavior (ie, water drinking) was subjected to a Poisson regression model with the same set of predictors as for *engagement*. Critically, we found that participants in the optimized gamification condition enacted the daily intention to drink water more often (mean 14.71, SD 6.57 times) than either the participants in the baseline condition (mean 11.64, SD 5.43 times) or the participants in the control condition with reminders and feedback (mean 11.64, SD 6.38 times; Table 4). Furthermore, reminders and feedback without points did not result in more water drinking than the *baseline condition* (incidence rate ratio=1.03, 95% CI 0.90-1.17; *P*=.70).

Table 4. Predicting the number of days on which participants drank water (daily intention enactment) from their condition and preintervention measurements of habit-related control variables (n=118)[a].

Predictor	Incidence rate ratio (95% CI)	*P* value
Intercept	16.02 (13.93-18.41)	<.001
Baseline vs optimized gamification	0.78 (0.69-0.89)	.001
Reminders and feedback vs optimized gamification	0.80 (0.71-0.91)	<.001
Preintervention retrospective intention enactment	1.01 (0.97-1.05)	.64
History of repetition	0.91 (0.84-0.99)	.02
Automaticity	1.07 (0.99-1.15)	.10

[a]Nagelkerke R^2=0.21. The treatment group with optimal points is the reference group.

Self-Reported Habit Strength

To test whether optimized gamification promoted habit formation, we compared the change in the SRHI self-report measures of *automaticity* and *history of repetition* from pretest to posttest to follow-up time points among the 3 experimental conditions (Figure S1 in Multimedia Appendix 3) using a multilevel model with fixed-effects predictors for time point, 2 dummy codes for the experimental condition with the *optimized gamification condition* as the reference, and all pairwise *time*×*condition* interactions (Tables S1 and S2 in Multimedia Appendix 3). We found that, compared with the pretest time point, both measures of habit strength were significantly higher immediately after the intervention (*P*<.001 in all cases) and at follow-up (*P*<.001 in all cases). However, these effects were no larger in the *optimized gamification condition* than in the baseline condition (*P*≥.20 in all cases) or the control condition with reminders and feedback (*P*≥.12 in all cases).

Behavior After the Intervention

As a further test of whether the behavior change we observed during the intervention was maintained, we analyzed the number of times participants reported having enacted their intention in the week before the intervention versus the first week after the intervention and the second week after the intervention (*retrospective intention enactment*; Figure S2 in Multimedia Appendix 3) using a multilevel model with fixed-effects predictors for time point, 2 dummy codes for the experimental condition with *the optimized gamification condition* as the reference, and all pairwise *time*×*condition* interactions (Table S3 in Multimedia Appendix 3). We found that, compared with the week before the intervention (mean 1.8, SD 2.1 times), participants enacted the desired behavior significantly more often after the intervention (1-week follow-up: mean 5.2, SD 1.9 times, t_{299}=15.21, and *P*<.001; 2-week follow-up: mean 5.1, SD 2.1 times, t_{299}=14.86, and *P*<.001). However, these effects were no larger in the optimized gamification condition than they were in the baseline condition (*P*≥.18 in all cases) or the control condition with reminders and feedback (*P*≥.32 in all cases).

Discussion

Principal Findings

In this study, we derived a mathematical principle for designing the point systems of gamified behavior change interventions. Our proof-of-concept study suggests that this principled approach to gamifying behavior change can be beneficial. We found that our gamified behavior change chatbot fostered positive behavior change during the intervention. This is consistent with previous findings that goal setting, reinforcement, reminders, and self-monitoring are effective techniques for promoting behavior change [7,33,39,41,44].

Moreover, we found that the behavior change that occurred during the intervention was maintained in all 3 conditions. The elements that the behavior change interventions in all 3 groups shared were goal setting and self-monitoring. Therefore, goal setting and self-monitoring may be sufficient for sustained behavior change. Adding reinforcement to goal setting and self-monitoring was beneficial during the intervention, but the additional benefits of reinforcement ceased to be statistically significant ($P \geq .17$ in all cases) in the week following the intervention. However, as our study had a small sample size, this apparent discrepancy could be an artifact of us having used different methods to measure behavior change during versus after the intervention. During the intervention, we measured behavior change through daily self-reports. After the intervention, we asked participants to retrospectively report on their behavior in the previous week, which is less accurate because of participants' fallible memory, and complete self-report questionnaires about their perceived habit strength, which are less objective than reports on behavior. Consistent with the interpretation that our study had insufficient statistical power for detecting retention effects, the measures we used to assess the maintenance of behavior change consistently showed a nonsignificant trend in favor of optimized gamification (Figures S1 and S2 in Multimedia Appendix 3). Moreover, previous studies have found that gamification-induced behavior change can persist over extended periods [63].

Limitations

From a theoretical perspective, the main limitation of our study is that it did not compare the effectiveness of the points computed by our optimized gamification method with alternative point schemes. Previous work has found optimized gamification to be more effective than several heuristic methods for designing point systems in contexts in which people choose among several activities [51]. However, the decisions that people face in behavior change applications appear to be simpler. Therefore, it remains unclear how much of the beneficial effects of optimized gamification on behavior change are specific to optimized gamification. Relatedly, it remains unclear which property of the points generated through optimized gamification is responsible for their effects on behavior change. Future work could address these questions by comparing optimized gamification of behavior change with simpler, alternative feedback mechanisms such as always awarding the same number of points or a streak-based point system.

However, we did evaluate optimized gamification against rewarding each enactment of the desired behavior using the same positive feedback message and punishing each failure to enact the desired behavior using the same negative feedback message (control condition with reminders and feedback). From a reinforcement learning perspective, this condition is equivalent to always awarding the same number of points when the behavior is enacted and always deducting the same number of points when the user fails to enact the behavior. We found that optimized gamification is more effective than this alternative feedback mechanism. This suggests that optimized gamification might be more effective than awarding the same number of points for each instance of the desired behavior. However, whether this interpretation is correct remains to be tested.

We illustrated the application of the general framework of optimized gamification to behavior change using a simple model of habit formation, which assumes that the user will indefinitely maintain the good habit once it has been established. This assumption is highly simplistic. In reality, maintenance is far from automatic. On the contrary, people may experience backsliding, and the strength of the habit may continue to wax and wane [64]. This could be captured by letting the process of deciding whether to perform the behavior continue indefinitely until the user dies. In this way, lapses could occur at any time and weaken the habit no matter how strong it is. In such a model, the health benefits of the behavior could be modeled explicitly in terms of its effects on a state variable that models the user's health status. Refining our method's model in this way would reduce the reinforcement for engaging in the behavior when the habit is weak and increase it when the habit is strong. This may make our method even more effective.

From an empirical perspective, the main limitation of the work presented in this paper is the relatively small sample size of our field experiment. Given that we collected <30 complete data sets per condition, the power of some of the statistical tests is relatively low. Therefore, the absence of significant differences in self-reported habit strength and retroactive intention enactment after the intervention does not provide strong evidence against the maintenance of the benefits of optimized gamification. Moreover, given that habit formation can take a very long time, our intervention may have been too short to fully capture the effects of the 3 different interventions on habit formation.

One flaw in our experiment was that some of the chatbot's messages were not worded in perfect, idiomatic English (Multimedia Appendix 2). We think that it is unlikely that our participants misunderstood any of the messages. However, it is possible that participants would have taken the messages more seriously if all of them had been written in perfect, idiomatic English. Another minor limitation of our chatbot is that its users started with a score of 0. Thus, if they failed to enact the desired behavior, their score fell to negative values, which might be demotivating. Therefore, future versions of our chatbot will award users a number of points (eg, 20) for setting the intention to build a good habit.

Finally, another weakness of our study design is that our intervention sought to strengthen water drinking in healthy people. Therefore, we cannot draw conclusions about the potential utility of our chatbot for clinical populations for which developing a healthy water-drinking habit might be crucial [3,4]. Moreover, it remains unclear whether our findings can be generalized to other habits that are more vital to people's health. In addition, our study was not specifically about water drinking as a weight loss strategy as only some participants tied water drinking to their meals. However, we introduced a general method that can be used to improve digital interventions for many critical behavior change applications.

Comparison With Prior Work

This study builds on previous work on optimized gamification [51]. Optimized gamification has been previously applied to help people decide what to work on [51] and which goals to set [52,55,56,65]. Moreover, optimized gamification has also been applied to help people stay focused on their work [58]. However, the work presented in this paper is the first application of optimized gamification to support habit formation. Moreover, it is one of the first real-world applications of optimized gamification as most previous work was confined to controlled laboratory experiments.

As our chatbot combined 4 established behavior change techniques (ie, goal setting, reminders, support for self-monitoring, and reinforcement), its design and effectiveness are therefore consistent with several extant theories of behavior change [66]. In particular, our optimized gamification method is consistent with behavior change methods that acknowledge the importance of providing positive feedback on improvements in behavior [28,66-69].

The most similar gamified digital intervention for behavior change that we are aware of is the SMS text messaging–based WalkIT intervention for promoting physical exercise [7,33,70]. Participants of the WalkIT trial received SMS text messages with daily physical exercise goals. Physical exercise (walking) was measured using the accelerometers of their smartphones and reported to a server. Participants received feedback on whether they met the exercise goal via an SMS text message that included points that were converted into money. Depending on the stage of the experiment, the number of points for achieving a goal was either constant or determined at random. The number of points awarded for failing to achieve a goal was 0. In contrast, our optimized gamification method provides a principled way to choose the exact number of points that a person should be awarded for meeting their daily goal or lose for failing to meet that goal. Therefore, our method could be used to enhance the WalkIT intervention with a more principled way of choosing the number of points depending on the user's history of successful and unsuccessful goal achievement. Conversely, WalkIT has many sophisticated features that go beyond the chatbot we introduced here. This includes an algorithm for adaptive goal setting, automatically delivered financial incentives, and an evidence-based sequence of different reward schedules that differ in the probability that goal achievement will be rewarded and whether the magnitude of the reward will be fixed or random. Consistent with our finding that the more performance-contingent feedback of the optimized gamification condition was more effective than less informative feedback or no feedback, the WalkIT studies found that immediate, behavior-contingent reinforcement was more effective in promoting behavior change than delayed, behavior-independent reinforcement.

Another gamified digital intervention for supporting behavior change is Habitica. Habitica embeds working through one's to-do list into a role-playing game in which the user's character can earn points by completing their daily to-dos. The points serve as an in-game currency that the player can use to buy weapons and armor for their avatar. Conversely, when the user does not complete a daily to-do, they lose points. As far as we can tell, the developers of Habitica chose the number of points the user gains for completing a to-do and the number of points they lose for failing to complete a daily to-do somewhat arbitrarily based on their intuitions. A recent study found that only 49% of Habitica's users rate its rewards as (rather) appropriate and that most experience counterproductive effects of Habitica's approach to gamification [17]. Given that optimized gamification was effective in our study and in previous studies, it is possible that redesigning Habitica's point system according to optimized gamification could alleviate some of the counterproductive effects of their users' experience.

The method introduced in this study mitigates the adverse effects of temporal discounting on people's health behavior [26]. Its approach is to add immediate rewards that are aligned with the behavior's long-term consequences for the user's health; that is, optimized gamification redesigns the decision environment so that people's shortsighted biases lead to decisions that are good for them in the long run [51]. Recent work on this topic introduced a computational model of intertemporal choice and applied it to compute personalized incentives for helping people more patiently work toward obtaining a larger reward later instead of abandoning the project in favor of a smaller immediate reward [71,72]. Similar to optimized gamification, their approach uses an MDP framework. However, their problem formulation and solution are different. The main difference lies in the application area. Sukumar et al [71] focused on canonical delay-of-gratification tasks, whereas we modeled habit formation. They tested their approach in online experiments in which participants played a *queue waiting game* and found that personalized incentives can increase people's patience while waiting in a simulated queue. In contrast, we conducted a field experiment on behavior change in which the incentives motivated people to act more farsightedly in the real world. Despite this critical difference, investigating whether modeling and measuring individual differences can be used to make optimized gamification more effective is an exciting direction for future work. Moreover, a computational model such as the one proposed by Sukumar et al [71] could be used to simulate the effects of alternative incentive schemes.

Previous work has found that drinking water before meals is an effective weight loss strategy for adults with obesity [3,4]. In the randomized controlled trial by Parretti et al [4], participants were instructed to use the water-drinking strategy in a face-to-face weight loss consultation. They did not receive any additional support in implementing this strategy. The chatbot we developed could be used to augment those weight loss consultations with a digital tool that helps people follow through on their resolutions. Alternatively, an appropriately adapted version of our chatbot could be used as a highly scalable, low-cost alternative to face-to-face weight loss consultations.

Conclusions

In conclusion, optimized gamification is a practically helpful theoretical principle for designing the point systems of (digital) behavior change tools and interventions. It can be implemented in just a few lines of code, and the point values can be computed instantaneously. It can be applied to improve or augment many existing (digital) behavior change interventions and can also be used to create new ones. Thus, optimized gamification can help tackle many challenging behavior change problems using scalable digital interventions. Testing whether, when, and how optimized gamification can make a positive difference in critical practical applications is an exciting direction for future research. A crucial next step will be to test our point system against simpler heuristic point systems for supporting behavior change. Moreover, our chatbot can be extended to support various health behaviors and other forms of positive behavior change.

Acknowledgments

This work was supported by grant 1757269 from the National Science Foundation.

Data Availability

The data generated in this study, the R scripts used to analyze the data, and the MATLAB (MathWorks) implementation of the optimized gamification method are available on the project's Open Science Framework repository [73].

Authors' Contributions

FL conceived and designed the study, developed the optimized gamification method, and wrote the manuscript. PZC designed and implemented the Good Habit Bot and co-designed and conducted the field experiment. MP conducted the data analysis. VA supported PZC in designing and conducting the field experiment and in the design of the Good Habit Bot. MT contributed to the implementation of the Good Habit Bot.

Conflicts of Interest

None declared.

Multimedia Appendix 1
An optimal feedback method for accelerating positive behavior change.
[DOCX File , 16 KB - games_v12i1e43078_app1.docx]

Multimedia Appendix 2
Details about the Good Habit Bot.
[DOCX File , 17 KB - games_v12i1e43078_app2.docx]

Multimedia Appendix 3
Details about the statistical results and supplementary analyses.
[DOCX File , 122 KB - games_v12i1e43078_app3.docx]

References

1. Physical activity. World Health Organization. 2022 Oct 5. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2024-02-08]
2. Bull ER, McCleary N, Li X, Dombrowski SU, Dusseldorp E, Johnston M. Interventions to promote healthy eating, physical activity and smoking in low-income groups: a systematic review with meta-analysis of behavior change techniques and delivery/context. Int J Behav Med 2018 Dec 12;25(6):605-616 [FREE Full text] [doi: 10.1007/s12529-018-9734-z] [Medline: 30003476]
3. Daniels MC, Popkin BM. Impact of water intake on energy intake and weight status: a systematic review. Nutr Rev 2010 Sep;68(9):505-521 [FREE Full text] [doi: 10.1111/j.1753-4887.2010.00311.x] [Medline: 20796216]
4. Parretti HM, Aveyard P, Blannin A, Clifford SJ, Coleman SJ, Roalfe A, et al. Efficacy of water preloading before main meals as a strategy for weight loss in primary care patients with obesity: RCT. Obesity (Silver Spring) 2015 Sep 03;23(9):1785-1791 [FREE Full text] [doi: 10.1002/oby.21167] [Medline: 26237305]
5. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
6. Berardi V, Hovell M, Hurley JC, Phillips CB, Bellettiere J, Todd M, et al. Correction to: variable magnitude and frequency financial reinforcement is effective at increasing adults' free-living physical activity. Perspect Behav Sci 2020 Sep 12;43(3):539-541 [FREE Full text] [doi: 10.1007/s40614-020-00254-y] [Medline: 33031485]

7. Adams MA, Todd M, Angadi SS, Hurley JC, Stecher C, Berardi V, et al. Adaptive goals and reinforcement timing to increase physical activity in adults: a factorial randomized trial. Am J Prev Med 2022 Feb;62(2):e57-e68 [FREE Full text] [doi: 10.1016/j.amepre.2021.09.014] [Medline: 35000693]
8. Berardi V, Bellettiere J, Nguyen B, Klepeis NE, Hughes SC, Adams MA, et al. Assessing reinforcing versus aversive consequences in a real-time secondhand smoke intervention. Transl Behav Med 2021 Aug 13;11(8):1558-1566 [FREE Full text] [doi: 10.1093/tbm/ibab004] [Medline: 33823045]
9. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
10. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
11. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Dec;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
12. Schmidt-Kraepelin M, Warsinsky S, Thiebes S, Sunyaev A. The role of gamification in health behavior change: a review of theory-driven studies. In: Proceedings of the 53rd Hawaii International Conference on System Sciences. 2020 Presented at: HICSS 2020; January 7-10, 2020; Maui, Hawaii URL: https://scholarspace.manoa.hawaii.edu/items/3e3f2688-4f1c-42a2-840e-ccefedcd85e8 [doi: 10.24251/hicss.2020.155]
13. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
14. Werbach K, Hunter D. For the Win: The Power of Gamification and Game Thinking in Business, Education, Government, and Social Impact. Philadelphia, PA: Wharton School Press; 2020.
15. Zichermann G, Cunningham C. Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. Sebastopol, CA: O'Reilly Media; 2011.
16. Toda AM, Valle PH, Isotani S. The dark side of gamification: an overview of negative effects of gamification in education. In: Proceedings of the First International Workshop on Social, Semantic, Adaptive and Gamification Techniques and Technologies for Distance Learning, HEFA 2017. 2017 Presented at: First International Workshop on Social, Semantic, Adaptive and Gamification Techniques and Technologies for Distance Learning, HEFA 2017; March 20-24, 2017; Maceió, Brazil URL: https://link.springer.com/book/10.1007/978-3-319-97934-2 [doi: 10.1007/978-3-319-97934-2_9]
17. Diefenbach S, Müssig A. Counterproductive effects of gamification: an analysis on the example of the gamified task manager Habitica. Int J Hum Comput Stud 2019 Jul;127:190-210. [doi: 10.1016/j.ijhcs.2018.09.004]
18. Nicholson S. A RECIPE for meaningful gamification. In: Reiners T, Wood L, editors. Gamification in Education and Business. Cham, Switzerland: Springer; 2015.
19. Dolan RJ, Dayan P. Goals and habits in the brain. Neuron 2013 Oct 16;80(2):312-325 [FREE Full text] [doi: 10.1016/j.neuron.2013.09.007] [Medline: 24139036]
20. Baumeister RF, Heatherton TF. Self-regulation failure: an overview. Psychol Inq 1996 Jan;7(1):1-15. [doi: 10.1207/s15327965pli0701_1]
21. Critchfield TS, Kollins SH. Temporal discounting: basic research and the analysis of socially important behavior. J Appl Behav Anal 2001 Feb 27;34(1):101-122 [FREE Full text] [doi: 10.1901/jaba.2001.34-101] [Medline: 11317983]
22. O'Donoghue S, Rabin M. Doing it now or later. Am Econ Rev 1999;89(1):103-124 [FREE Full text] [doi: 10.1257/aer.89.1.103]
23. Berns GS, Laibson D, Loewenstein G. Intertemporal choice--toward an integrative framework. Trends Cogn Sci 2007 Nov;11(11):482-488. [doi: 10.1016/j.tics.2007.08.011] [Medline: 17980645]
24. Steel P, König CJ. Integrating theories of motivation. Acad Manage Rev 2006 Oct;31(4):889-913. [doi: 10.2307/20159257]
25. Hennessy EA, Johnson BT, Acabchuk RL, McCloskey K, Stewart-James J. Self-regulation mechanisms in health behavior change: a systematic meta-review of meta-analyses, 2006-2017. Health Psychol Rev 2020 Mar 12;14(1):6-42 [FREE Full text] [doi: 10.1080/17437199.2019.1679654] [Medline: 31662031]
26. Story GW, Vlaev I, Seymour B, Darzi A, Dolan RJ. Does temporal discounting explain unhealthy behavior? a systematic review and reinforcement learning perspective. Front Behav Neurosci 2014 Mar 12;8:76 [FREE Full text] [doi: 10.3389/fnbeh.2014.00076] [Medline: 24659960]
27. Rebar AL, Gardner B. Habit formation and behavior change. In: Oxford Research Encyclopedia of Psychology. Oxford, UK: Oxford University Press; 2019.
28. Skinner BF. Science And Human Behavior. New York, NY: Macmillan Publishers; 1953.
29. Staddon JE, Cerutti DT. Operant conditioning. Annu Rev Psychol 2003;54:115-144 [FREE Full text] [doi: 10.1146/annurev.psych.54.101601.145124] [Medline: 12415075]
30. Miller KJ, Shenhav A, Ludvig EA. Habits without values. Psychol Rev 2019 Mar;126(2):292-311 [FREE Full text] [doi: 10.1037/rev0000120] [Medline: 30676040]

31. Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: a meta-analysis. Addiction 2006 Nov 10;101(11):1546-1560. [doi: 10.1111/j.1360-0443.2006.01581.x] [Medline: 17034434]
32. Higgins ST, Budney AJ, Bickel WK, Foerg FE, Donham R, Badger GJ. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. Arch Gen Psychiatry 1994 Jul 01;51(7):568-576. [doi: 10.1001/archpsyc.1994.03950070060011] [Medline: 8031230]
33. Berardi V, Hovell M, Hurley JC, Phillips CB, Bellettiere J, Todd M, et al. Variable magnitude and frequency financial reinforcement is effective at increasing adults' free-living physical activity. Perspect Behav Sci 2020 Sep 05;43(3):515-538 [FREE Full text] [doi: 10.1007/s40614-019-00241-y] [Medline: 33029578]
34. Pfund RA, Ginley MK, Rash CJ, Zajac K. Contingency management for treatment attendance: a meta-analysis. J Subst Abuse Treat 2022 Feb;133:108556 [FREE Full text] [doi: 10.1016/j.jsat.2021.108556] [Medline: 34210566]
35. Griffith JD, Rowan-Szal GA, Roark RR, Simpson DD. Contingency management in outpatient methadone treatment: a meta-analysis. Drug Alcohol Depend 2000 Feb 01;58(1-2):55-66. [doi: 10.1016/s0376-8716(99)00068-x] [Medline: 10669055]
36. Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. Addiction 2006 Feb;101(2):192-203. [doi: 10.1111/j.1360-0443.2006.01311.x] [Medline: 16445548]
37. Roll JM, Higgins ST, Badger GJ. An experimental comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. J Appl Behav Anal 1996 Feb 27;29(4):495-504; quiz 504 [FREE Full text] [doi: 10.1901/jaba.1996.29-495] [Medline: 8995832]
38. Taj F, Klein MC, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: 10.2196/13311] [Medline: 31833836]
39. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]
40. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. Am J Public Health 2011 Apr;101(4):751-758 [FREE Full text] [doi: 10.2105/AJPH.2010.194381] [Medline: 21330590]
41. Eckerstorfer LV, Tanzer NK, Vogrincic-Haselbacher C, Kedia G, Brohmer H, Dinslaken I, et al. Key elements of mHealth interventions to successfully increase physical activity: meta-regression. JMIR Mhealth Uhealth 2018 Nov 12;6(11):e10076 [FREE Full text] [doi: 10.2196/10076] [Medline: 30425028]
42. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: a meta‐analysis of effects and processes. Adv Exp Soc Psychol 2006;38:119. [doi: 10.1016/S0065-2601(06)38002-1]
43. Gollwitzer PM. Implementation intentions: strong effects of simple plans. Am Psychol 1999 Jul;54(7):493-503. [doi: 10.1037/0003-066X.54.7.493]
44. Figueroa CA, Deliu N, Chakraborty B, Modiri A, Xu J, Aggarwal J, et al. Daily motivational text messages to promote physical activity in university students: results from a microrandomized trial. Ann Behav Med 2022 Feb 11;56(2):212-218. [doi: 10.1093/abm/kaab028] [Medline: 33871015]
45. Hamari J, Koivisto J, Sarsa H. Does gamification work? -- a literature review of empirical studies on gamification. In: Proceedings of the 47th Hawaii International Conference on System Sciences. 2014 Presented at: HICSS 2014; January 6-9, 2014; Waikoloa, HI. [doi: 10.1109/hicss.2014.377]
46. Mekler ED, Brühlmann F, Opwis K, Tuch AN. Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of common gamification elements. In: Proceedings of the First International Conference on Gameful Design, Research, and Applications. 2013 Presented at: Gamification '13; October 2-4, 2013; Toronto, Ontario. [doi: 10.1145/2583008.2583017]
47. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput Hum Behav 2017 Jun;71:525-534. [doi: 10.1016/j.chb.2015.08.048]
48. Antonaci A, Klemke R, Specht M. The effects of gamification in online learning environments: a systematic literature review. Informatics 2019 Aug 12;6(3):32. [doi: 10.3390/informatics6030032]
49. Schmidt-Kraepelin M, Thiebes S, Stepanovic S, Mettler T, Sunyaev A. Gamification in health behavior change support systems - a synthesis of unintended side effects. In: Proceedings of the 14th International Conference on Wirtschaftsinformatik. 2019 Presented at: WI2019; February 24-27, 2019; Siegen, Germany.
50. Nacke LE, Deterding S. The maturing of gamification research. Comput Hum Behav 2017 Jun;71:450-454. [doi: 10.1016/j.chb.2016.11.062]
51. Lieder F, Chen OX, Krueger PM, Griffiths TL. Cognitive prostheses for goal achievement. Nat Hum Behav 2019 Oct 19;3(10):1096-1106. [doi: 10.1038/s41562-019-0672-9] [Medline: 31427788]
52. Consul S, Stojcheski J, Lieder F. Leveraging AI for effective to-do list gamification. In: Proceedings of the 5th International Workshop "Gam-R – Gamification Reloaded". 2022 Presented at: 5th International Workshop "Gam-R – Gamification Reloaded"; September 04, 2022; Darmstadt, Germany URL: https://dl.gi.de/items/b17fae53-fdd1-4d91-a122-7df976acc37d [doi: 10.1515/icom-2021-0025]

53. Steel P. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. Psychol Bull 2007 Jan;133(1):65-94. [doi: 10.1037/0033-2909.133.1.65] [Medline: 17201571]
54. Murayama K. A reward-learning framework of knowledge acquisition: an integrated account of curiosity, interest, and intrinsic-extrinsic rewards. Psychol Rev 2022 Jan;129(1):175-198. [doi: 10.1037/rev0000349] [Medline: 35099213]
55. Xu L, Wirzberger M, Lieder F. How should we incentivize learning? an optimal feedback mechanism for educational games and online courses. In: Proceedings of the 41st Annual Meeting of the Cognitive Science Society. 2019 Presented at: CogSci 2019; July 24-27, 2019; Montreal, Canada.
56. Pauly R, Heindrich L, Amo V, Lieder F. What to learn next? aligning gamification rewards to long-term goals using reinforcement learning. In: Proceedings of the Multi-disciplinary Conference on Reinforcement Learning and Decision Making 2022. 2022 Presented at: RLDM; June 8-11, 2022; Providence, RI. [doi: 10.1002/9781118266502.ch3]
57. Callaway F, Jain YR, van Opheusden B, Das P, Iwama G, Gul S, et al. Leveraging artificial intelligence to improve people's planning strategies. Proc Natl Acad Sci U S A 2022 Mar 22;119(12):e2117432119 [FREE Full text] [doi: 10.1073/pnas.2117432119] [Medline: 35294284]
58. Wirzberger M, Lado A, Eckerstorfer L, Oreshnikov I, Passy J, Stock A, et al. How to navigate everyday distractions: leveraging optimal feedback to train attention control. In: Proceedings of the 42nd Annual Meeting of the Cognitive Science Society. 2020 Presented at: CogSci 2020; July 29-August 1, 2020; Online. [doi: 10.21203/rs.3.rs-3355487/v1]
59. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press; 2018.
60. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game on? Smoking cessation through the gamification of mHealth: a longitudinal qualitative study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]
61. Verplanken B, Orbell S. Reflections on past behavior: a self - report index of habit strength. J Appl Soc Pyschol 2006 Jul 31;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
62. Prentice M, Lieder F. Habit formation REG. Open Science Framework. 2022 Jun 8. URL: https://osf.io/h7jn8 [accessed 2024-02-09]
63. Harris MA. Maintenance of behaviour change following a community-wide gamification based physical activity intervention. Prev Med Rep 2019 Mar;13:37-40 [FREE Full text] [doi: 10.1016/j.pmedr.2018.11.009] [Medline: 30510892]
64. Dunton GF, Leventhal AM, Rebar AL, Gardner B, Intille SS, Rothman AJ. Towards consensus in conceptualizing and operationalizing physical activity maintenance. Psychol Sport Exerc 2022 Jul;61:102214. [doi: 10.1016/j.psychsport.2022.102214]
65. Lieder F, Chen PZ, Stojcheski J, Consul S, Pammer-Schindler V. A cautionary tale about AI-generated goal suggestions. In: Proceedings of the Mensch und Computer 2022. 2022 Presented at: MuC '22; September 4-7, 2022; Darmstadt, Germany URL: https://dl.acm.org/doi/fullHtml/10.1145/3543758.3547539 [doi: 10.1145/3543758.3547539]
66. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Rimer BK, Viswanath V, editors. Health Behavior: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass Inc; 2015.
67. Hovell M, Wahlgren D, Adams M. The logical and empirical basis for the behavioral ecological model. In: DiClemente RJ, Crosby RA, Kegler MC, editors. Emerging Theories in Health Promotion Practice and Research. San Francisco, CA: Jossey-Bass Inc; 2009.
68. Sallis JF, Owen N, Fisher EB. Ecological models of health behavior. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass Inc; 2008.
69. Bandura A. Self-Efficacy: The Exercise of Control. New York, NY: Worth Publishers; 1997.
70. Adams MA, Hurley JC, Phillips CB, Todd M, Angadi SS, Berardi V, et al. Rationale, design, and baseline characteristics of WalkIT Arizona: a factorial randomized trial testing adaptive goals and financial reinforcement to increase walking across higher and lower walkable neighborhoods. Contemp Clin Trials 2019 Jun;81:87-101 [FREE Full text] [doi: 10.1016/j.cct.2019.05.001] [Medline: 31063868]
71. Sukumar S, Ward AF, Elliott-Williams C, Hakimi S, Mozer MC. Overcoming temptation: incentive design for intertemporal choice. arXiv. Preprint posted online March 11, 2022 2023 [FREE Full text]
72. Mozer MC, Sukumar S, Elliott-Williams C, Hakimi S, Ward AF. Overcoming temptation: incentive design for intertemporal choice. In: Proceedings of the 3rd Multidisciplinary Conference on Reinforcement Learning and Decision Making. 2017 Presented at: RLDM '17; June 11-14, 2017; Ann Arbor, MI.
73. Prentice M, Lieder F, Chen PZ. Habit formation REG. Open Science Framework. 2022 Jun 5. URL: https://osf.io/5qg2n [accessed 2024-02-08]

Abbreviations

MDP: Markov decision process
SRHI: Self-Report Habit Index

Edited by G Eysenbach, T Leung; submitted 30.09.22; peer-reviewed by V Berardi, M Kapsetaki; comments to author 25.01.23; revised version received 12.06.23; accepted 31.08.23; published 22.03.24.

Please cite as:
Lieder F, Chen PZ, Prentice M, Amo V, Tošić M
Gamification of Behavior Change: Mathematical Principle and Proof-of-Concept Study
JMIR Serious Games 2024;12:e43078
URL: https://games.jmir.org/2024/1/e43078
doi:10.2196/43078
PMID:38517466

The Rationale Behind the Design Decisions in an Augmented Reality Mobile eHealth Exergame to Increase Physical Activity for Inactive Older People With Heart Failure

Aseel Berglund[1], PhD; Leonie Klompstra[1], PhD; Helena Orädd[1], MSc; Johan Fallström[1], MSc; Anna Strömberg[2,3], PhD; Tiny Jaarsma[2], PhD; Erik Berglund[1], PhD

1

2

3

Corresponding Author:
Aseel Berglund, PhD

Abstract

Physical activity is important for everyone to maintain and improve health, especially for people with chronic diseases. Mobile exergaming has the potential to increase physical activity and to specifically reach people with poor activity levels. However, commercial mobile exergames are not specially designed for older people with chronic illnesses such as heart failure. The primary aim of this viewpoint is to describe the underlying reasoning guiding the design choices made in developing a mobile exergame, Heart Farming, tailored specifically for sedentary older people diagnosed with heart failure. The goal of the exergame is to increase physical activity levels by increasing the daily walking duration of patients with heart failure by at least 10 minutes. The rationale guiding the design decisions of the mobile exergame is grounded in the thoughtful integration of gamification strategies tailored for application in cardiovascular care. This integration is achieved through applying gamification components, gamification elements, and gamification principles. The Heart Farming mobile exergame is about helping a farmer take care of and expand a virtual farm, with these activities taking place while the patient walks in the real world. The exergame can be adapted to individual preferences and physical condition regarding where, how, when, and how much to play and walk. The exergame is developed using augmented reality so it can be played both indoors and outdoors. Augmented reality technology is used to track the patients' movement in the real world and to interpret that movement into events in the exergame rather than to augment the mobile user interface.

(JMIR Serious Games 2024;12:e50066) doi:10.2196/50066

KEYWORDS

sedentary; exercise; exertion; exergames; technology; training; inactivity; eHealth application; heart disease; physical activity

Introduction

Background

Physical activity is important for everyone, especially for people with chronic disease [1]. Low levels of physical activity are commonly reported by patients with cardiac disease, which raises the risk of both illness progression and death [2-4]. Patients with heart failure benefit from physical exercise, but the activity must be tailored to the patient's baseline condition and the intensity of their symptoms [5]. A promising, innovative way to motivate people to become more physically active is the use of exergames. Exergames are games that combine gameplay with physical activity, requiring people to move in the real world [6]. They lead to an increase in physical activity [7] and have the potential to reach people with poor activity levels [8]. Nintendo Wii and the X-Box Kinect are well-known commercial exergames that have been tested for effectiveness in older people and for patients in cardiac rehabilitation programs [9-11]. Although some exergame studies show benefits for physical activity, a meta-analysis showed that exergaming, when compared to conventional cardiac rehabilitation programs, showed equal effects in terms of improving exercise capacity, quality of life, and mental health [11]. Nevertheless, commercially available exergames, using a "one size fits all" approach, fail to cater adequately to the specific requirements of the older adult population in general [12,13], and more specifically, to individuals dealing with heart failure [14]. A randomized controlled trial using a commercially available, off-the-shelf product demonstrated ineffectiveness in enhancing submaximal exercise capacity among patients with heart failure [15]. Based on interview studies regarding mostly commercial exergames, adult players perceived barriers to exergaming that might be related to the fact that the exergames were initially designed for children or young adults [14,16,17]. Barriers that could hinder older peoples' abilities to exergame were poor vision, physical function limitations, and other health conditions [14,17].

For this reason, the design of exergames for older people should focus on suitability, usability, and accessibility, as this will increase acceptance. An important aspect to consider in designing exergames for older people is that the location of play should be flexible, for example, offering the potential to be played both indoors and outdoors. This can be realized by developing the exergame for a smartphone, instead of a stationary computer. Another benefit is that smartphones are accessible to the target group. Even though older people might be interested in using mobile phones [18] and mobile health applications can help older people in self-care of illnesses and lifestyle improvement [19], there are no specific mobile exergames designed for older people with heart failure.

The aim of this paper is to describe the principles and technology used in Heart Farming, a mobile exergame designed with the objective of increasing physical activity for inactive older people with heart failure by increasing their walking time by at least 10 minutes each day, as this has been shown to reduce the risk of death or hospitalization by 4% [20]. The Heart Farming exergame was designed and developed using an adapted version of the player-centered, iterative, interdisciplinary, and integrated framework for serious games [21].

Mobile Exergames for Older People

A previous study shows that outdoor mobile exergames increase older peoples' engagement, satisfaction, and interest in physical activity [22]. Furthermore, previous research shows that the quality of life for patients with heart conditions can be improved by using mobile phones as health interventions to stimulate physical activities, medication management, dermatology, fall detection, and exercise monitoring [23]. Smartphones are recommended for non–face-to-face physical activity interventions for older people [24]. Smartphones contain built-in technologies, for example, accelerometers, cameras, and sensors, that measure users' movement and can provide the users with real-time monitoring and feedback on their physical activity without the use of a separate device [25-28]. Furthermore, smartphones provide connectivity, enabling users to get social support that can impact individual motivation positively and consequently improve goal attainment [26,29,30]. Flowie is one of the first mobile exergames to promote increased physical activity among older people and is designed to motivate them to enhance levels of physical activity using a virtual coach [31]. During an 11-day exploratory intervention, participants expressed positive feedback regarding the virtual coach, highlighting increased motivation to engage in exercise. However, quantitative data did not demonstrate a timely increase in physical activity, possibly due to the prototype's limited activity-sensing capabilities and varying weather conditions during the study. Another mobile exergame is Solitaire Fitness, an asynchronous exergame designed to enhance cognitive and physical ability for older people [32]. The exergame is built on a well-established card game, solitaire, since familiarity can impact the experience positively. The exergame gathers gameplay data so researchers and health professionals can assess project success by measuring user exercise levels and the frequency of features used in the exergame. The authors argue that the general attractiveness and engagement of the game might be improved, while offering health advantages, if new mechanisms are included while keeping it accessible and familiar.

To provide older people with an exergame experience that is perceived as meaningful play, it is important to understand their needs and why they would want to play the exergame [33]. Exergames need to be both attractive to the players as well as effective as an exercise [34]. There are several guidelines for game design [35], exergame design [14,36], and fitness applications [37], as well as motivational factors for mobile games for older adult users [38], all of which need to be considered when designing exergames for older people with heart failure. For example, the benefits gained from playing the exergame need to be clear since they will impact older adults' use of the exergame [38]. Furthermore, the exergame topic should be adjusted to the interests of older people [36,39,40], and the players' physical condition should be considered with a suitable difficulty setting [36,39]. Furthermore, the user interface should be adapted to older people [36,39] by, for example, having an easy-to-use user interface and avoiding small objects by, for example, using bigger game characters [36]. Feedback should be both auditory and visual [36], and no feedback should be provided about unachieved goals or unperformed activities, since only positive reinforcement should be used to improve behaviors in exergames for older people [40].

Gamification Features in Exergames

Overview

Gamification can be used to enhance user motivation and engagement with systems [41] by using game design to create gamified experiences [42]. Therefore, gamification features can be integrated in exergames [43]. Matallaoui et al [43] consider exergames to consist of 3 gamification components [41,42] that can be applied for people with heart failure together with gamification principles and gamification elements [44].

Gamification Components

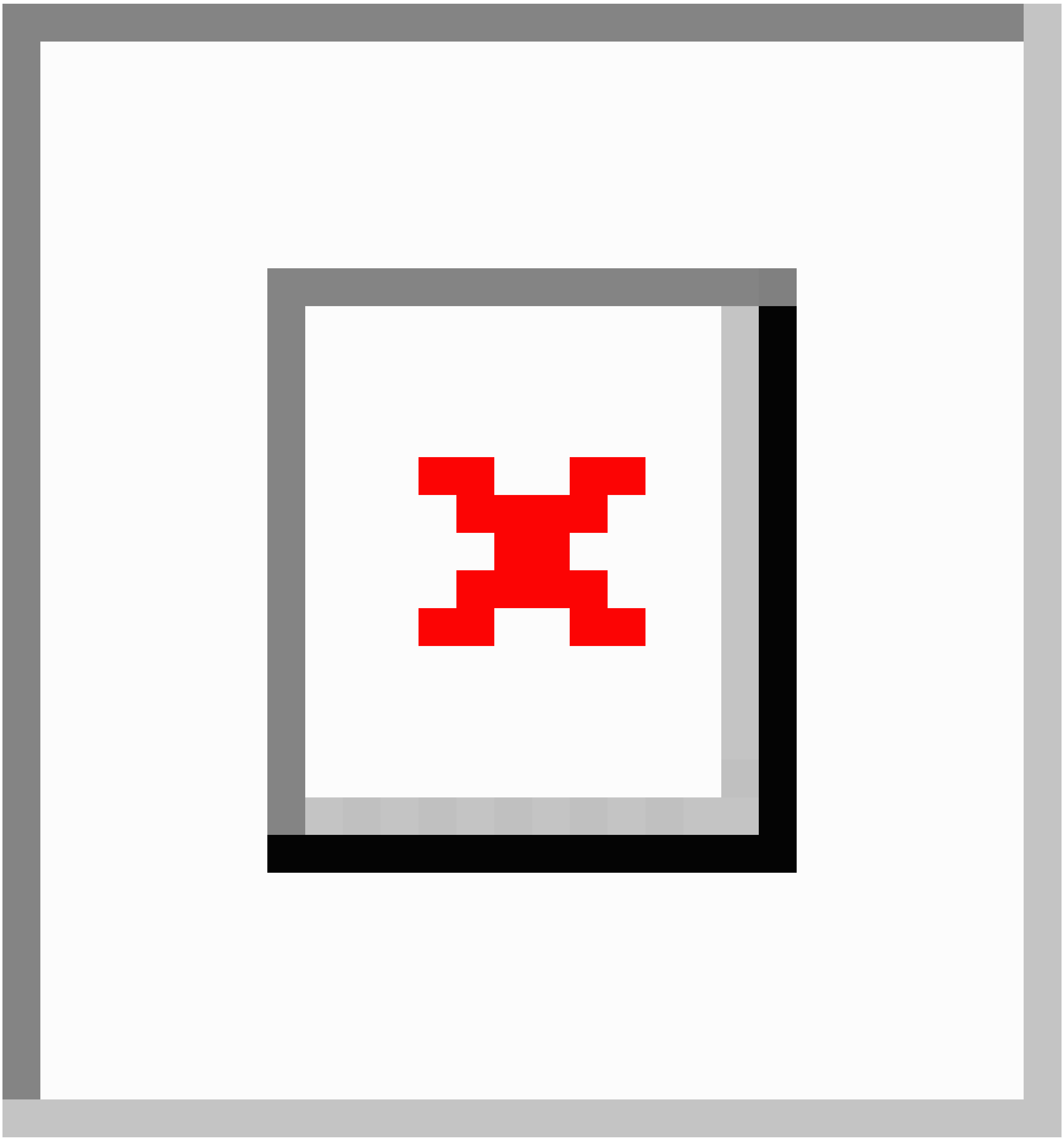
Gamification can be presented as a process with 3 main components, and these components are the fundamental building blocks that make up a gamified experience [41,44] (Figure 1):

1. *Implemented motivational affordance*: Objects in the user interface of the exergame have properties that allow the players to experience the satisfaction of their psychological, internal, motivational needs. These are used to motivate users to engage with a gamified experience and provide users with a sense of challenge, accomplishment, and reward.
2. *Psychological outcomes*: These are the psychological experiences promoted by the exergame to motivate the user to perform activities. These are the emotional or cognitive effects that gamification can have on users. Examples of psychological outcomes include increased engagement and a sense of accomplishment.
3. *Behavioral outcomes*: These are the behaviors that the exergame will change and that will affect the player's health. These are the specific actions and behaviors that users take as a result of engaging with a gamified experience. Examples of behavioral outcomes include

increased participation, the completion of tasks, and adherence to a desired behavior.

The gamification process starts with motivational affordances that facilitate the psychological experiences that will, together with the implemented affordances, in turn influence the behavioral outcomes (ie, the physical activity for exergames).

Figure 1. Gamification components adapted from Berglund et al [44], which is published under Creative Commons Attribution 4.0 International License [45].



Gamification Principles

Gamification principles are the underlying concepts and strategies that guide the design and implementation of gamification. These principles are the fundamental rules that govern how a game is played and are used to create and extend the player experience [46] and increase overall enjoyment. There are 5 principles that impact the gamification process [47]:

1. *Meaningful purpose*: The goal of the game should align with the users' motivations [47].
2. *Meaningful choice*: Players have agency over how they achieve their goals and make progress in the game [47] to satisfy their motivational needs [48].
3. *Supporting player archetypes*: The game's mechanics allow for individual play styles and characteristics [47].
4. *Feedback*: Feedback on how actions performed in the game affect the player's progress is clearly communicated to the

player [47]. Instant feedback makes the game responsive and helps maintain the game's appeal [49]. Accumulated feedback provides information about the progress of players, which makes it suitable for comparison [49] and can be a source of motivation when players evaluate how close their current behavior is in relation to the goal, thus facilitating self-monitoring [50].

5. *Visibility*: The player can see the status of their current progress, including how much remains to meet their goals or finish a task [47].

Gamification Elements

Gamification elements are the specific parts that make up the game. The main difference between gamification principles and gamification elements is that gamification principles define how a game works, are used to create an engaging and fun gameplay experience for players, and refer to the overarching design concepts that guide the creation of a game, whereas gamification elements are the specific implementations used to achieve the principles. The 6 most common gamification elements used in health care are as follows [51]:

1. *Leaderboards*: Leaderboards (or high score lists) present the rank of players based on their experience points [51]. Leaderboards enable players to compare themselves with others and evaluate their performance [52].
2. *Levels of achievement or rank*: These are sums of points or credits associated with levels. Levels could evoke a feeling of progression [51].
3. *Digital rewards*: Rewards for achievements in the game can have a positive impact on interest and enjoyment [53] and influence the player's behavior in different ways [54]. There are 8 forms of rewards [49]:
   - 3.1. *Score systems*: Numbers are used to mark player performance.
   - 3.2. *Experience points reward systems*: Experience points are earned during gameplay and lead to "leveling up" when specified goals are achieved.
   - 3.3. *Item-granting system rewards*: These are virtual items that can be used by players.
   - 3.4. *Resources*: Virtual resources can be collected and used in the game in a manner that affects gameplay.
   - 3.5. *Achievement systems*: Players can collect achievements by fulfilling clearly stated conditions.
   - 3.6. *Feedback messages for immediate rewards*: Some examples include text, sound effects, and animations.
   - 3.7. *Plot animations and pictures*: These are provided after important events in the game such as clearing a new level.
   - 3.8. *Unlocking mechanisms*: Players are rewarded by getting access to game content, for example, new levels achieved when the required experience points are gained.
4. *Real-world prizes*: Players can exchange credits or points earned in the exergame for real-world prizes such as vouchers, goods, or services.
5. *Competitions*: Competitions can be invoked through comparisons with oneself or with peers [50]. Competition appeals to the player's social needs, which impacts their intrinsic motivation [48].
6. *Social or peer pressure*: Cooperation appeals primarily to the player's psychological need for relatedness [48] and can complement competition [50].

Exergames and Augmented Reality

Augmented reality (AR) is a technology that involves the use of sensors and cameras to collect and interpret data from the surrounding environment. AR combines virtual information with the real world and can be used for intelligent displays, 3D registration technology, or intelligent interaction technology [55]. AR technology is usually used to present information (eg, graphics, video, or sound) in a dynamic and interactive manner by overlaying digital content onto the real-world environment [56], and it is highly recommended for health applications that bridge the virtual and real worlds [57]. Furthermore, AR has been used for indoor navigation [28,58,59] by using the technology to determine the user's position in the physical environment. AR can support localization, but for the technology to work properly, users need to hold the smartphone device in a specific orientation and position [59].

The application of AR in fitness is a novel approach for patient intervention in clinical environments, promotes physical activity, and protects the public from noncommunicable illnesses [60]. Furthermore, AR has the potential to play a role in encouraging and sustaining an active lifestyle, fostering overall well-being among the older adult population [61]. Since AR mobile exergames have the potential to enhance cardiovascular health and physical exercise, they are appropriate for person-centered care interventions [62]. AR technologies have the potential to complement physical interventions in altering human behaviors and influencing physical activity [63]. However, there are several challenges to successfully implement AR techniques in exergames for older people, and usability is one of the challenges. According to a study by Stamm et al [60], older people found it challenging to use complex gestures and navigate menus in AR exergames. Another challenge is safety, and the danger of falling has been identified as a safety challenge that impacts patients' use of AR exergames [60]. The acceptance of technology is another challenge that needs to be considered in AR exergames [60,64].

To create exergames that are fun to play, the physical environment needs to be considered, since it influences physical activity [65] and the appropriate technology that needs to be used. AR has been used for both outdoor [66] and indoor [67] exergaming. GPS, a satellite-based radio navigation system, is another technology used in many mobile location-based exergames intended for outdoor use [8,68,69]. The error margin of most mobile GPS trackers is in the 5- to 10-m range, which means registering shorter distances when walking back and forth (as people playing in smaller areas might have to do) is not possible. However, GPS has a reduced signal strength or lack of signals in indoors environments due to the fact that physical obstructions such as buildings may weaken and even block GPS signals [70,71]. Therefore, GPS is more suitable for outdoors positioning while AR can be used both indoors [72,73] and outdoors [74]. Using the camera for movement tracking results in spatial movement detection nearly 10 times more precise than that achieved by a GPS sensor [58]. AR and GPS can be combined in exergaming. One of the most well-known mobile exergames using both GPS and AR is Pokémon GO, which was shown to increase physical activity [8,68]. Pokémon GO increased user steps by an average of 1473 steps per day over a period of 30 days [8]. However, Pokémon GO has been criticized for the potential harm it could cause via traffic accidents due to players' distraction or through players placing themselves in dangerous situations or environments while playing [75,76].

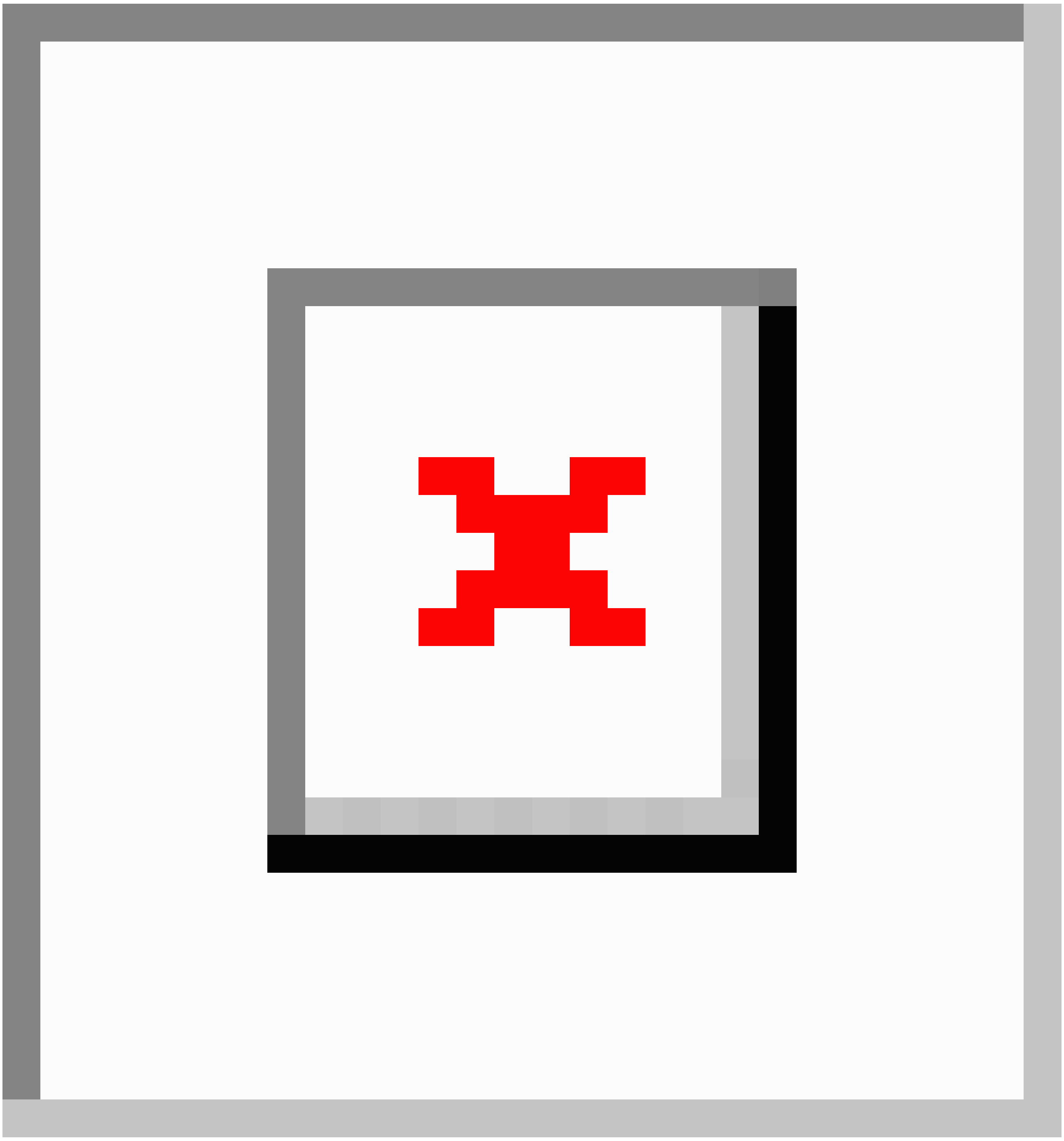
In summary, older people with heart failure often lead sedentary lifestyles. AR mobile exergames have the potential to encourage increased physical activity, yet the current lack of mobile exergames designed for both indoor and outdoor use, incorporating gamification features to influence patient behavior, presents a notable gap in the available solutions. In this viewpoint, we describe the rationale behind the design decisions in Heart Farming, a mobile exergame tailored for older people with heart failure.

Heart Farming: A Mobile Exergame

Overview

The Heart Farming exergame uses a farming theme, which involves helping a farmer take care of and expand a farm (Figure 2A and B). The players can plan activities in the exergame and then walk in the physical world, indoors or outdoors, to perform the activities they have planned without the need to look at the screen while walking (Figure 2B and C). Available activities are planting and collecting crops of various kinds (eg, sowing carrots, watering strawberries, and harvesting oranges), buying new fields, foraging in the forest, fishing, and selling crops to neighbors on the farm (see Figure 2A-C). Fields and neighbors are visible on a map (Figure 2A and B), and the player must walk in the real world, both to perform the planned activities and to move between locations. As players take care of the farm, they gain experience points, which, over time, unlock new types of crops to farm as well as new neighbors to sell to.

Figure 2. Images from the Heart Farming exergame: (A) main view of the exergame showing the farm before any planning of activities, (B) farm scene as the player plans activities, (C) planned activities are shown in a list view with a progress bar as the player walks, and (D) player statistics and progress in the exergame.



Gamification Features in Heart Farming

The gamification components applied in Heart Farming are as follows:

1. *Implemented motivational affordance*: The desired behavior of being physically active is paired with rewards, for example, numeric rewards such as walked distance and experience points. Walked distance builds up to complete the daily goal and is only earned while performing activities in the Heart Farming exergame that require walking. Sums of experience points are associated with levels, which evoke a feeling of progression (Figure 2D). The player gets experience points by performing activities in the exergame, which are based on how many meters of movement the AR system has registered, balanced to match the player's own physical activity levels. Experience points can be earned from all activities, including sit-to-stand activities, and can be given as a reward, for example, for reaching the daily goal.
2. *Psychological outcomes*: The purpose is to increase players' motivation and self-efficacy for physical activity, since these factors are important for patients with heart failure [77]. It is common for patients with low exercise capacity to be motivated but lack the confidence to become more physically active [77]. Achieving daily goals in the Heart

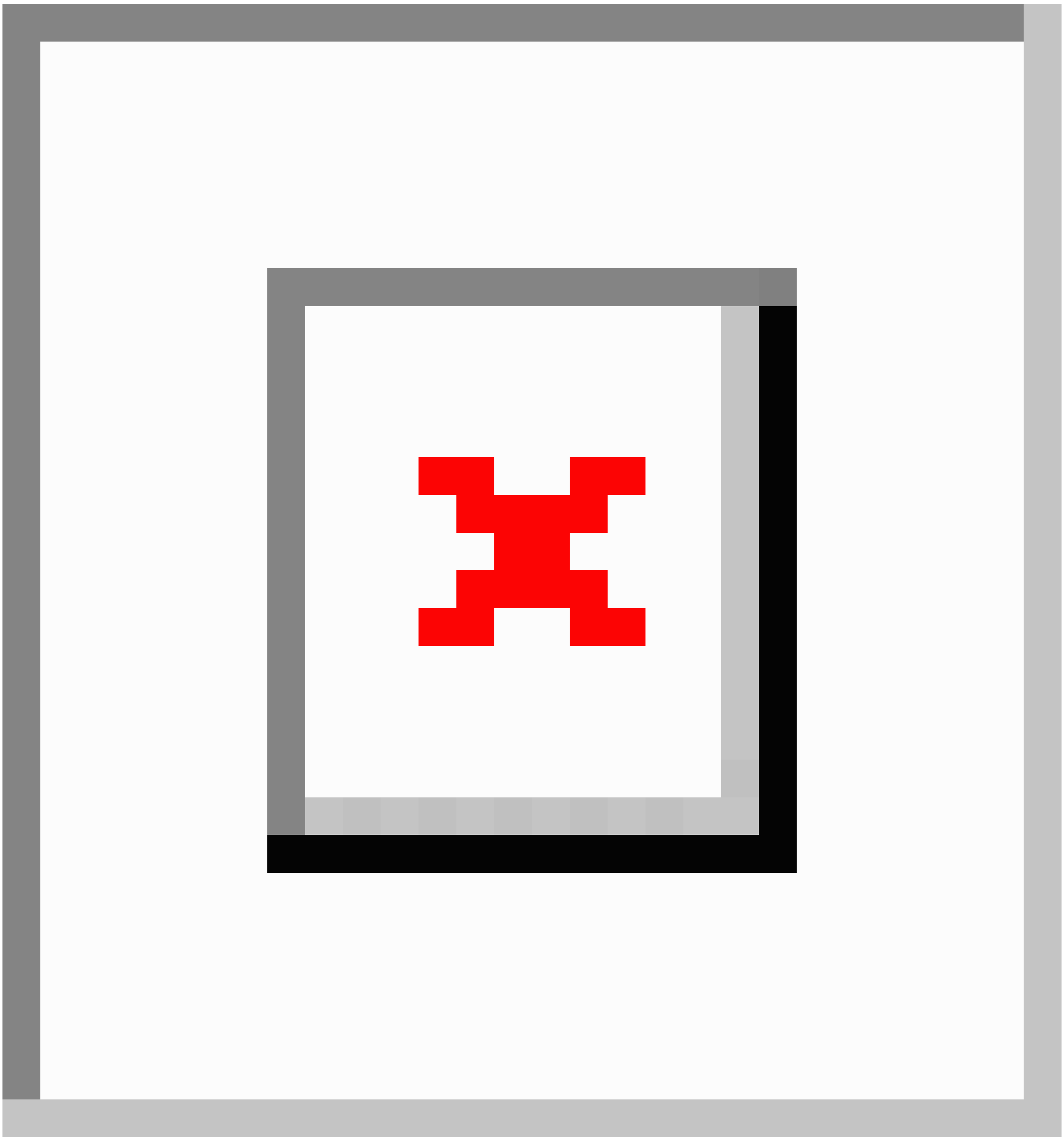
Farming exergame means that patients are following the physical health recommendation, which might increase their confidence in walking.

3. *Behavioral outcomes*: The goal is to increase physical activity by walking for 10 minutes every day with the Heart Farming exergame. To create a trigger for the player to help create the habit of playing every day, the exergame logs whether players have played or not and sends a reminder at 12 PM if they have not yet played that day. A mission book is also included to give players regular missions; the missions trigger them to do specific activities in the exergame and encourage curiosity about later content (see Figure 3D).

The gamification principles applied in Heart Farming are as follows:

1. *Meaningful purpose*: Real-life physical activity is used to progress in Heart Farming. Players can do 2 types of physical activities: walking in real life and doing sit-to-stand activities. Both activities are performed using the smartphone, which registers the movements and converts them into virtual exergame progress and rewards. Walking is required to meet the player's daily goal.
2. *Meaningful choice*: Players can choose freely how, when, and where to play. Players who are not interested in optimizing the rewards can choose the automatic mode, which only requires the press of a button before the patient can begin to walk. Players who want more challenges and to perform well in the exergame can instead use the strategic gameplay mode to plan farming activities, which are chosen freely with little or no real analysis of the exergame rules and progress. How players choose to plan their tasks affects the number of rewards they get. Players can also choose to play to complete missions in a mission book. Completing the missions is not obligatory to progress in the exergame, and many tasks in the missions can be completed in more than 1 way to give players choices in how to play the exergame, while others are more specific to encourage players to try new things or extend their farm further. An optional sit-to-stand activity that gives the players experience points or crops is also included to facilitate players with varying degrees of mobility.
3. *Supporting player archetypes*: The goal distance to walk each day is adapted to the players' exercise capacity. Players are represented in Heart Farming by an avatar in the form of a hat seen from above. As players progress in the exergame, new hats are unlocked, so players can choose a hat that they can identify themselves with. The avatar reflects in-game activities conducted by the player, such as moving on the map when the player walks. Heart Farming is also designed to be adaptable to the players' exercise capacity. The goal of 10 minutes of walking per day can be based on the number of meters walked during a 6-minute walk test [78]. When players walk their daily goal in the exergame, they receive a trophy. When they have walked twice or 3 times their daily goal, they are rewarded with a smaller experience point boost and a flag animation. After the player has walked 3 times their daily goal, the walking distance is registered and displayed to the player, but in a toned-down fashion without motivational components such as rewards and with the progress bar completely filled.
4. *Feedback*: The players receive multiple types of immediate feedback for performing actions in Heart Farming that communicates clearly how their actions affect progress in the exergame. Instant feedback is provided as visual and auditory information about the behavior that is currently being performed. For example, while walking with a plan in motion, the player can see a progress bar fill up on the current task and hear sounds corresponding to the action being performed (eg, watering sounds while watering a field; see Figure 2C).
5. *Visibility*: All progress in the form of number of meters walked is displayed visually so players can see their overall progress in the Heart Farming exergame as well as their current progress on their daily goal at any time (see Figure 2B-D).

Figure 3. Digital rewards in Heart Farming: (A) score system, (B) experience points, (C) daily gifts, (D) achievement systems, and (E) crops that can be sold to a neighbor.



The gamification elements applied in Heart Farming are as follows (Figures 3 and 4):

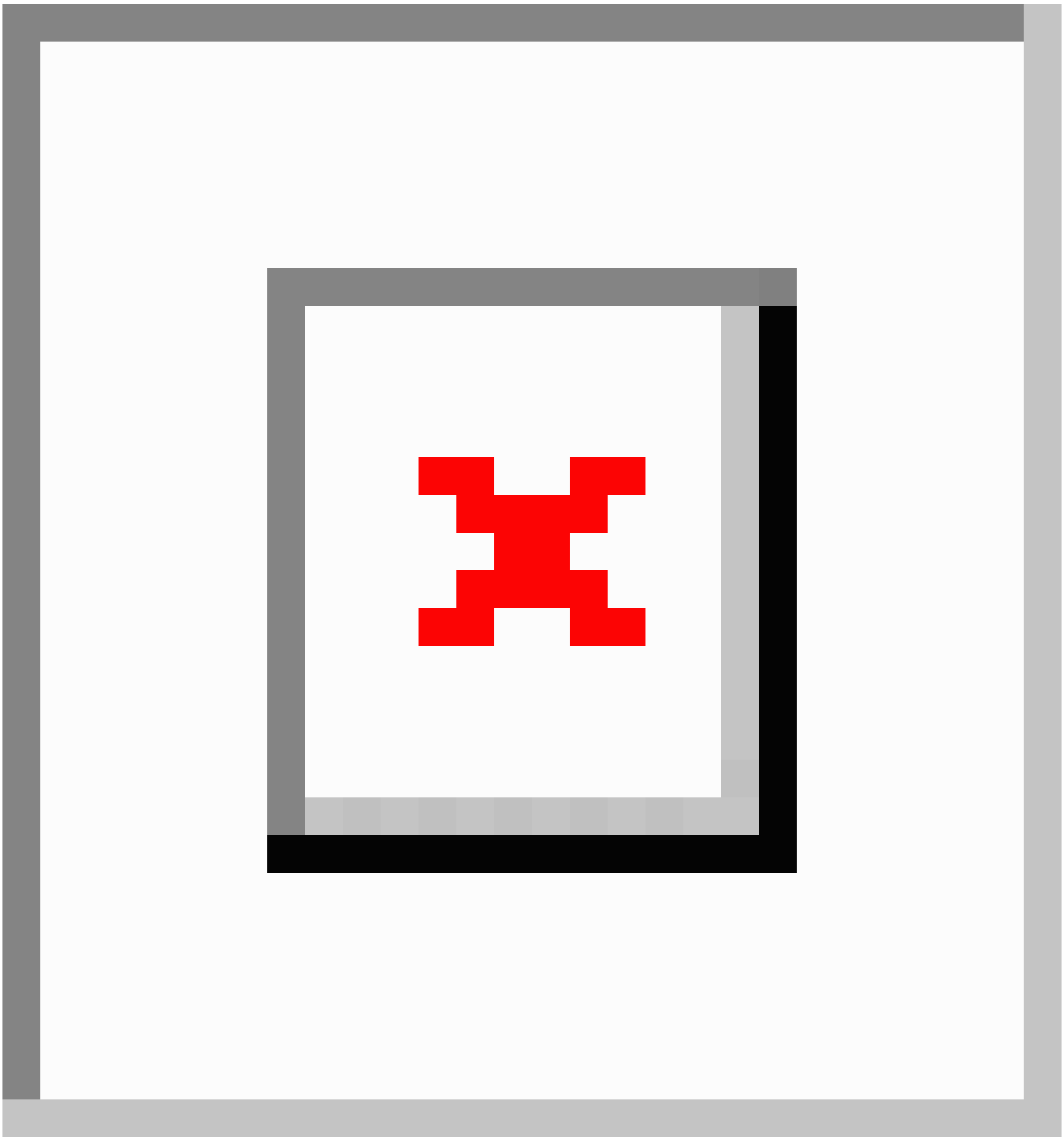
1. *Leaderboards*: A leaderboard displays the in-game postal codes of the players that gained the most experience points the previous day (Figure 4A). Since low-performing players can be demotivated by a leaderboard when they fall behind and feel that they cannot catch up with others [79], our leaderboard is designed so that players have the same opportunity to be included in it every day, regardless of their performance in the days prior to the current one.
2. *Levels of achievement or rank*: Sums of experience points are associated with levels that evoke a feeling of progression . Each level also unlocks a new type of crop to farm or forage, a new tool that makes farming easier , a new neighbor to whom the player can sell crops (), or a new decorative hat for the farmer to use (Figure 4B). The actual level is always displayed in the upper left corner (Figure 3).
3. *Digital rewards*: The following digital rewards are included in Heart Farming exergame (Figure 3):
   - 3.1. *Score systems*: The total distance walked adds up to reach the daily goal. It is displayed as a progress bar in the lower part of the screen (Figure 3A).
   - 3.2. *Experience points reward systems*: Each activity in the exergame is translated to experience points, which are required for the player to level up over time.

A star symbol is used for experience points (Figure 3B).

- 3.3. *Item-granting system rewards*: Every day, players get a gift of free crops to encourage them to come back to the farm each day (Figure 3C). This incentive fosters a daily habit of playing, which in turn promotes walking every day.
- 3.4. *Resources*: Coins are needed to buy new fields and are earned from selling crops (Figure 3A).
- 3.5. *Achievement systems*: Completing specific tasks unlocks unique animals. Each neighbor has hearts that are filled by selling them specific crops, and if all hearts are filled before the end of the week, the player gets a special decoration (Figure 3D).
- 3.6. *Feedback messages for immediate rewards*: Both animations and sound effects are used to provide feedback when the players perform actions. When players complete the daily goal, they get a reward in the form of an animated trophy and a boost of extra experience points.
- 3.7. *Plot animations and pictures*: These were not used.
- 3.8. *Unlocking mechanisms*: The players unlock new tools and crops at higher levels (Figure 4B).

4. *Real-world prizes*: The exergame does not include any real-world prizes.
5. *Competitions*: The exergame does not include any competition.
6. *Social or peer pressure*: Each player has their own in-game postal code that can be used to communicate with other players by sending emojis (eg, smile messages) with positive reinforcement and packages with crops to help each other progress in the exergame (Figure 4C).

Figure 4. Gamification elements in Heart Farming: (A) leaderboards, (B) levels, and (C) social or peer pressure.

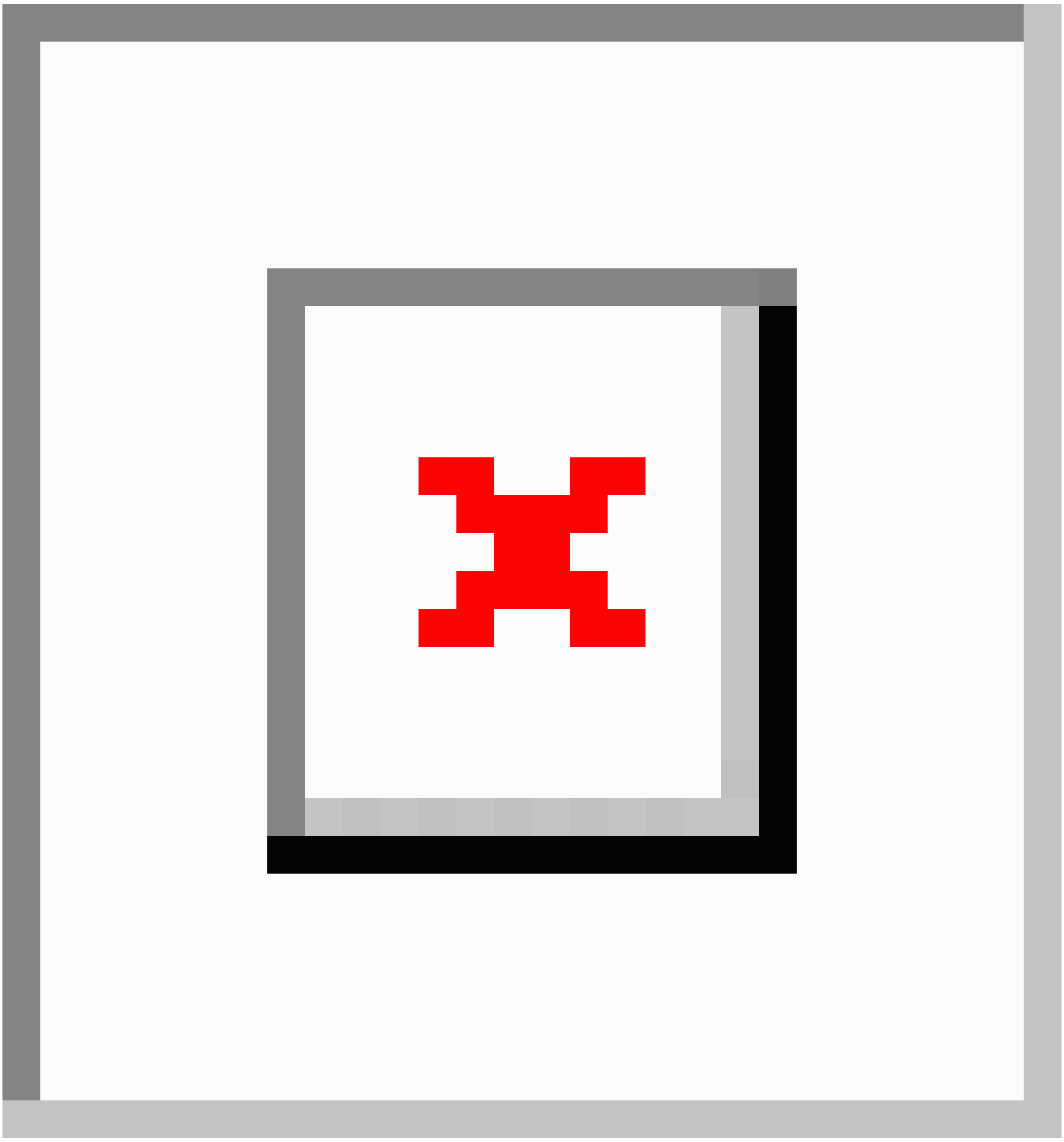


The Exergame Development

The game development engine Unity was used for the exergame development. The player data are stored in an SQLite database that is connected to a Flask server. The exergame was developed for both the Android and the iOS smartphone operating systems that support AR. The AR is only used to track the players' movements when walking in the real world. AR enables acceleration-independent tracking by analyzing visual discrepancies among identified landmarks by calculating the change through the AR program's SLAM solution [80,81]. The status of the AR program is displayed for the player using a stop traffic light with 2 colors: a green light means that the AR can track and register the player's movement so the player can walk, whereas a red light means that the AR is not working. A red light may either indicate that the AR program is loading or that it is unable to analyze the current camera feed for some reason, for example, it is too dark, the player is moving too fast, or the camera feed is covered. A warning pop-up is displayed for the player and proposes a solution based on the problem. If the camera feed is too dark, the player is prompted to relocate to a brighter place, and if the movement speed is too high (biking or driving speed), the player is prompted with 1 of the 3 different messages: "bad camera image (dark/covered)," "too fast motions for camera," and "initializing/restarting tracking" (Figure 5).

To mitigate the risk of falling, the player does not need to look at the screen while walking. The player creates a plan with activities and can then walk to perform all activities in the plan without the need to look at the screen. Each activity has a sound to provide feedback for the player about what is happening on the farm.

Figure 5. (A) Green traffic light and (B-D) red traffic light with warning messages.



Discussion

Principal Findings

Research indicates that older people with heart failure can potentially reduce the risk of death or hospitalization by 4% simply by adding a daily 10-minute walk to their routine [20]. In response to this, we developed a mobile exergame, Heart Farming, specifically designed to increase their physical activity levels. The rationale that directed the design decisions of the mobile exergame described in this work is based on well-established game design methods to support behavioral changes for inactive patients with heart failure in an enjoyable way. Gamification components, gamification principles, and gamification elements have been applied for designing the exergame. Heart Farming is based on a farming theme, and to play the game, players need to walk in the real world.

Exergames can persuade people to exercise more when they are designed for exercising [82], but defining appropriate goals is difficult, and therefore, official institutional guideline values can be used appropriately as a starting reference point [64]. Since regular physical activity can help preserve health and enhance the quality of life for the patient group, the exergame is designed so that the patient comes back to the exergame and plays every day, for example, using daily gifts with different crops and daily reminders when the patient has not walked.

The gamification components in Heart Farming are provided to increase the patients' physical activity, which requires behavioral changes. Since motivation drives humans to perform and can arise from physiological or psychological needs, thoughts, or emotions [83], the design and technical implementation is applied to balance the patients' physical needs with an elusive fun factor. To perform a behavior, a person must be motivated, have sufficient ability, and have a trigger to prompt the behavior [84]. Thus, the design of Heart Farming aims to increase the motivation to engage in physical activity and keep it at a level where the patients' abilities are sufficient to perform the activity. For the trigger, the exergame sends a reminder at 12 PM if the player has not yet played that day. By setting the reminder around lunch time, the patient can associate playing the exergame with an existing habit of eating lunch, making it easier to create a new habit either before or after eating. According to Stawarz et al [85], reminders can, however, hinder habit formation if the user gets reliant on the reminders, which causes the habit to die out if the user stops using the app. By only sending reminders to players who have not played yet, only those who need the reminder get them, and as they start connecting a trigger event in their lives to the new habit, the reminders might phase out naturally over time. A potential problem with having a set time for the reminders is that the time might not work for some patients, as they might be busy or might not have any trigger events such as having lunch around that time. Allowing patients to set their own reminder time might make it easier for them to form a habit.

The only way for players to customize their avatar is to change the hat that is visible from above. Furthermore, the players only unlock new hats after reaching higher levels and thus have very limited choices in the beginning. This could make it harder to relate to and identify with the avatar, which has been shown to be important for games that aim to change behavior [82]. However, changing to a more human avatar could raise issues regarding how much personalization is needed to ensure that nobody feels alienated by the available choices. It may therefore be better to stick to an avatar where the patient must imagine what the farmer looks like on their own.

To avoid endangering the patients' health and to avoid physical overexertion, as demonstrated with Pokémon GO [86], players will not be encouraged to walk after reaching 3 times their daily goal, but they are not prohibited from continuing to walk if they wish to exercise further, and the data will still be collected and stored by the game. This allows for individual play styles and characteristics, which has been identified as an important principle [44]. Players can also decide which play style they want to use when playing the exergame (automatic plan or strategic plan).

Players that wish to compare themselves to others as a way of challenging themselves to do better can use the leaderboard to see the scores from other players from the previous day and work toward reaching the top of the board. By resetting scores daily and collecting scores 1 day at a time, new players or those who have not played the exergame in a while still have an opportunity to top the leaderboard. Players can compare their results by looking at the leaderboard or by talking to other players about their scores and progress. Players can also interact with other players in the exergame by sending messages. The possibility to play with others has been identified as a desired feature in exergaming for patients with heart failure [14,17]. The question of whether sending messages is an optimal format for interaction should be further explored.

To not endanger the patients' health, the AR technique is used to track the players' movements when walking rather than to enhance the real world. The AR technology is used to detect the players' motion in the real world so players can choose to be physically active indoors, outdoors, or both. Using AR technology may have some practical constraints. When using the AR as a tool for motion detection, it is important that the players understand why and how the camera in the smartphone is used, so they feel confident when they play and can interact with the exergame as required. For Heart Farming, the players can play the exergame by either holding the smartphone in front of them or hanging the smartphone in a bag around the neck. In both cases, the front-facing camera must have a clear view of the surroundings. However, the player does not need to look at the screen while walking, which mitigates the risk of falling. The players get audio feedback informing them about the performed activities. AR is not the only technique available for tracking the players' movement. GPS and step counting are 2 possible alternatives, both of which could be used even if the smartphone is kept inside a pocket. Step counting is an offline data stream that is reported occasionally, so it is not directly accessible by the system for instant feedback. Step counting also does not provide good data for users that move slowly or generate little vibration from their steps and is therefore unsuitable for the target patient group. Since GPS lacks precision or fails entirely indoors, it is not suitable for indoor use. However, GPS could be incorporated for outdoor use only, with AR still used for indoor movement, although making the choice between the technologies in an easy way could be challenging.

The effects of playing the Heart Farming exergame will be investigated in a randomized control trial with patients with heart failure, and the data are currently being collected. In the study, the patients' experiences from playing the exergame will also be investigated.

Conclusion

Heart Farming is a mobile exergame tailored to increase physical activity in inactive older people with heart failure by encouraging them to walk for at least 10 minutes each day. The Heart Farming exergame is about helping a farmer take care of and expand a virtual farm. The advantages of the proposed exergame can be summarized as follows:

- The Heart Farming exergame detects patients' physical activities by using AR technology to detect patients' movements in the real world, both indoors and outdoors.
- The AR technology in the Heart Farming exergame continuously registers patients' movements in the real world and interprets them into events in the exergame, which then generates both visual and auditory feedback.

- The Heart Farming exergame can support patients' own individual preferences and conditions regarding where (indoors or outdoors), how (automatic or strategic gameplay), how much (patients are rewarded for once, twice, or 3 times their daily goal, and after that, progress is registered and displayed without rewards), and when (patients are remined at 12 PM if they have not played that day) to play.

Acknowledgments

The authors would like to thank the the Swedish National Science Council (2020-01109); The Swedish Heart and Lung Foundation (20200190); the Kamprad Foundation (20210074) and the Research Council in South East Sweden (FORSS-940933 and FORSS-995300) for their support of this work.

Authors' Contributions

TJ, AB, EB, LK, and AS acquired the funding for the study. The conception of the study was carried out by AB, AS, EB, LK, and TJ. The design was collaboratively developed by AB, AS, EB, HO, JF, LK, and TJ. AS and EB were responsible for the software development and analysis, while the interpretation of data was conducted by AB, AS, EB, HO, LK, and TJ. The initial draft of the work was written by AS and EB. Substantive revisions were made by AS, EB, HO, JF, LK, and TJ. Finally, the submitted version was reviewed and approved by AB, TJ, HO, JF, AS, and EB.

Conflicts of Interest

None declared.

References

1. Pedersen BK, Saltin B. Exercise as medicine–evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports 2015 Dec;25 Suppl 3:1-72. [doi: 10.1111/sms.12581] [Medline: 26606383]
2. Moholdt T, Wisløff U, Nilsen TIL, Slørdahl SA. Physical activity and mortality in men and women with coronary heart disease: a prospective population-based cohort study in norway (the HUNT study). Eur J Cardiovasc Prev Rehabil 2008 Dec;15(6):639-645. [doi: 10.1097/HJR.0b013e3283101671] [Medline: 18779734]
3. Stewart RAH, Held C, Hadziosmanovic N, et al. Physical activity and mortality in patients with stable coronary heart disease. J Am Coll Cardiol 2017 Oct 3;70(14):1689-1700. [doi: 10.1016/j.jacc.2017.08.017] [Medline: 28958324]
4. Jeong SW, Kim SH, Kang SH, et al. Mortality reduction with physical activity in patients with and without cardiovascular disease. Eur Heart J 2019 Nov 14;40(43):3547-3555. [doi: 10.1093/eurheartj/ehz564] [Medline: 31504416]
5. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021 Sep 21;42(36):3599-3726. [doi: 10.1093/eurheartj/ehab670] [Medline: 34447992]
6. Mueller F, Khot RA, Gerling K, Mandryk R. Exertion games. FNT Hum-Comput Interact 2016 Dec 29;10(1):1-86. [doi: 10.1561/1100000041]
7. Kari T. Can exergaming promote physical fitness and physical activity?: a systematic review of systematic reviews. Int J Gaming Comput-Mediat Simul 2014;6(4):59-77. [doi: 10.4018/ijgcms.2014100105]
8. Althoff T, White RW, Horvitz E. Influence of Pokémon GO on physical activity: study and implications. J Med Internet Res 2016 Dec 6;18(12):e315. [doi: 10.2196/jmir.6759] [Medline: 27923778]
9. Verheijden Klompstra L, Jaarsma T, Strömberg A. Exergaming in older adults: a scoping review and implementation potential for patients with heart failure. Eur J Cardiovasc Nurs 2014 Oct;13(5):388-398. [doi: 10.1177/1474515113512203] [Medline: 24198306]
10. Zeng N, Pope Z, Lee JE, Gao Z. A systematic review of active video games on rehabilitative outcomes among older patients. J Sport Health Sci 2017 Mar;6(1):33-43. [doi: 10.1016/j.jshs.2016.12.002] [Medline: 30356538]
11. Blasco-Peris C, Fuertes-Kenneally L, Vetrovsky T, Sarabia JM, Climent-Paya V, Manresa-Rocamora A. Effects of exergaming in patients with cardiovascular disease compared to conventional cardiac rehabilitation: a systematic review and meta-analysis. Int J Environ Res Public Health 2022 Mar 15;19(6):3492. [doi: 10.3390/ijerph19063492] [Medline: 35329177]
12. Wollersheim D, Merkes M, Shields N, et al. Physical and psychosocial effects of Wii video game use among older women. Int J Emerg Technol Soc 2010;8(2):85-98 [FREE Full text]
13. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
14. Klompstra L, Jaarsma T, Mårtensson J, Strömberg A. Exergaming through the eyes of patients with heart failure: a qualitative content analysis study. Games Health J 2017 Jun;6(3):152-158. [doi: 10.1089/g4h.2016.0087] [Medline: 28422516]

15. Jaarsma T, Klompstra L, Ben Gal T, et al. Effects of exergaming on exercise capacity in patients with heart failure: results of an international multicentre randomized controlled trial. Eur J Heart Fail 2021 Jan;23(1):114-124. [doi: 10.1002/ejhf.1754] [Medline: 32167657]
16. Ning H, Jiang D, Du Y, et al. Older adults' experiences of implementing exergaming programs: a systematic review and qualitative meta-synthesis. Age Ageing 2022 Dec 5;51(12):afac251. [doi: 10.1093/ageing/afac251] [Medline: 36571772]
17. Cacciata MC, Stromberg A, Klompstra L, et al. Facilitators and challenges to exergaming: perspectives of patients with heart failure. J Cardiovasc Nurs 2022;37(3):281-288. [doi: 10.1097/JCN.0000000000000825] [Medline: 34091566]
18. Zhou J, Rau PLP, Salvendy G. Age-related difference in the use of mobile phones. Univ Access Inf Soc 2014 Nov;13(4):401-413. [doi: 10.1007/s10209-013-0324-1]
19. Debon R, Coleone JD, Bellei EA, de Marchi ACB. Mobile health applications for chronic diseases: a systematic review of features for lifestyle improvement. Diabetes Metab Syndr 2019;13(4):2507-2512. [doi: 10.1016/j.dsx.2019.07.016] [Medline: 31405669]
20. Conraads VM, Spruit MA, Braunschweig F, et al. Physical activity measured with implanted devices predicts patient outcome in chronic heart failure. Circ Heart Fail 2014 Mar 1;7(2):279-287. [doi: 10.1161/CIRCHEARTFAILURE.113.000883] [Medline: 24519908]
21. Berglund A, Jaarsma T, Orädd H, et al. The Application of a Serious Game Framework to Design and Develop an Exergame for Patients With Heart Failure. JMIR Form Res 2024;8:e50063. [doi: 10.2196/50063]
22. Birn T, Holzmann C, Stech W. MobileQuiz: a serious game for enhancing the physical and cognitive abilities of older adults. In: Stephanidis C, Antona M, editors. Universal Access in Human-Computer Interaction Aging and Assistive Environments UAHCI 2014. Lecture Notes in Computer Science, Vol 8515: Springer; 2014:3-14. [doi: 10.1007/978-3-319-07446-7_1]
23. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954. [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
24. Müller AM, Khoo S. Non-face-to-face physical activity interventions in older adults: a systematic review. Int J Behav Nutr Phys Act 2014 Mar 10;11(1):35. [doi: 10.1186/1479-5868-11-35] [Medline: 24612748]
25. Ryder J, Longstaff B, Reddy S, Estrin D. Ambulation: a tool for monitoring mobility patterns over time using mobile phones. Presented at: 2009 International Conference on Computational Science and Engineering; Aug 29-31, 2009; Vancouver, BC p. 927-931. [doi: 10.1109/CSE.2009.312]
26. Paul L, Brewster S, Wyke S, et al. Increasing physical activity in older adults using STARFISH, an interactive smartphone application (app); a pilot study. J Rehabil Assist Technol Eng 2017 Apr 5;4:2055668317696236. [doi: 10.1177/2055668317696236] [Medline: 31186925]
27. Consolvo S, Klasnja PV, McDonald DW, et al. Flowers or a robot army? encouraging awareness & activity with personal, mobile displays. In: UbiComp '08: Proceedings of the 10th International Conference on Ubiquitous Computing: Association for Computing Machinery; 2008:54-63. [doi: 10.1145/1409635.1409644]
28. Rehman U, Cao S. Augmented-reality-based indoor navigation: a comparative analysis of handheld devices versus Google Glass. IEEE Trans Human-Mach Syst 2017 Feb;47(1):140-151. [doi: 10.1109/THMS.2016.2620106]
29. Consolvo S, Everitt K, Smith I, Landay JA. Design requirements for technologies that encourage physical activity. In: Grinter R, Rodden T, Aoki P, editors. CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2006:457-466. [doi: 10.1145/1124772.1124840]
30. Greaves CJ, Sheppard KE, Abraham C, et al. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11:119. [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
31. Albaina IM, Visser T, van der Mast C, Vastenburg MH. Flowie: a persuasive virtual coach to motivate elderly individuals to walk. Presented at: 3d International ICST Conference on Pervasive Computing Technologies for Healthcare; Apr 1 to 3, 2009; London, United Kingdom. [doi: 10.4108/ICST.PERVASIVEHEALTH2009.5949]
32. Garcia JA, Sundara N, Tabor G, Gay VC, Leong TW. Solitaire Fitness: design of an asynchronous exergame for the elderly to enhance cognitive and physical ability. Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); Aug 5-7, 2019; Kyoto, Japan p. 1-6. [doi: 10.1109/SeGAH.2019.8882471]
33. Ijsselsteijn W, Nap HH, de Kort Y, Poels K. Digital game design for elderly users. In: Future Play '07: Proceedings of the 2007 Conference on Future Play: Association for Computing Machinery; 2007:17-22. [doi: 10.1145/1328202.1328206]
34. Sinclair J, Hingston P, Masek M. Considerations for the design of exergames. In: GRAPHITE '07: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia: Association for Computing Machinery; 2007:289-295. [doi: 10.1145/1321261.1321313]
35. Marston HR. Design recommendations for digital game design within an ageing society. Educ Gerontol 2013 Feb;39(2):103-118. [doi: 10.1080/03601277.2012.689936]
36. Planinc R, Nake I, Kampel M. Exergame design guidelines for enhancing elderly's physical and social activities. In: Proccedings of The Third International Conference on AMBIENT Computing, Applications, Services and Technologies (AMBIENT 2013): IARIA; 2013:58-63 URL: http://hdl.handle.net/20.500.12708/54918 [accessed 2024-07-17]

37. Campbell T, Ngo B, Fogarty J. Game design principles in everyday fitness applications. In: CSCW '08: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work: Association for Computing Machinery; 2008:249-252. [doi: 10.1145/1460563.1460603]
38. de Carvalho RNS, Ishitani L. Motivational factors for mobile serious games for elderly users. Presented at: SBGames 2012: XI Simpósio Brasileiro de Games e Entretenimento Digital; Nov 2-4, 2012; Brasília, Brazil p. 19-28 URL: https://www.sbgames.org/sbgames2012/proceedings/papers/cultura/C_F3.pdf [accessed 2024-07-17]
39. Li J, Xu X, Pham TP, Theng YL, Katajapuu N, Luimula M. Exergames designed for older adults: a pilot evaluation on psychosocial well-being. Games Health J 2017 Dec;6(6):371-378. [doi: 10.1089/g4h.2017.0072] [Medline: 29131678]
40. Brox E, Fernandez-Luque L, Evertsen G, González-Hernández J. Exergames for elderly: social exergames to persuade seniors to increase physical activity. Presented at: 5th International ICST Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshop; May 23 to 26, 2011; Dublin, Republic of Ireland. [doi: 10.4108/icst.pervasivehealth.2011.246049]
41. Hamari J, Koivisto J, Sarsa H. Does gamification work?--a literature review of empirical studies on gamification. Presented at: 2014 47th Hawaii International Conference on System Sciences (HICSS); Jan 6-9, 2014; Waikoloa, HI p. 3025-3034. [doi: 10.1109/HICSS.2014.377]
42. Huotari K, Hamari J. Defining gamification: a service marketing perspective. In: MindTrek '12: Proceeding of the 16th International Academic MindTrek Conference: Association for Computing Machinery; 2012:17-22. [doi: 10.1145/2393132.2393137]
43. Matallaoui A, Koivisto J, Hamari J, Zarnekow R. How effective is “exergamification”? a systematic review on the effectiveness of gamification features in exergames. In: Proceedings of the 50th Hawaii International Conference on System Sciences: AIS Electronic Library (AISeL); 2017:3316-3325. [doi: 10.24251/HICSS.2017.402]
44. Berglund A, Jaarsma T, Berglund E, Strömberg A, Klompstra L. Understanding and assessing gamification in digital healthcare interventions for patients with cardiovascular disease. Eur J Cardiovasc Nurs 2022 Aug 29;21(6):630-638. [doi: 10.1093/eurjcn/zvac048] [Medline: 35709297]
45. Creative Commons License - Attribution 4.0 International - CC BY 4.0 Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2024-08-02]
46. Robson K, Plangger K, Kietzmann JH, McCarthy I, Pitt L. Is it all a game? understanding the principles of gamification. Bus Horiz 2015;58(4):411-420. [doi: 10.1016/j.bushor.2015.03.006]
47. Floryan M, Chow PI, Schueller SM, Ritterband LM. The model of gamification principles for digital health interventions: evaluation of validity and potential utility. J Med Internet Res 2020 Jun 10;22(6):e16506. [doi: 10.2196/16506] [Medline: 32519965]
48. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Dec;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
49. Wang H, Sun CT. Game reward systems: gaming experiences and social meanings. In: Proceedings of DiGRA 2011 Conference: Think Design Play: DiGRA; 2011. URL: https://dl.digra.org/index.php/dl/article/view/594 [accessed 2024-07-17]
50. Weiser P, Bucher D, Cellina F, De Luca V. A taxonomy of motivational affordances for meaningful gamified and persuasive technologies. In: Proceedings of EnviroInfo and ICT for Sustainability 2015: Atlantis Press; 2015:271-280. [doi: 10.2991/ICT4S-ENV-15.2015.31]
51. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? gamification in health and fitness apps. JMIR Serious Games 2014 Aug 4;2(2):e9. [doi: 10.2196/games.3413] [Medline: 25654660]
52. Costa JP, Wehbe RR, Robb J, Nacke LE. Time's up: studying leaderboards for engaging punctual behaviour. In: Gamification '13: Proceedings of the First International Conference on Gameful Design, Research, and Applications: Association for Computing Machinery; 2013:26-33. [doi: 10.1145/2583008.2583012]
53. Phillips C, Johnson D, Klarkowski M, White MJ, Hides L. The impact of rewards and trait reward responsiveness on player motivation. In: CHI PLAY '18: Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play: Association for Computing Machinery; 2018:393-404. [doi: 10.1145/3242671.3242713]
54. Mekler ED, Brühlmann F, Tuch AN, Opwis K. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Comput Human Behav 2017 Jun;71:525-534. [doi: 10.1016/j.chb.2015.08.048]
55. Chen Y, Wang Q, Chen H, Song X, Tang H, Tian M. An overview of augmented reality technology. J Phys: Conf Ser 2019 Jun;1237(2):022082. [doi: 10.1088/1742-6596/1237/2/022082]
56. Garcia JA, Navarro KF. The mobile RehAppTM: an AR-based mobile game for ankle sprain rehabilitation. Presented at: 2014 IEEE 3rd International Conference on Serious Games and Applications for Health (SeGAH); May 14-16, 2014; Rio de Janeiro, Brazil p. 1-6. [doi: 10.1109/SeGAH.2014.7067087]
57. Fu Y, Hu Y, Sundstedt V. A systematic literature review of virtual, augmented, and mixed reality game applications in healthcare. ACM Trans Comput Healthcare 2022 Mar 3;3(2):1-27. [doi: 10.1145/3472303]
58. Rajeev S, Wan Q, Yau K, Panetta K, Agaian SS. Augmented reality-based vision-aid indoor navigation system in GPS denied environment. In: Agaian SS, Asari VK, DelMarco SP, editors. Proceedings Volume 10993, Mobile Multimedia/Image Processing, Security, and Applications 2019: SPIE; 2019. [doi: 10.1117/12.2519224]

59. Möller A, Kranz M, Diewald S, et al. Experimental evaluation of user interfaces for visual indoor navigation. In: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2014:3607-3616. [doi: 0.1145/2556288.255700]
60. Stamm O, Vorwerg S, Müller-Werdan U. Exergames in augmented reality for older adults with hypertension: a qualitative study exploring user requirements. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population. Social Media, Games and Assistive Environments HCII 2019. Lecture Notes in Computer Science (LNISA, Vol 11593): Springer; 2019:232-244. [doi: 10.1007/978-3-030-22015-0_18]
61. Chirico A, Avellone M, Palombi T, et al. Exploring the psychological nexus of virtual and augmented reality on physical activity in older adults: a rapid review. Behav Sci (Basel) 2023 Dec 31;14(1):31. [doi: 10.3390/bs14010031] [Medline: 38247683]
62. Ferguson C, Davidson PM, Scott PJ, Jackson D, Hickman LD. Augmented reality, virtual reality and gaming: an integral part of nursing. Contemp Nurse 2015;51(1):1-4. [doi: 10.1080/10376178.2015.1130360] [Medline: 26678947]
63. Ng YL, Ma F, Ho FK, Ip P, Fu KW. Effectiveness of virtual and augmented reality-enhanced exercise on physical activity, psychological outcomes, and physical performance: a systematic review and meta-analysis of randomized controlled trials. Comput Human Behav 2019 Oct;99:278-291. [doi: 10.1016/j.chb.2019.05.026]
64. Schlomann A, von Storch K, Rasche P, Rietz C. Means of motivation or of stress? the use of fitness trackers for self-monitoring by older adults. HBScience 2016 Aug;7(3):111-116. [doi: 10.1007/s16024-016-0275-6]
65. Knöll M, Dutz T, Hardy S, Göbel S. Urban exergames: how architects and serious gaming researchers collaborate on the design of digital games that make you move. In: Ma M, Jain LC, Anderson P, editors. Virtual, Augmented Reality and Serious Games for Healthcare: Springer; 2014:191-207. [doi: 10.1007/978-3-642-54816-1_11]
66. Lindeman RW, Lee G, Beattie L, Gamper H, Pathinarupothi R, Akhilesh A. Geoboids: a mobile AR application for exergaming. Presented at: 2012 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH); Nov 5-8, 2012; Altanta, GA p. 93-94. [doi: 10.1109/ISMAR-AMH.2012.6483997]
67. Yang H, Li J, Liu J, Bian Y, Liu J. Development of an indoor exergame based on moving-target hitting task for COVID-19 epidemic: a comparison between AR and VR mode. IEEE Trans Games 2021 Sep;14(3):511-521. [doi: 10.1109/TG.2021.3118035]
68. LeBlanc AG, Chaput JP. Pokémon GO: a game changer for the physical inactivity crisis? Prev Med 2017 Aug;101:235-237. [doi: 10.1016/j.ypmed.2016.11.012] [Medline: 27856340]
69. Boulos MNK, Yang SP. Exergames for health and fitness: the roles of GPS and geosocial apps. Int J Health Geogr 2013 Apr 5;12:18. [doi: 10.1186/1476-072X-12-18] [Medline: 23561306]
70. Bui V, Le NT, Vu TL, Nguyen VH, Jang YM. GPS-based indoor/outdoor detection scheme using machine learning techniques. Appl Sci 2020;10(2):500. [doi: 10.3390/app10020500]
71. Kunhoth J, Karkar A, Al-Maadeed S, Al-Ali A. Indoor positioning and wayfinding systems: a survey. Hum Cent Comput Inf Sci 2020 May 2;10:18. [doi: 10.1186/s13673-020-00222-0]
72. Gerstweiler G, Vonach E, Kaufmann H. HyMoTrack: a mobile AR navigation system for complex indoor environments. Sensors (Basel) 2015 Dec 24;16(1):17. [doi: 10.3390/s16010017] [Medline: 26712755]
73. Ng XH, Lim WN. Design of a mobile augmented reality-based indoor navigation system. Presented at: 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT); Oct 22-24, 2020; Istanbul, Turkey p. 1-6. [doi: 10.1109/ISMSIT50672.2020.9255121]
74. Karan E, Morkoc ME, Şireci AE, Fidanboylu K. Indoor and outdoor navigation application development with augmented reality technology. Presented at: 5th International Conference on Computational Mathematics and Engineering Sciences; Jun 8-10, 2021; Van, Turkey p. 211-220 URL: https://www.researchgate.net/publication/359282706 [accessed 2024-07-17]
75. Wagner-Greene VR, Wotring AJ, Castor T, Kruger J, Mortemore S, Dake JA. Pokémon GO: healthy or harmful? Am J Public Health 2017 Jan;107(1):35-36. [doi: 10.2105/AJPH.2016.303548] [Medline: 27925813]
76. Ayers JW, Leas EC, Dredze M, Allem JP, Grabowski JG, Hill L. Pokémon GO-a new distraction for drivers and pedestrians. JAMA Intern Med 2016 Dec 1;176(12):1865-1866. [doi: 10.1001/jamainternmed.2016.6274] [Medline: 27635638]
77. Klompstra L, Jaarsma T, Strömberg A, Evangelista LS, van der Wal MHL, HF-Wii Study Team. Exercise motivation and self-efficacy vary among patients with heart failure - an explorative analysis using data from the HF-Wii study. Patient Prefer Adherence 2021 Oct 19;15:2353-2362. [doi: 10.2147/PPA.S314414] [Medline: 34703217]
78. Guyatt GH, Sullivan MJ, Thompson PJ, et al. The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. Can Med Assoc J 1985 Apr 15;132(8):919-923. [Medline: 3978515]
79. Preist C, Massung E, Coyle D. Competing or aiming to be average?: normification as a means of engaging digital volunteers. In: CSCW '14: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing: Association for Computing Machinery; 2014:1222-1233. [doi: 10.1145/2531602.2531615]
80. Durrant-Whyte H, Bailey T. Simultaneous localization and mapping: part I. IEEE Robot Automat Mag 2006 Jun;13(2):99-110. [doi: 10.1109/MRA.2006.1638022]
81. Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): part II. IEEE Robot Automat Mag 2006 Sep;13(3):108-117. [doi: 10.1109/MRA.2006.1678144]

82. Brox E, Fernandez-Luque L, Tøllefsen T. Healthy gaming - video game design to promote health. Appl Clin Inform 2011 Apr 27;2(2):128-142. [doi: 10.4338/ACI-2010-10-R-0060] [Medline: 23616865]
83. Baumeister RF. Toward a general theory of motivation: problems, challenges, opportunities, and the big picture. Motiv Emot 2016 Feb;40(1):1-10. [doi: 10.1007/s11031-015-9521-y]
84. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology: Association for Computing Machinery; 2009:1-7. [doi: 10.1145/1541948.1541999]
85. Stawarz K, Cox AL, Blandford A. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. In: CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2015:2653-2662. [doi: 10.1145/2702123.2702230]
86. Höchsmann C, Walz SP, Schäfer J, Holopainen J, Hanssen H, Schmidt-Trucksäss A. Mobile exergaming for health-effects of a serious game application for smartphones on physical activity and exercise adherence in type 2 diabetes mellitus-study protocol for a randomized controlled trial. Trials 2017 Mar 6;18(1):103. [doi: 10.1186/s13063-017-1853-3] [Medline: 28264717]

Abbreviations

AR: augmented reality

Edited by G Eysenbach, N Cahill, T Leung; submitted 18.06.23; peer-reviewed by J Garcia, K Blondon, P Bamidis; revised version received 29.04.24; accepted 11.06.24; published 21.08.24.

Please cite as:
Berglund A, Klompstra L, Orädd H, Fallström J, Strömberg A, Jaarsma T, Berglund E
The Rationale Behind the Design Decisions in an Augmented Reality Mobile eHealth Exergame to Increase Physical Activity for Inactive Older People With Heart Failure
JMIR Serious Games 2024;12:e50066
URL: https://games.jmir.org/2024/1/e50066
doi:10.2196/50066

Viewpoint

Is the Apple Vision Pro the Ultimate Display? A First Perspective and Survey on Entering the Wonderland of Precision Medicine

Jan Egger[1,2,3], PhD; Christina Gsaxner[1,4,5,6], PhD; Gijs Luijten[1,6], MSc; Jianxu Chen[7], PhD; Xiaojun Chen[8,9], PhD; Jiang Bian[10], PhD; Jens Kleesiek[1,3,11,12], PhD, MD; Behrus Puladi[4,5], MD, DMD

[1]Institute for Artificial Intelligence in Medicine, Essen University Hospital (AöR), Essen, Germany
[2]Center for Virtual and Extended Reality in Medicine (ZvRM), Essen University Hospital (AöR), Essen, Germany
[3]Cancer Research Center Cologne Essen (CCCE), University Medicine Essen (AöR), Essen, Germany
[4]Department of Oral and Maxillofacial Surgery & Institute of Medical Informatics, University Hospital RWTH Aachen, Aachen, Germany
[5]Institute of Medical Informatics, University Hospital RWTH Aachen, Aachen, Germany
[6]Institute of Computer Graphics and Vision, Graz University of Technology, Graz, Austria
[7]Leibniz-Institut für Analytische Wissenschaften (ISAS), Dortmund, Germany
[8]Institute of Biomedical Manufacturing and Life Quality Engineering, State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China
[9]Institute of Medical Robotic, Shanghai Jiao Tong University, Shanghai, China
[10]Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, United States
[11]German Cancer Consortium (DKTK), Partner Site Essen, Essen, Germany
[12]Department of Physics, TU Dortmund University, Dortmund, Germany

Corresponding Author:
Behrus Puladi, MD, DMD
Department of Oral and Maxillofacial Surgery & Institute of Medical Informatics
University Hospital RWTH Aachen
Pauwelsstraße 30
Aachen, 52074
Germany
Phone: 49 241 80 88231
Fax: 49 241 80 82430
Email: bpuladi@ukaachen.de

Abstract

At the Worldwide Developers Conference in June 2023, Apple introduced the Vision Pro. The Apple Vision Pro (AVP) is a mixed reality headset; more specifically, it is a virtual reality device with an additional video see-through capability. The video see-through capability turns the AVP into an augmented reality (AR) device. The AR feature is enabled by streaming the real world via cameras on the (virtual reality) screens in front of the user's eyes. This is, of course, not unique and is similar to other devices, such as the Varjo XR-3 (Varjo Technologies Oy). Nevertheless, the AVP has some interesting features, such as an inside-out screen that can show the headset wearer's eyes to "outsiders," and a button on the top, called the "digital crown," that allows a seamless blend of digital content with the user's physical space by turning it. In addition, it is untethered, except for the cable to the battery, which makes the headset more agile, compared to the Varjo XR-3. This could actually come closer to "The Ultimate Display," which Ivan Sutherland had already sketched in 1965. After a great response from the media and social networks to the release, we were able to test and review the new AVP ourselves in March 2024. Including an expert survey with 13 of our colleagues after testing the AVP in our institute, this Viewpoint explores whether the AVP can overcome clinical challenges that AR especially still faces in the medical domain; we also go beyond this and discuss whether the AVP could support clinicians in essential tasks to allow them to spend more time with their patients.

(JMIR Serious Games 2024;12:e52785) doi:10.2196/52785

KEYWORDS

Apple Vision Pro; mixed reality; augmented reality; virtual reality; health care; mobile phone

Introduction

In his highly influential paper "The Ultimate Display" [1], Ivan Sutherland, the pioneer of computer graphics and creator of the first augmented reality (AR) head-mounted display (HMD) in the early 1960s, predating the advent of PCs, described aspects that Apple's Vision Pro aims to implement. The Apple Vision Pro (AVP) can span the wide range of the reality-virtuality continuum, as explained by Milgram et al [2], starting from unaltered *reality*, achieved by streaming the real world without additional information to the user; progressing through *mixed reality* (MR) [3], involving the integration of digital information into the real world (referred to as AR) and vice versa (augmented virtuality); and extending to complete immersion, also known as virtual reality (VR), where the user experiences only computer-generated virtual content.

In the realm of VR, there is currently a broader selection of headsets at various price points, including offerings from Meta, Vive, Varjo, PlayStation, and Google (with the most economical probably being the now-discontinued Google Cardboard or a similar variant from another vendor), compared to AR. AR faces a distinct challenge because of the need to analyze *reality* to augment it at the *right* position with digital content for the user. An exception is AR used for a *pure* simulation, where the AR hologram is indeed shown in the real world but has no meaningful relation to or interactions with real-world objects. An example of such a scenario is a surgical simulation [4]. The AR hologram is shown in front of the user to allow them to inspect and interact with it. Still, the underlying hardware and software need to capture and model the real world to *anchor* the hologram within it so that the hologram does not *drift* away from the user's view.

In a medical scenario, AR can enable equips a physician with x-ray vision into a patient [5]. This capability is a valuable advancement from the current practice of viewing patient information, such as preoperative imaging, on PC monitors, contributing to the "switching focus problem," where a physician must divide their attention between the patient and the digital information. Hansen et al [6] outlined that such a division leads to an increased mental workload, disorientation, and deteriorated hand-eye coordination. Instead, AR can display patient information directly in the physician's view of the patient, additionally offering spatial perception even before any incisions are made. A recent systematic review of the AR headset HoloLens (Microsoft Corp) in medicine [7] showed its broad applications and the massive research efforts during the recent years in all these areas (medicine and health care with applications for patients, physicians or surgeons and students, which should ultimately lead to the development of an intelligent health care metaverse [8,9]). Hence, we highlight the user groups and applications targeted by AR in the next sections and discuss how common challenges in AR can be overcome with the AVP's capabilities. Furthermore, we share our findings about the new AVP gathered through a structured survey among ourselves and colleagues who had the opportunity to try out the AVP.

Target Groups

Physicians and Health Care Professionals

Physicians and health care professionals comprise, by far, the largest target group for AR applications [7]. The adoption and use cases of AR within nursing departments are growing, focusing on wound care, information retrieval, injections, and education [10]. In this context, the accuracy of the visualization is not of utmost importance, but having a wearable "sensing" computer can increase time at the bedside through automatic information retrieval and documentation. However, especially for AR-supported image guidance and navigation, such as in surgery, very high accuracy and reliability are needed [11]. Applications for which submillimeter precision is not necessary are, for example, ablations, ventriculostomies [12-16], and certain orthopedic interventions [17]. Here, the HoloLens is, with its inside-out tracking, already a promising tool, but for applications that need, for example, submillimeter precision, it cannot be used reliably yet. An example is the deep brain stimulation procedure used for treating essential tremors and Parkinson disease, where millimeter-to-submillimeter accuracy in deep brain stimulation targeting (an electrode placement inside the brain) can be important [18]. Another study exploring the clinical accuracy of the HoloLens for neuronavigation concluded that it is currently not within clinically acceptable levels [19]. The same holds true for some application scenarios in orthopedic surgery [20], such as screw placement, where there is still room for improvement [21]. This improvement can be achieved through external tracking hardware [22,23], but it may come at the cost of mobility and increased expenses. The AVP is promising due to its superior sensing technology, as discussed in the Accuracy and Reliability subsection. More importantly, the powerful Apple M1 and R1 chipsets can handle the sensor data appropriately to exploit them completely. This technology comes at a high price, but the often-criticized end-user price of US $3499 (without tax) for the AVP does not seem to be an issue for health professionals who want to use it in health care. The price is similar to that of the HoloLens and much cheaper than (a fraction of) those of existing and clinically used conventional medical navigation systems, for example, those from Brainlab (Stefan Vilsmeier) or Medtronic (Earl Bakken and Palmer Hermundslie). In this context, it is important to mention that the use of AR HMDs with direct patient reference, outside of research in the context of feasibility studies, must obtain appropriate regulatory approval, and the AT HMDs must be certified and classified as medical devices. The Magic Leap 2 (Rony Abovitz), for example, received certification for use in an operating room last year (2023). The requirements to be met are still up to debate, for example, when it comes to active or passive cooling. In contrast to the HoloLens 2, both AVP and Magic Leap have active cooling, potentially compromising the sterile airflow. Furthermore, sterilization is a problem because Apple recommends not using isopropyl alcohol or other solvents for cleaning. This should not be an absolute impediment, given the approval of the Magic Leap 2 for use in the operating theater and the use of regular glasses in the operating theater. We do, however, think that the time is coming when clinicians, technicians, and manufacturers should come together to meet binding guidelines.

Another challenge for clinical adoption is the integration of the device within the digital infrastructure of the hospital, for example, to retrieve data from and transfer data to the electronic health records. Moreover, and especially in a medical context, the display may also require an image focus at surgical table distance [24]. In combination with the outward sensors, this would be ideal for medical scenarios.

Students

Students are the second most common intended target group [7] for AR applications, for example, in educational training, such as the HoloPointer, a virtual AR pointer for laparoscopic surgery training [25]. Another example is the use of AR to teach medical students catheter placement [26] and the use of AR for training on central venous access [27]. Notably, the aforementioned partially outward-oriented sensors, unlike those in other headsets, allow for the accurate tracking of all "normal" (hand) movements for interaction, regardless of (head) position. This is also valuable for health care professionals when practicing or working at the patient's bedside. When combined with speech commands and eye-controlled interactions, physicians can expectantly implement the AVP easily and in a sterile manner.

However, we found that the effects of HoloLens-based learning compared to those of conventional learning, for example, using cadavers or other computerized methods, seem to be rather small [7]. A reason for this is that the learning tools usually used are relatively simple, conventional 3D models, and more innovative visualizations, including interactive, dynamic content, which cannot be easily delivered through regular computerized methods, have not been explored in depth yet [7]. Here, we believe that the AVP can raise the bar because of its high-resolution visualization (consisting of 2 micro-OLED displays with 23 megapixels) that uses eye-tracking to make the headset also usable as a desktop screen with its fine textual details. In addition, the 3D user interface, using precise and intuitive finger tracking in combination with eye interactions, can be a game changer. Despite these advances, it remains to be seen whether an AR headset has a major advantage over a conventional screen. The cost of the AVP is also a concern, especially for students with a limited financial budget, but the "pro" might indicate that a "regular" version will also be released. For medical simulation training, however, an examination of costs [28] suggests that using the AVP is a financially viable option. The AVP is, however, highly personalized to the individual user to optimize the experience and accuracy. Luckily for the student group, the AVP can be shared, and the interpupillary distance setting (IDP) is adaptable, compromising the experience minimally, but sharing greatly reduces cost for simulation purposes.

Patients

Patients are currently the least frequented target group for extended reality (XR) devices [7]. An example is the use of MemHolo, which provides MR experiences for patients with Alzheimer disease [29]. Other examples in this area include the use of the HoloLens as an assistance and monitoring tool for medication adherence [30] and the use of a HoloLens-based system for functional mobility assessment [31]. We expect the advanced eye-tracking to enhance the user experience because visual cues will not merely exist within the user's field of view but can also be precisely situated in the user's focal point. However, many interesting assistance and monitoring applications are limited by the restricted possible use time of untethered AR headsets, primarily due to battery life, and this problem seems to exist in the AVP, which has a similar battery life to that of the HoloLens, approximately 2 hours. The only consolation is the "external" battery pack of the AVP, which can be swapped. Finally, the Apple M2 chip and the new R1 chip of the AVP could potentially boost the relatively small number of applications that have been explored so far, for example, by increasing the frames per second, minimizing display latency, and thus reducing motion sickness, especially useful for older patients, who are more susceptible to MR sickness than younger users [32,33].

Differences Between Target Groups

An example of different use cases per user group is shown in Textbox 1. The main target group is health care professionals, which can be divided into several subgroups, such as surgeons, physicians, and nurses. Unlike the other groups, price is not a major issue for this group, but the use cases are different from those of the other groups. For nurses, data retrieval, documentation, and guidance without the need for high accuracy are the estimated use cases, freeing up time at the bedside. The AVP's high data transfer rate and processing power are more important. For a surgeon, however, accuracy is critical for surgical navigation [34]. For the general practitioner, seamless integration and unparalleled view during transport may be the greatest values of the AVP. For other groups, price is an important consideration, with the caveat that it could reduce overall costs for medical simulation training. The average patient is not expected to be able to afford this device on their own. For now, it could be used for therapies in the clinic, where it can be shared among several patients. However, this makes cleaning the device more important than if there is a single user. We expect to see silicon-replaceable parts for the device, which would ultimately benefit all groups. The AVP has similar challenges to other headsets in terms of battery life and motion sickness, which affects all groups, especially nurses during long, hectic shifts with movement and sometimes no breaks. The integration of an external battery makes battery life easy to address in future hardware updates. Regarding motion sickness, we expect Apple to offer a lower resolution for a higher refresh rate in combination with its foveated rendering to combat this. In summary, the biggest differences between target groups are in the need for high accuracy and the willingness to accept the high price.

Textbox 1. Possible health care scenarios with the Apple Vision Pro.

Health care professionals

- Patient record retrieval
- Automated documentation
- Guidance in medical procedures and equipment
- Surgical preoperative planning
- Surgical navigation
- Peerless public medical record viewer

Students

- Complex anatomy education
- Complex pathophysiology education
- Training simulators
- Automated skill evaluation

Patients

- Patient treatment preparation
- Patient education
- (Automated) diagnostics
- Rehabilitation
- Improved patient well-being

Challenges

Despite significant research efforts in medical AR, clinical translation remains limited so far [35]. Several technical challenges require ongoing multidisciplinary efforts among software developers, hardware manufacturers, regulatory entities, and clinical personnel to create safe, effective, and seamlessly integrated AR solutions in health care. Some of these challenges could be addressed by the AVP's hardware and user interface.

Accuracy and Reliability

The most significant challenge in medical AR systems is their accuracy and reliability in registration and calibration, which are critical for aligning virtual objects with real-world elements. Failures in tracking or registration components can misguide physicians, potentially leading to serious medical accidents. Although recent developments in tracking technology are encouraging [36,37], clinical settings pose additional challenges with limited space, unconventional lighting, and dynamic environments. A solution is to combine an AR device with external tracking systems to increase precision [22]. On the downside, such combinations require complicated setups and the calibration of bulky (external) devices, diminishing the simplicity and lightweight nature of purely head-worn systems. The latest sensor technology, combined with state-of-the-art tracking and registration algorithms, could eliminate the need for external systems to meet clinical requirements. We expect the AVP to move the needle in terms of accuracy because of its strong inside-out tracking through 12 built-in cameras and light detection and ranging sensing method, which is the key aspect for increasing the accuracy of AR. We expect that enhanced eye-tracking capabilities and the incorporation of multiple outward-facing sensors will substantially enhance the reliability and stability of user interactions. The increase in both the quantity and range of sensors, coupled with a more potent chipset, is bound to increase spatial resolution and visualization accuracy significantly in comparison to untethered alternatives. Future research should focus on exploiting this improved hardware setup.

Usability and Human Factors

In complex clinical environments, a seamless integration of AR systems into existing clinical processes is crucial. Existing AR solutions often add undue complexity and disrupt established workflows, requiring intricate hardware, many manual steps, and elevated technical knowledge and understanding. Health care professionals, who possess varying degrees of technical background, or patients, who represent diverse demographics, cannot be reasonably expected to meet these demands. Hence, AR systems need to be intuitive and user-friendly. Apple products are known for their user-centric design and seamless integration of hardware and software. We expect that the AVP's user experience, in terms of interface, interaction, and feedback, will be more intuitive for a diverse audience than existing HMDs and VR glasses.

As a demonstration of this ease of use, the AVP connects automatically to nearby Apple MacBooks, commonly used by physicians, to display the screen and access files in real time without any effort. An obvious use case is secure, peerless access to patient files while using public transportation. This is valuable for working with all medical records. Apple products work seamlessly with each other, but in a clinical setting, not all products may come from Apple or its partners. A combined challenge and focus should be to enable cross-manufacturer integration.

In addition, as mentioned earlier, more natural and user-guided interactions are possible with the outward and eye-tracking sensors. This will allow developers to create more intuitive applications. The improved hardware should not be underestimated. For example, Gsaxner et al [38] developed stand-alone marker-less image registration, which has potential in many medical applications, but the setup of an additional laptop is cumbersome and another potential risk for data breach. With the AVP, this type of setup is most likely not needed.

Visualization and Perception

Current AR devices face limitations in terms of view, resolution, and brightness, impacting visual fidelity [39]. In addition to these hardware constraints, challenges include depth perception and occlusion between virtual and real content [40]. Current optical see-through (OST) displays, such as that of the popular HoloLens, struggle to address these issues adequately, as they cannot modify the reality presented. By contrast, the AVP's video see-through (VST) system provides significantly more possibilities for seamlessly and adaptively integrating virtual elements with the real environment. Visual accuracy is difficult to achieve and user dependent, but eye-tracking and automatic eye calibration, combined with the high-resolution displays inside the headset, make us believe that the AVP could set new standards in terms of the quality and perception of visualization.

In summary, hardware limitations are the main bottleneck in developing AR applications suitable for the clinical translation of AR that cannot be addressed by software developers. Another bottleneck is the regulation and list of requirements for 3D smart glasses in surgery. Combining hardware technology and software from different companies, especially for data transfer, is the final major issue, however, but it is not limited to 3D smart glasses and has been a persistent obstacle in health care [41].

We believe that the AVP can provide significant advancements as the new state-of-the-art device, with much-needed improvements in precision, reliability, usability, workflow, and perception. The combination of AR and VR in an XR device has the potential to transform health care by aiding in diagnostics, improving surgical procedures, facilitating remote patient care, and enhancing medical training. Many other novel applications may emerge because of the advancements in AVP hardware and software platform. Returning to our main question, the AVP has the potential to enhance medical education and therapy, medical record evaluation, data retrieval, and documentation automation while improving patient outcomes and freeing up clinicians to spend more time at the patient's bedside. Of course, this will require a collaborative effort among engineers, researchers, and clinicians to develop or enhance these applications and to prove cost-effectiveness to ensure clinical adoption.

The New AVP

Highlighted Features

Apple announced some interesting (and still unique) features of the AVP. In the upcoming sections, we introduce and discuss these features in the context of medical MR with a focus on AR.

Digital Crown

The digital crown is a turning button that allows the user to blend digital content easily and seamlessly with the physical environment surrounding them. The button is named the same as that in the Apple Watch, but it serves a different function in the AVP by providing an easy option to quickly *escape* VR without the need to remove the headset. Depending on the capabilities and accessibility of developers to exploit the digital crown, it could also enable diminished reality (DR) [3,42]. This feature would allow users, such as physicians, students, and patients, to remove distracting real-world objects or environments, thereby reducing visual overstimulation. Notably, DR may even extend to diminishing or removing the headsets of other users. To elaborate, the AVP can create realistic avatars of its users. When using the FaceTime application, the movement and direction of the eyes, tongue, and facial muscles are correctly tracked and visualized for other users of the AVP inside the same FaceTime application. This could be used to give the wearer of the headset the ability to perform DR, that is, to diminish the headset and in paint the face when communicating with other AVP users in the real world but via the pass-through mode. The idea of DR is not new, but its implementation, especially in medicine, is a relatively novel [43]. The AVP may be the first untethered device capable of providing DR in a medically relevant and accurate context without the need for external hardware.

Notably, DR needs, in general, a VST capability and cannot be realized by OST displays, such as that in the HoloLens. The medical HMD from VOSTARS has addressed this by combining an OST display, that is, the view through a semitransparent glass, with a VST function [44]. VOSTARS has realized this with a semitransparent glass that can transition into a nontransparent mode. The AVP, by contrast, does not conform to the OST versus VST dichotomy but rather operates on the VST versus VR HMD spectrum, aligning with the reality-virtuality continuum, postulated by Milgram and Kishino in 1994 [45].

Virtual Eyes

Spooky at first sight, the outward display reveals a user's eyes while wearing the AVP. According to Apple, this feature is meant to let others know when the AVP wearer is using apps or is completely immersed in a virtual world. This feature is not needed in OST AR devices, such as the HoloLens, because the headset wearer's eyes can be seen through the transparent display. This feature is particularly crucial when considering patients who are dealing with serious illnesses. The transparency ensures that medical professionals, while using the device, do not create an additional layer of detachment by being obscured behind a screen, potentially alleviating concerns about the impact on patient-physician interactions.

3D Camera

The AVP features a 3D camera, which allows a user to capture spatial photos and spatial videos in 3D. This could hold significant potential for the professional health care sector. The ability to capture spatial photos and videos in 3D could prove immensely valuable for documenting medical interventions. This is certainly also of interest for medical training, allowing residents to review and reenact treatments in 3D from different viewpoints. It could also serve as a tool to uncover malpractice. The 3D capabilities can also provide an alternative or an addition to the 3D documentation of crime and crash scenes, which is common practice during forensic and medicolegal investigations [46].

Interaction

According to Apple, a user can interact with virtual content by simply looking at an element, tapping the fingers together to select it, and using the virtual keyboard or dictation to type. There also seems to be a “visual search” function, which may be similar to the “visual look up” feature found on iPhones and iPads. “Visual search” will allow users to interact with items and text for opening webpages and translating into other languages in real time. Interestingly, this was already described by Sutherland [1] as “the language of glances” to interact with a computer, with an example where looking at a corner of a triangle makes it become rounded. Looking something up is definitely of interest, for example, to surgeons who have their hands full with (surgical) instruments or nonverbal users.

Optic ID

The AVP is supposed to support the optical identification of a user for authentication, such the iPhone’s face identification. The biometric method is enabled by scanning the iris of the user wearing the headset. This is not unique to the AVP because the HoloLens 2 already supported an iris authentication. However, this is also an important feature for the AVP in health care, where you have sensitive patient data, and unlocking the device and accessing the data should be permitted only for authorized users. Furthermore, it enables hands-free log-in in a potentially sterile environment.

ZEISS Optical Inserts

Carl Zeiss AG is working together with Apple to provide precision optics for users who require vision correction. This will improve the visual experience and provide greater comfort, as users do not need to wear their prescription glasses and frames inside the limited space within the AVP. In addition, it may help avoid side effects such as “overheating,” headaches, and light bleeding. The additional lenses can be attached magnetically to the main lens, which makes insertion and replacement easy.

Head Strap

The head strap of the AVP seems to be easily exchangeable, and there are already several alternatives (in terms of materials, colors, etc) available on the web. This will ultimately enable more comfort in comparison to the HoloLens and the Varjo XR-3, which are adjustable but still consist of quite stiff (plastic) frames around the head. It will definitely be more comfortable than the Sword of Damocles, developed by Ivan Sutherland and his student Bob Sproull [47], where the user had to be strapped into the AR or VR HMD device, which was suspended from the ceiling because it was so heavy. The head strap also contains the AVP’s speaker, which is placed directly over the user’s ears and is supposed to virtualize surround sound. However, the strap means that all main components are integrated with the visor. The HoloLens 2, for example, distributes the weight more evenly by integrating components in a plastic case located at the rear, including the battery and system on a chip board. This was done by Microsoft based on its experiences with the HoloLens 1 to make the HoloLens 2 more comfortable. On the contrary, the AVP seems to be worn more like a VR device that fits tightly onto the face instead of partly hovering in front of the user’s eyes, such as other AR devices that are more similar to prescription glasses. The AVP has a similar weight range (600-650 g) to other popular AR and VR headsets, even though the battery is not integrated into the device and is separate. In part, this could be attributed to the material, for example, the aluminum alloy frame. Furthermore, in a medical context, it is usually necessary to be able to easily disinfect the equipment used. A special version with a disinfectable head strap could meet this requirement.

3D Persona Avatars

During the setup of the AVP, the user’s face is scanned by the headset to create a photorealistic avatar, which is then used by the device’s operating system, called VisionOS. For better fitting purposes, the user’s face can also be scanned with the TrueDepth camera of an iPhone, as seen at a public demo at the Worldwide Developers Conference in June 2023. Furthermore, the user’s ears can be scanned to optimize the speakers that are inside of the head strap because, aside from the shape and size of the head, the ear sizes, positions, and distances can vary a lot between users. All these user-specific options will make the headset more comfortable to wear, and this will finally increase the acceptability of the new device, not only in the health care sector but also in other sectors.

Unity

The Unity engine already supports pretty much all common headset devices, whether it be AR, VR, or MR, such as the Oculus (Palmer Luckey), HoloLens, or Google VR (Google LLC). Unity also supports the OpenVR application programming interface that has already been used in medicine [48]. Therefore, the support of the Unity engine of the AVP makes the development and porting of apps much easier. This also makes a comparison between and an evaluation of devices with the “same” app easier. Notably, Unity currently already allows the development of apps for iPhone, iPad, and Mac while deployment occurs via the Apple app store or Xcode, suggesting future development for the AVP. Overall, Unity support will enable faster development of apps, hopefully fast enough to reach a critical mass before end users may turn their backs, which is needed for the long-term success of the new Apple device.

Medical Device

Talking about the health care sector, we need to mention that the AVP is not certified as a medical device yet. We are not aware of whether Apple has plans in this direction, at least not publicly. Apple, however, promoted Complete HeartX as an education app for medical students by providing hyperrealistic 3D models and animations of the heart and medical issues. In comparison, Microsoft advertised the HoloLens 2 specifically for medical scenarios, for example, with a plenary presentation by Bernard Kress (who was at that time Principal Optical Architect, HoloLens team, Microsoft Corp) at the International Society for Optics and Photonics Medical Imaging in February 2020. Nevertheless, we are sure that the AVP will be used by researchers and companies for medical applications, similar to other devices, if Apple does not explicitly prohibit this somehow, for example, by monitoring their devices and locking them. The integration of the AVP with other software and hardware commonly used in clinical settings, for example, integration with electronic health records, will be interesting [49,50].

First Experience

In March 2024, we had the opportunity to try out the AVP ourselves for the first time (Figure 1). We, therefore, conducted a standardized expert survey among colleagues (and ourselves) at our institute (the Institute for Artificial Intelligence in Medicine). The inclusion criterion was an experience of at least 10 minutes with the AVP. After obtaining written consent, we conducted a structured survey on the AVP. We collected baseline data from the respondents including experience with HMDs (Table 1), as well as asked Likert questions on how far the AVP corresponds to an “Ultimate Display” (Table 2). We also administered the User Experience Questionnaire (UEQ; Figure 2) as well as asked open-ended questions about what is missing from an “Ultimate Display” (Textbox 2).

Figure 1. The illustration of a possible Apple Vision Pro scenario with one of the authors.

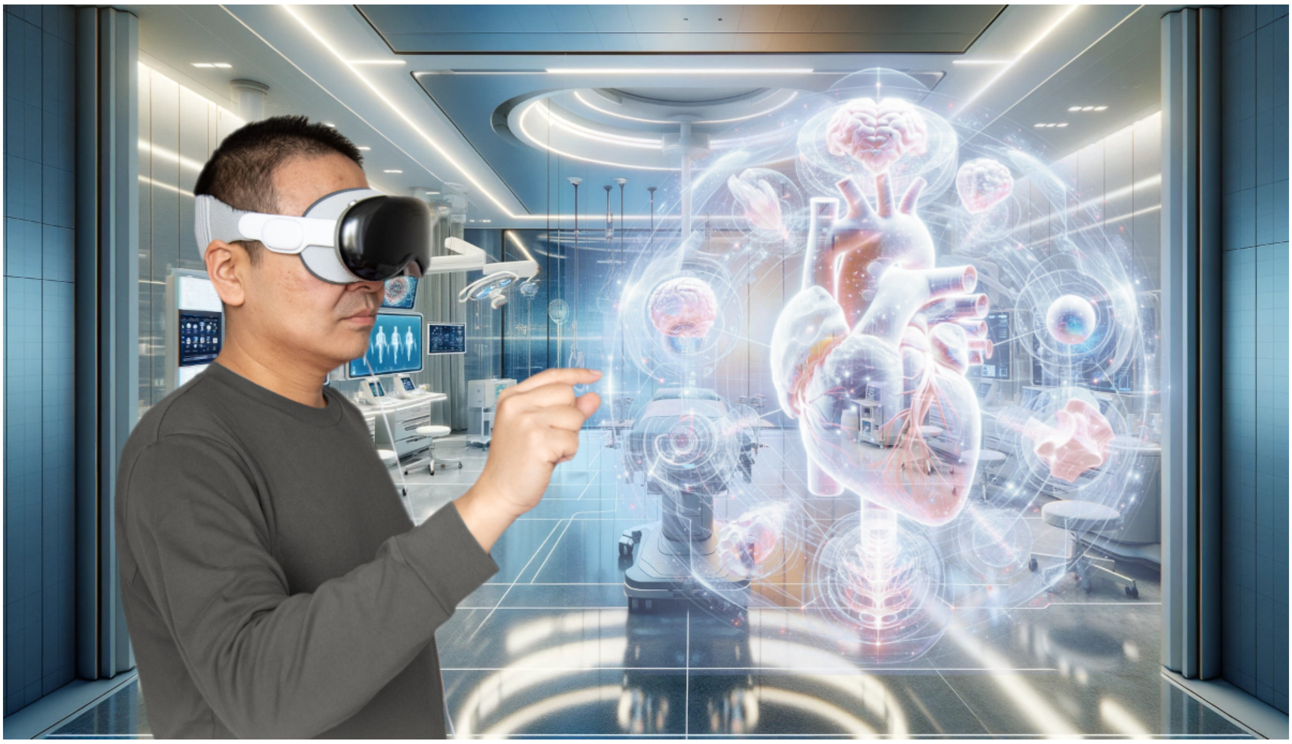


Table 1. Characteristics of the experts interviewed (N=13).

Parameter	Value
Sex	
Female	3 (23)
Male	10 (77)
Age (y), mean (SD; range)	30.9 (6.5; 23-48)
Work experience (y), mean (SD; range)	4.3 (5.6; 0-20)
Previous experience with HMDs[a]	
Yes	9 (69)
No	4 (31)
Previous experience with HMDs (hours)	
Mean (SD)	213.3 (400.8)
Median (range)	1.0 (0-1000)

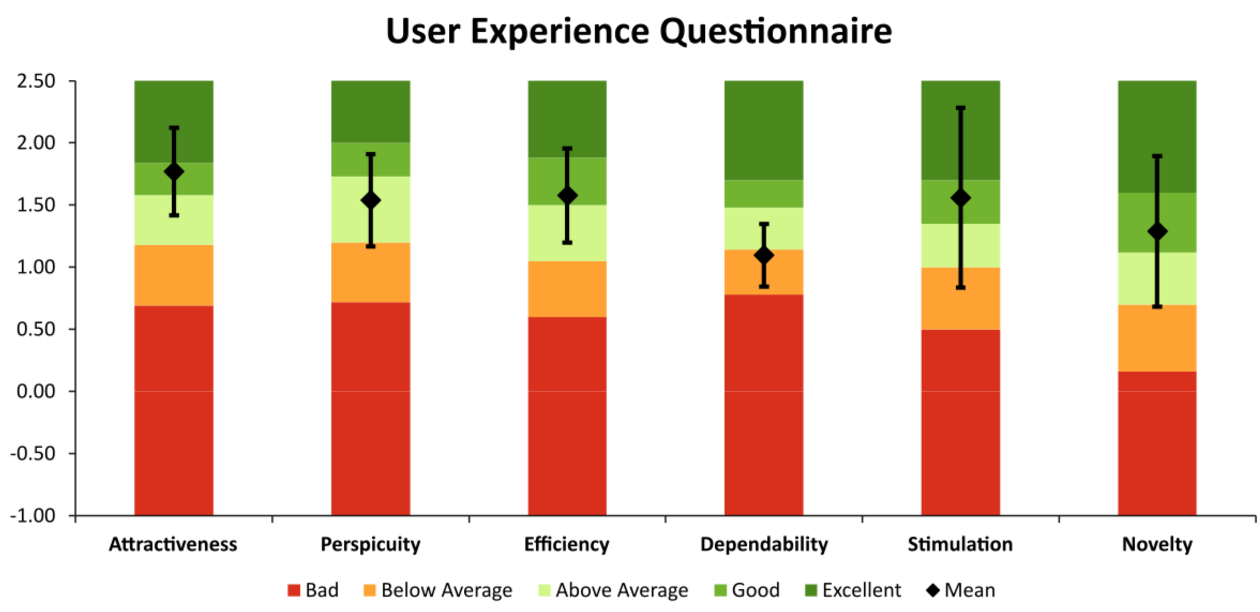
[a]HMD: head-mounted display.

Table 2. Likert questionnaire (N=13).

	Likert item (1=disagree; 7=agree)	Value, mean (SD)
1.	The Apple Vision Pro comes very close to an “Ultimate Display” in terms of realism.	5.5 (1.3)
2.	The Apple Vision Pro comes very close to an “Ultimate Display” in terms of interactivity.	5.0 (1.1)
3.	The Apple Vision Pro is more comfortable than other HMDs[a].	4.9 (1.5)
4.	The Apple Vision Pro is more user-friendly than other HMDs.	5.1 (1.9)
5.	The Apple Vision Pro surpasses other HMDs in immersion and sense of presence.	5.4 (1.6)
6.	The Apple Vision Pro represents a significant technological advance.	5.7 (1.4)

[a]HMD: head-mounted display.

Figure 2. The results (means and SDs) of the User Experience Questionnaire (UEQ) regarding the attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty of the Apple Vision Pro compared to those of a benchmark (bar plot in the background).



Textbox 2. Open questionnaire summary.

Negatives

- Suboptimal weight distribution
- Comparatively heavy among head-mounted display (HMDs)
- Detachable light seal cushion poses potential damage risks and necessitates screen touching before and after use
- Requirement for specialized prescription glasses
- Short battery life (2 hours under normal use)
- Forced interaction method with eyes and pinch
- Inconsistent functionality of eye navigation near the borders of the field of view
- The absence of intelligent automatic transition between real and virtual worlds via the crown button
- Limited field of view compared to other HMDs
- Blurry resolution of real-world pass-through, hindering text legibility
- Clear differentiation between virtual and real environmental components by users

Positives

- Lightweight compared to the currently used hardware and materials
- Natural interaction via outward sensors
- Effective eye-tracking for seamless user interaction
- Rapid and simple adjustment of glasses for user interpupillary distance
- Innovative crown button design with wheel and push functionalities
- Smooth transition between real and virtual worlds facilitated by the crown button
- High-resolution internal display for realistic virtual objects
- High-quality spatial sound experience
- Seamless visualization of the real world
- Minimal app start-up and file loading times
- Autofocus and person blending and unblending for safety precautions
- Simple and fast to learn

Responses to the Likert questions on whether the AVP comes close to an "Ultimate Display" comprised values between 4.9 and 5.7 on a scale of 1 (disagree) to 7 (agree). The UEQ showed good mean results for attractiveness (1.77, SD 0.65), efficiency (1.58, SD 0.70), stimulation (1.56, SD 1.33), and novelty (1.29, SD 1.11) but above average mean result for perspicuity (1.54, SD 0.68) and below average mean result for dependability (1.10, SD 0.46) compared to a benchmark. This fits in with the general perception (Textbox 2) that the AVP is a significant improvement over other HMDs but still behind the "Ultimate Display."

Discussion

To summarize, we see Apple entering the MR headset market with its own device very positively. Another company's contribution can benefit the technical progress in this area, not only from a hardware perspective but also from a software perspective, by enabling new user input concepts and apps. With the current hardware specifications and Apple's history of delivering cutting-edge devices and beyond, there is a very good chance of another AR "evolution" or even "revolution" (Table 3). This is a very important aspect because (1) AR devices are sparse in comparison to VR devices, and (2) the product life cycle to a possible successor is in the range of years, which feels like ages when compared to the current pace of technological developments, especially in (medical) computer science [51]. For example, the first version of the HoloLens was introduced by Microsoft in 2016, and it took Microsoft >3 years to introduce its successor, the HoloLens 2 (released in limited numbers on November 7, 2019).

Table 3. Side-by-side comparison of the specifications of several existing mixed reality headsets with those of the upcoming Apple Vision Pro [52].

Parameter	Apple Vision Pro	Pico 4	Meta Quest 3	HTC Vive XR Elite	Varjo XR-3	HoloLens 2	Magic Leap 2
AR[a]	VST[b]	VST	VST	VST	VST	OST[c]	OST
VR[d]	Yes	Yes	Yes	Yes	Yes	No	No
Resolution (per eye)	3800×3000	2160×2160	2064×2208	1920×1920	1920×1920 focus and 2880 ×2720 peripheral	2048×1080	1440×1760
Refresh rate (Hz)	100	90	120	90	90	120	60
Visible field of view	Estimate 100 horizontal	104 horizontal and 103 vertical	110 horizontal and 96 vertical	102.13 horizontal and 91,27 vertical	115 horizontal and 90 vertical	43 horizontal and 29 vertical	44 horizontal and 53 vertical
Camera	Dual 6.5 MP[e] pass-through cameras	16 MP RGB[f] camera	Dual 18 PPD[g] color pass-through cameras	16 MP RGB camera	Dual 12 MP 90 Hz cameras	Native pass-through 8 MP camera 30 Hz	Native pass-through 12.6 MP 60 fps RGB camera
Tracking type	6 DoF[h] inside-out tracking via 12 integrated cameras; includes depth sensor and LIDAR[i]	6 DoF inside-out tracking	6 DoF inside-out tracking via 4 integrated cameras; includes depth sensor	6 DoF inside-out tracking via 4 integrated cameras; includes depth sensor	6 DoF marker-based tracking and non–marker-based tracking; includes LiDAR sensor	6 DoF inside-out tracking via 4 integrated cameras	6 DoF inside-out tracking
Tracking	Eyes, face, and hands	Hands	Hands, upper body, and leg position	Hands	Eyes and hands	Eyes and hands	Eyes and hands
Depth sensing	LiDAR+TrueDepth	Depth camera	Depth camera	Depth camera	LiDAR+RGB fusion	ToF[j] depth	iToF[k] depth
CPU[l]	Apple M2 and Apple R1	Qualcomm Snapdragon XR2	Qualcomm Snapdragon XR2 Gen 2	Qualcomm Snapdragon XR2 Gen 2	None	Qualcomm Snapdragon 850	Quad-core Zen2
GPU[m]	Apple M2 10-core GPU	Adreno 650	Adreno 740	Adreno 650	None	Adreno 630	AMD GFX10.2
RAM	Up to 16 GB	8 GB	8GB	12 GB	None	4 GB	16 GB
Storage	64 GB	128 or 256 GB	128, 256, and 512 GB	128 GB	None	64 GB	256 GB
OS[n]	visionOS	PICO OS 5.0 (Android)	Android	Android	None	Universal Windows Platform	Lumin OS
SDK[o]	visionOS	PICO Unity Integration SDK and OpenXR	Meta XR All-in-One SDK and OpenXR	Wave andOpenXR	Varjo Native and OpenXR	Windows Mixed Reality platform and OpenXR	Magic Leap and OpenXR
Battery life	2 hours	3 hours	2 hours	2 hours	Wired	3 hours	3.5 hours
Weight	600 g + headband 50 g + external battery 353 g	295 g + headband 291 g	509 g	625 g	594 g + headband 386 g	566 g	260 g
Untethered	Yes	Yes	Yes	No	No	Yes	Yes
Price (US $)	3499	430	499	1100	6500	3500	3299

[a]AR: augmented reality.
[b]VST: video see-through.
[c]OST: optical see-through.
[d]VR: virtual reality.
[e]MP: megapixel.
[f]RGB: red, green, blue.
[g]PPD: peak pixel density.

[h]DoF: degrees of freedom.
[i]LIDAR: light detection and ranging.
[j]ToF: time-of-flight.
[k]iToF: indirect time-of-flight.
[l]CPU: central processing unit.
[m]GPU: graphics processing unit.
[n]OS: operating system.
[o]SDK: software development kit.

Steve Jobs released the iPhone in 2007, creating a new category of smartphones. What made it special was that Apple combined existing technologies and software in a manner that made previously unexpected things possible, which Apple repeated with the iPad in 2010 and then with the Apple Watch in 2015. Hence, the AVP could lead to the breakthrough of not only AR, VR, or MR but also DR (because it is a pass-through technique) due to its enormous technological progress. Moreover, technological progress can not only enable more precise, patient-specific treatments but also make these treatments more efficient, for example, through hands-free, instant authentications; documentation; and the visualization of search queries and their results directly within the field of vision of the HMD user. Further, well-engineered AR can help overcome communication barriers in a face-to-face manner. Thus, it supports clinicians in a range of essential tasks and reduces the administrative burden on clinicians to allow them to spend more time with their patients, which is a key challenge for upcoming artificial intelligence (AI)-based foundation models [53].

Moreover, the recent rise of large language models [54], such as ChatGPT, cannot be ignored in the context of the upcoming AVP because it changed how we communicate with computers, and the AVP can certainly be enhanced with the integration of a ChatGPT-like bot. Even though ChatGPT is still in its infancy, when it comes to hard facts, especially in health care [55,56], it already enables a fluent, partly fact-based, conversation. This will, in the short or long run, also completely enter the headset market, where someone can discuss upcoming treatment steps on a humanlike level based on a trained large language model. However, we do not think that this will be specific to the AVP, and with ChatGPT, Microsoft seems to currently have the edge.

Another application is the metaverse, which integrates physical reality and VR [8]. The idea is that users and their avatars interact in an environment with access to an unlimited amount of health data. Metaverse applications focus primarily on AI-based medical practice and medical image–guided diagnosis and therapy [8]. The AVP could make an important contribution as a hardware enabler that allows users to move and interact in the metaverse in a more intuitive manner.

Nevertheless, digitization holds promise for making medical care more efficient. But will screens come between doctors and patients [57]? At the same time, for every hour doctors spend with their patients, they spend 2 hours in front of the computer [58]. The AVP's untethered, external battery, and spooky-eye design with improved sensor and computing hardware has the potential to help health care professionals improve efficiency and patient outcomes, allowing more time for patient care. The improved hardware should allow the levels of accuracy and computing needed for the device to truly function as a stand-alone device for surgery or preoperative planning. Integration with patient records combined with AI allows for patient recognition and provides the health care professional with accurate data retrieval and documentation. The provision of guidance for devices and medical procedures can occur naturally through its outward-facing sensors, making reading books a thing of the past. The VST design opens up futuristic possibilities for DR, while the ghostly eyes maintain a sense of patient contact. We recognize that this is an optimistic outlook and will require the combined efforts of experts in the field.

Whether the AVP will blur the line between patient encounters and computer time also depends on patient and physician acceptance. In ophthalmology, the AVP is seen as a breathtaking application [59]. However, this must first be confirmed in clinical trials with rigorous scientific methodology and hard clinical end points. When the AVP is used for patient scenarios, the focus must be on the patient, not the technology, which is always a vehicle for existing problems.

Conclusions

To sum up, the AVP enriches the current headset market and provides another alternative, especially in the limited AR segment. The AVP will likely be at the technological forefront in this sector, which puts pressure on other vendors to compete by drawing level with and surpassing Apple. The price is certainly a burden for entertainment consumers, but for professionals, especially in the health care sector, with the Microsoft HoloLens being in the same price range and the Varjo XR-3 costing twice as much, it is not a deal-breaker. This becomes even more *irrelevant* if seen in relation to the other hardware costs in health care, such as those of navigation systems and imaging equipment (computed tomography and magnetic resonance imaging scanners), which can reach millions. In Sutherland's [1] sense and our experience, it is not yet the Wonderland that Alice or the doctors (added by us) walked into, but perhaps a bit closer.

Acknowledgments

This work was supported by the REACT-EU project KITE (Plattform für KI-Translation Essen; EFRE-0801977), FWF enFaced 2.0 (KLI 1044), National Institutes of Health (NIH) and National Institute of Allergy and Infectious Diseases (NIAID;

R01AI172875), and NIH and NCATS (UL1 TR001427). BP was funded by the Medical Faculty of the RWTH Aachen University in Germany as part of the Clinician Scientist Program. CG was funded by the Advanced Research Opportunities Program from the RWTH Aachen University. JC was funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung [BMBF]) in Germany under the funding reference 161L0272 and supported by the Ministry of Culture and Science of the State of North Rhine-Westphalia (Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen [MKW NRW]). Furthermore, the authors acknowledge the Center for Virtual and Extended Reality in Medicine (ZvRM) of the University Hospital in Essen, Germany. Finally, the authors want to thank ZhaoDi Deng, Ahmad Idrissi-Yaghir, and Rainer Röhrig. Parts of Figure 1 were created using the generative artificial intelligence tool DALL·E 3 by OpenAI [60].

Conflicts of Interest

BP is an associate editor of the *Journal of Medical Internet Research*. All other authors declare no other conflicts of interest.

References

1. Sutherland IE. The ultimate display. In: Proceedings of the 3rd IFIP Congress 1965. 1965 Presented at: IFIP Congress 1965; May 24-29, 1965; New York, NY.
2. Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: a class of displays on the reality-virtuality continuum. In: Proceedings of SPIE's 1995 Symposium on Microlithography. 1995 Presented at: SPIE 1995; February 19-24, 1995; Santa Clara, CA. [doi: 10.1117/12.197321]
3. Egger J, Gsaxner C, Kleesiek J, Puladi B. What is diminished virtuality? A directional and layer-based taxonomy for the reality-virtuality continuum. JMIR XR Spatial Comput 2024 Jan 31;1:e52904. [doi: 10.2196/52904]
4. Lungu AJ, Swinkels W, Claesen L, Tu P, Egger J, Chen X. A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery. Expert Rev Med Devices 2021 Jan 16;18(1):47-62. [doi: 10.1080/17434440.2021.1860750] [Medline: 33283563]
5. Livingston MA, Dey A, Sandor C, Thomas BH. Pursuit of "X-ray vision" for augmented reality. In: Huang W, Alem L, Livingston M, editors. Human Factors in Augmented Reality Environments. New York, NY: Springer; 2013.
6. Hansen C, Black D, Lange C, Rieber F, Lamadé W, Donati M, et al. Auditory support for resection guidance in navigated liver surgery. Int J Med Robot 2013 Mar 28;9(1):36-43. [doi: 10.1002/rcs.1466] [Medline: 23192891]
7. Gsaxner C, Li J, Pepe A, Jin Y, Kleesiek J, Schmalstieg D, et al. The HoloLens in medicine: a systematic review and taxonomy. Med Image Anal 2023 Apr;85:102757 [FREE Full text] [doi: 10.1016/j.media.2023.102757] [Medline: 36706637]
8. Wang G, Badal A, Jia X, Maltz JS, Mueller K, Myers KJ, et al. Development of metaverse for intelligent healthcare. Nat Mach Intell 2022 Nov 15;4(11):922-929 [FREE Full text] [doi: 10.1038/s42256-022-00549-6] [Medline: 36935774]
9. Randazzo G, Reitano G, Carletti F, Iafrate M, Betto G, Novara G, et al. Urology: a trip into metaverse. World J Urol 2023 Oct;41(10):2647-2657 [FREE Full text] [doi: 10.1007/s00345-023-04560-3] [Medline: 37552265]
10. Wüller H, Behrens J, Garthaus M, Marquard S, Remmers H. A scoping review of augmented reality in nursing. BMC Nurs 2019 May 16;18(1):19 [FREE Full text] [doi: 10.1186/s12912-019-0342-2] [Medline: 31123428]
11. Gsaxner C, Li J, Pepe A, Schmalstieg D, Egger J. Inside-out instrument tracking for surgical navigation in augmented reality. In: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology. 2021 Presented at: VRST '21; December 8-10, 2021; Osaka, Japan. [doi: 10.1145/3489849.3489863]
12. Ferraguti F, Minelli M, Farsoni S, Bazzani S, Bonfe M, Vandanjon A, et al. Augmented reality and robotic-assistance for percutaneous nephrolithotomy. IEEE Robot Autom Lett 2020 Jul;5(3):4556-4563. [doi: 10.1109/lra.2020.3002216]
13. Li Y, Chen X, Wang N, Zhang W, Li D, Zhang L, et al. A wearable mixed-reality holographic computer for guiding external ventricular drain insertion at the bedside. J Neurosurg 2019 Nov 01;131(5):1599-1606. [doi: 10.3171/2018.4.JNS18124] [Medline: 30485188]
14. Azimi E, Niu Z, Stiber M, Greene N, Liu R, Molina C, et al. An interactive mixed reality platform for bedside surgical procedures. In: Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. 2020 Presented at: MICCAI 2020; October 4-8, 2020; Lima, Peru. [doi: 10.1007/978-3-030-59716-0_7]
15. Van Gestel F, Frantz T, Vannerom C, Verhellen A, Gallagher AG, Elprama SA, et al. The effect of augmented reality on the accuracy and learning curve of external ventricular drain placement. Neurosurg Focus 2021 Aug;51(2):E8. [doi: 10.3171/2021.5.FOCUS21215] [Medline: 34333479]
16. Schneider M, Kunz C, Pal'a A, Wirtz CR, Mathis-Ullrich F, Hlaváč M. Augmented reality-assisted ventriculostomy. Neurosurg Focus 2021 Jan;50(1):E16. [doi: 10.3171/2020.10.FOCUS20779] [Medline: 33386016]
17. Dennler C, Bauer DE, Scheibler AG, Spirig J, Götschi T, Fürnstahl P, et al. Augmented reality in the operating room: a clinical feasibility study. BMC Musculoskelet Disord 2021 May 18;22(1):451 [FREE Full text] [doi: 10.1186/s12891-021-04339-w] [Medline: 34006234]
18. Frey J, Cagle J, Johnson KA, Wong JK, Hilliard JD, Butson CR, et al. Past, present, and future of deep brain stimulation: hardware, software, imaging, physiology and novel approaches. Front Neurol 2022 Mar 9;13:825178 [FREE Full text] [doi: 10.3389/fneur.2022.825178] [Medline: 35356461]

19. van Doormaal TP, van Doormaal JA, Mensink T. Clinical accuracy of holographic navigation using point-based registration on augmented-reality glasses. Oper Neurosurg (Hagerstown) 2019 Dec 01;17(6):588-593 [FREE Full text] [doi: 10.1093/ons/opz094] [Medline: 31081883]
20. Jud L, Fotouhi J, Andronic O, Aichmair A, Osgood G, Navab N, et al. Applicability of augmented reality in orthopedic surgery - a systematic review. BMC Musculoskelet Disord 2020 Feb 15;21(1):103 [FREE Full text] [doi: 10.1186/s12891-020-3110-2] [Medline: 32061248]
21. Elmi-Terander A, Nachabe R, Skulason H, Pedersen K, Söderman M, Racadio J, et al. Feasibility and accuracy of thoracolumbar minimally invasive pedicle screw placement with augmented reality navigation technology. Spine (Phila Pa 1976) 2018 Jul 15;43(14):1018-1023 [FREE Full text] [doi: 10.1097/BRS.0000000000002502] [Medline: 29215500]
22. Li R, Si W, Liao X, Wang Q, Klein R, Heng PA. Mixed reality based respiratory liver tumor puncture navigation. Comput Visual Media 2020 Jan 17;5(4):363-374. [doi: 10.1007/s41095-019-0156-x]
23. Frisk H, Lindqvist E, Persson O, Weinzierl J, Bruetzel LK, Cewe P, et al. Feasibility and accuracy of thoracolumbar pedicle screw placement using an augmented reality head mounted device. Sensors (Basel) 2022 Jan 11;22(2):522 [FREE Full text] [doi: 10.3390/s22020522] [Medline: 35062483]
24. Carbone M, Cutolo F, Condino S, Cercenelli L, D'Amato R, Badiali G, et al. Architecture of a hybrid video/optical see-through head-mounted display-based augmented reality surgical navigation platform. Information 2022 Feb 08;13(2):81. [doi: 10.3390/info13020081]
25. Heinrich F, Huettl F, Schmidt G, Paschold M, Kneist W, Huber T, et al. HoloPointer: a virtual augmented reality pointer for laparoscopic surgery training. Int J Comput Assist Radiol Surg 2021 Jan 23;16(1):161-168 [FREE Full text] [doi: 10.1007/s11548-020-02272-2] [Medline: 33095424]
26. Schoeb DS, Schwarz J, Hein S, Schlager D, Pohlmann PF, Frankenschmidt A, et al. Mixed reality for teaching catheter placement to medical students: a randomized single-blinded, prospective trial. BMC Med Educ 2020 Dec 16;20(1):510 [FREE Full text] [doi: 10.1186/s12909-020-02450-5] [Medline: 33327963]
27. Suzuki K, Morita S, Endo K, Yamamoto T, Fujii S, Ohya J, et al. Learning effectiveness of using augmented reality technology in central venous access procedure: an experiment using phantom and head-mounted display. Int J Comput Assist Radiol Surg 2021 Jun 16;16(6):1069-1074. [doi: 10.1007/s11548-021-02365-6] [Medline: 33864188]
28. Hippe DS, Umoren RA, McGee A, Bucher SL, Bresnahan BW. A targeted systematic review of cost analyses for implementation of simulation-based education in healthcare. SAGE Open Med 2020 Mar 19;8:2050312120913451 [FREE Full text] [doi: 10.1177/2050312120913451] [Medline: 32231781]
29. Aruanno B, Garzotto F. MemHolo: mixed reality experiences for subjects with Alzheimer's disease. Multimed Tools Appl 2019 Jan 5;78(10):13517-13537. [doi: 10.1007/S11042-018-7089-8]
30. Blusi M, Nieves JC. Feasibility and acceptability of smart augmented reality assisting patients with medication pillbox self-management. Stud Health Technol Inform 2019 Aug 21;264:521-525. [doi: 10.3233/SHTI190277] [Medline: 31437978]
31. Sun R, Aldunate RG, Sosnoff JJ. The validity of a mixed reality-based automated functional mobility assessment. Sensors (Basel) 2019 May 11;19(9):2183 [FREE Full text] [doi: 10.3390/s19092183] [Medline: 31083514]
32. Oh H, Son W. Cybersickness and its severity arising from virtual reality content: a comprehensive study. Sensors (Basel) 2022 Feb 09;22(4):1314 [FREE Full text] [doi: 10.3390/s22041314] [Medline: 35214216]
33. Li X, Luh DB, Xu RH, An Y. Considering the consequences of cybersickness in immersive virtual reality rehabilitation: a systematic review and meta-analysis. Appl Sci 2023 Apr 20;13(8):5159. [doi: 10.3390/app13085159]
34. Malhotra S, Halabi O, Dakua SP, Padhan J, Paul S, Palliyali W. Augmented reality in surgical navigation: a review of evaluation and validation metrics. Appl Sci 2023 Jan 27;13(3):1629 [FREE Full text] [doi: 10.3390/app13031629]
35. Seibold M, Spirig JM, Esfandiari H, Farshad M, Fürnstahl P. Translation of medical AR research into clinical practice. J Imaging 2023 Feb 14;9(2):44 [FREE Full text] [doi: 10.3390/jimaging9020044] [Medline: 36826963]
36. Macario Barros A, Michel M, Moline Y, Corre G, Carrel F. A comprehensive survey of visual SLAM algorithms. Robotics 2022 Feb 10;11(1):24. [doi: 10.3390/robotics11010024]
37. Andrews CM, Henry AB, Soriano IM, Southworth MK, Silva JR. Registration techniques for clinical applications of three-dimensional augmented reality devices. IEEE J Transl Eng Health Med 2021;9:1-14. [doi: 10.1109/jtehm.2020.3045642]
38. Gsaxner C, Pepe A, Wallner J, Schmalstieg D, Egger J. Markerless image-to-face registration for untethered augmented reality in head and neck surgery. In: Proceedings of the Medical Image Computing and Computer Assisted Intervention. 2019 Presented at: MICCAI 2019; October 13-17, 2019; Shenzhen, China. [doi: 10.1007/978-3-030-32254-0_27]
39. Xiong J, Hsiang EL, He Z, Zhan T, Wu ST. Augmented reality and virtual reality displays: emerging technologies and future perspectives. Light Sci Appl 2021 Oct 25;10(1):216 [FREE Full text] [doi: 10.1038/s41377-021-00658-8] [Medline: 34697292]
40. Bichlmeier C, Sielhorst T, Heining SM, Navab N. Improving depth perception in medical AR: a virtual vision panel to the inside of the patient. In: Proceedings of the Bildverarbeitung für die Medizin 2007. 2007 Presented at: BVM 2007; March 25-27, 2007; Munich, Germany. [doi: 10.1007/978-3-540-71091-2_44]
41. Innocente C, Ulrich L, Moos S, Vezzetti E. Augmented reality: mapping methods and tools for enhancing the human role in healthcare HMI. Appl Sci 2022 Apr 24;12(9):4295 [FREE Full text] [doi: 10.3390/app12094295]

42. Mori S, Ikeda S, Saito H. A survey of diminished reality: techniques for visually concealing, eliminating, and seeing through real objects. IPSJ T Comput Vis Appl 2017 Jun 28;9:17. [doi: 10.1186/s41074-017-0028-1]
43. Gsaxner C, van Meegdenburg T, Luijten G, Vu VD. DREAMING - diminished reality for emerging applications in medicine through inpainting dataset. Zenodo. 2024 Jan 8. URL: https://zenodo.org/records/10471365 [accessed 2024-09-06]
44. Cutolo F, Cattari N, Fontana U, Ferrari V. Towards the development of a quasi-orthoscopic hybrid video/optical see-through HMD for manual tasks. In: Proceedings of the Augmented Reality, Virtual Reality, and Computer Graphics. 2019 Presented at: AVR 2019; June 24-27, 2019; Santa Maria al Bagno, Italy. [doi: 10.1007/978-3-030-25999-0_15]
45. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inform Syst 1994;77:1321-1329.
46. Kottner S, Thali MJ, Gascho D. Using the iPhone's LiDAR technology to capture 3D forensic data at crime and crash scenes. Forensic Imaging 2023 Mar;32:200535. [doi: 10.1016/j.fri.2023.200535]
47. Sproull R, Brock DC. Interview of Ivan Sutherland. IEEE Ann Hist Comput 2023;45(1):64-76. [doi: 10.1109/MAHC.2023.3244258]
48. Egger J, Gall M, Wallner J, Boechat P, Hann A, Li X, et al. HTC Vive MeVisLab integration via OpenVR for medical applications. PLoS One 2017 Mar 21;12(3):e0173972 [FREE Full text] [doi: 10.1371/journal.pone.0173972] [Medline: 28323840]
49. Boonstra A, Versluis A, Vos JF. Implementing electronic health records in hospitals: a systematic literature review. BMC Health Serv Res 2014 Sep 04;14(1):370 [FREE Full text] [doi: 10.1186/1472-6963-14-370] [Medline: 25190184]
50. Jha AK, DesRoches CM, Campbell EG, Donelan K, Rao SR, Ferris TG, et al. Use of electronic health records in U.S. hospitals. N Engl J Med 2009 Apr 16;360(16):1628-1638. [doi: 10.1056/nejmsa0900592]
51. Egger J, Gsaxner C, Pepe A, Pomykala KL, Jonske F, Kurz M, et al. Medical deep learning-a systematic meta-review. Comput Methods Programs Biomed 2022 Jun;221:106874 [FREE Full text] [doi: 10.1016/j.cmpb.2022.106874] [Medline: 35588660]
52. VRcompare. URL: https://vr-compare.com/ [accessed 2024-02-22]
53. Moor M, Banerjee O, Abad ZS, Krumholz HM, Leskovec J, Topol EJ, et al. Foundation models for generalist medical artificial intelligence. Nature 2023 Apr 12;616(7956):259-265. [doi: 10.1038/s41586-023-05881-4] [Medline: 37045921]
54. Brants T, Popat AC, Xu P, Och FJ, Dean J. Large language models in machine translation. In: Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. 2007 Presented at: EMNLP-CoNLL 2007; June 28-30, 2007; Prague, Czech Republic.
55. Li J, Dada A, Puladi B, Kleesiek J, Egger J. ChatGPT in healthcare: a taxonomy and systematic review. Comput Methods Programs Biomed 2024 Mar;245:108013 [FREE Full text] [doi: 10.1016/j.cmpb.2024.108013] [Medline: 38262126]
56. Barnard F, Van Sittert M, Rambhatla S. Self-diagnosis and large language models: a new front for medical misinformation. arXiv. Preprint posted online July 10, 2023 2024 [FREE Full text]
57. Gawande A. Why doctors hate their computers: digitization promises to make medical care easier and more efficient. But are screens coming between doctors and patients? The New Yorker. 2018 Nov 5. URL: https://www.newyorker.com/magazine/2018/11/12/why-doctors-hate-their-computers#:~:text=A%202016%20study%20found%20that,screen%20to%20do%20electronic%20tasks [accessed 2024-09-06]
58. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Sep 06;165(11):753. [doi: 10.7326/m16-0961]
59. Waisberg E, Ong J, Masalkhi M, Zaman N, Sarker P, Lee AG, et al. The future of ophthalmology and vision science with the Apple Vision Pro. Eye (Lond) 2024 Feb;38(2):242-243 [FREE Full text] [doi: 10.1038/s41433-023-02688-5] [Medline: 37542175]
60. DALL·E 3. OpenAI. URL: https://openai.com/index/dall-e-3/ [accessed 2024-09-06]

Abbreviations

AI: artificial intelligence
AR: augmented reality
AVP: Apple Vision Pro
DR: diminished reality
HMD: head-mounted display
MR: mixed reality
OST: optical see-through
VR: virtual reality
VST: video see-through
XR: extended reality

Edited by T Leung, G Eysenbach; submitted 15.09.23; peer-reviewed by M Hirota, N Silva; comments to author 12.11.23; revised version received 26.03.24; accepted 02.07.24; published 18.09.24.

Please cite as:
Egger J, Gsaxner C, Luijten G, Chen J, Chen X, Bian J, Kleesiek J, Puladi B
Is the Apple Vision Pro the Ultimate Display? A First Perspective and Survey on Entering the Wonderland of Precision Medicine
JMIR Serious Games 2024;12:e52785
URL: https://games.jmir.org/2024/1/e52785
doi:10.2196/52785
PMID:39292499

Assessing the Importance of Content Versus Design for Successful Crowdfunding of Health Education Games: Online Survey Study

Hong Huang[1], PhD; Han Yu[2], PhD; Wanwan Li[3], PhD

1
2
3

Corresponding Author:
Hong Huang, PhD

Abstract

Background: Health education games make health-related tasks enjoyable and interactive, thereby encouraging user participation. Entrepreneurs and health educators can leverage online crowdfunding platforms, such as Kickstarter, to transform their innovative ideas into funded projects.

Objective: This research focuses on health education game initiatives on Kickstarter. Through an online user survey, it aims to understand user perceptions and evaluate the significance of 8 distinct components that may influence the success of such crowdfunding initiatives.

Methods: A total of 75 participants evaluated games using 8 dimensions: game rules, learning objectives, narrative, content organization, motivation, interactivity, skill building, and assessment and feedback. The survey data were analyzed using descriptive statistical analysis, exploratory factor analysis, the Wilcoxon-Mann-Whitney test, and multivariate analysis.

Results: Exploratory data analysis showed that, among the 8 dimensions, skill building, content organization, and interactivity were the top-ranking dimensions most closely associated with crowdfunding health education game. The 8 dimensions can be grouped into 3 categories from the exploratory factor analysis: game content–, instruction-, and game design–related components. Further statistical analysis confirmed the correlation between these dimensions with the successful crowdfunding of health education games.

Conclusions: This empirical analysis identified critical factors for game proposal design that can increase the likelihood of securing crowdfunding support.

(JMIR Serious Games 2024;12:e39587) doi:10.2196/39587

KEYWORDS

game-based learning; rubrics; Kickstarter; learning game campaign; collaboration; user perception; design; health; learning; gaming; game; evaluation; organization; user; engagement; skill; feedback; assessment; analysis; correlation; crowdfunding; support

Introduction

Background

Digital strategies, particularly gamification, have introduced a refreshing dynamic to health education [1,2]. Platforms, such as Kickstarter [3], champion these tech-infused health games, providing a unique avenue for their development. By leveraging the power of crowdfunding, Kickstarter and similar platforms facilitate the evolution of health education games. This allows entrepreneurs, educators, game developers, and supporters to access essential resources and connect with audiences eager for meaningful health support and intervention.

Gamification in Health

Gamification in health integrates game-design elements into nongame health scenarios, aiming to boost user engagement and immersion in health solutions. This transforms routine health tasks into enjoyable, competitive activities. This approach leads to positive behavioral changes, improving overall health, fitness, and adherence to medical treatments and programs [1,2,4,5]. Gamification has been applied to a wide range of medical fields, including health education, medical therapy, obesity, and mental health [1,2,4,5].

Health education games are interactive digital tools specifically designed to impart knowledge or skills related to health and wellness [1,2]. These games transform traditional health-related lessons into enjoyable and engaging tasks, aiming to enhance retention and application of health information in daily life [1,2,4]. Serious health games, created primarily for specific health objectives rather than solely for entertainment, use gaming components to create an educational environment [1,2]. They use gaming components to facilitate a teaching environment, enabling users to learn specific health skills or gain valuable health-related information [1,2]. Especially beneficial for long-term health and chronic-related applications, these games foster positive emotional or empathetic connections among users, leading to improved medical treatment plans and behavior changes [1,2,4].

The Role of Crowdfunding in Promoting Health Education Games

Given the modest initial investment required and the scale of crowdfunding, it is advocated that crowdfunding serves as a primary method to promote and support the development of health education games. With the recent success of platforms such as Kickstarter, researchers and health care advocates are turning to these tools to fund their projects [6,7]. Through crowdfunding, health educators, entrepreneurs, and other stakeholders can conduct their work to meet community needs while also achieving financial and community outreach goals. This method attracts a varied group of participants who contribute financially, participate in the development, and offer social support [8-12].

Health education games, similar to other game-based learning tools, motivate users by making health-related tasks more enjoyable [1,13]. Online crowdfunding can assist entrepreneurs and health educators with limited resources to translate their innovative ideas into solid and appealing content and formats [14,15]. Crowdfunding platforms help individuals transform ideas into fundable and actionable projects [16,17].

Crowdfunding for health education games benefits users' self-efficacy, well-being, chronic disease management, and physical activity [9-11]. Rewards, feedback, and socialization elements are frequently used to gamify eHealth in crowdfunding-based health education games. Furthermore, health education games can positively change their health behaviors, benefiting their overall health and wellness [2]. Successful health education crowdfunding projects elicit both intrinsic (altruistic) and extrinsic (rewards and feedback) motivation in order to attract a diverse range of crowdfunding donors, and they work by effortlessly facilitating online digital health engagement [18]. This study aims to explore 8 critical evaluation dimensions from the user's perspective that influence the success of crowdfunding campaigns for health education games. The findings will guide practitioners and entrepreneurs in strategizing and designing impactful crowdfunding campaigns for health education games.

Related Works

To understand the intricacies of successful crowdfunding for health education games, we performed a literature review to acquire insights on the various dimensions related to the subject. The literature review enabled us to systematically explore the dynamics of crowdfunding, the principles of game-based learning, and the factors that influence the success of health education games.

Dynamics and Success Factors of Crowdfunding Initiatives

To develop and promote content for successful crowdfunding campaigns, extensive planning, outreach, and marketing are required. Data suggest that the most popular crowdfunding projects are those that are creative, participatory, or consumable, such as games, technology, film and video, and art and design [19]. In general, crowdfunding projects have small funding sizes and offer various donor incentives, small gifts, or awards, which leads to a higher success rate for the projects [19]. Such success not only mirrors financial objectives but also nurtures the emergence of communities with shared interests [20]. Numerous game developers have used crowdfunding to fund the initial investment in educational applications [21]. This then encourages more entrepreneurs to participate in collaborative crowdfunding platforms and launch their projects.

Unlike a traditional purchase, crowdfunding involves a high level of social capital influence, particularly the status and reach on social network sites [17]. Social capital creates an online environment that combines collective knowledge, appeal, and emotional responses, enabling investors to make well-informed decisions [17]. This investment process shapes perception and investment behavior. The interaction mechanism has a broader and more pervasive contextual impact, and the crowdfunding campaign design and features also influence decisions [21].

Crowdfunding initiatives require both content richness and ownership diversity [22]. Several studies have explored strategies to optimize the success of such crowdfunding efforts [22,23]. Notably, during crowdfunding, potential investors often evaluate founders based on their personal communication skills and presentation, both of which influence investment decisions [24]. In addition, the use of specific language, the length of campaign text, the frequency of updates, and the inclusion of video in campaign texts have all been correlated with the success of crowdfunding campaigns [25,26]. Reducing the cognitive effort needed to understand campaign content has been shown to result in increased funding [15].

Researchers have also linked crowdfunding success to the trustworthiness and reputation of developers, as well as their experiences on social crowdfunding networks [27,28]. However, the quality of the presented information also plays an important role in determining crowdfunding success [9,29,30]. Factors that contribute to successful crowdfunding factors include the content of the campaign, audience participation, and the timing of fundraising development [31].

Health Education Game Development and User Experience

Gamification has been proven to enhance medication and treatment adherence among patients with chronic disease [4,32]. Health serious games, on the other hand, have been praised for their ability to help people with chronic illnesses improve their behavior [2,33]. These games mirror real-life challenges, allowing players to develop coping strategies [17]. They educate players about their condition and the necessary lifestyle alterations, with compelling storylines that ensure better engagement [15,17]. Game interactivity allows players to make decisions, learn from outcomes, and receive feedback on health implications [2,17].

When evaluating the feasibility of a game proposal, it is important to consider both the organization and narrative of the content, as well as the effectiveness of interactive games as a learning tool [34]. A well-organized and clearly written proposal can help the investor understand the purpose, goals, and potential value of the project [31]. Interactive health games can educate users with content and skills [34]. Users can also actively engage with the material, explore and experiment with different concepts and strategies, and receive immediate feedback on their progress [35]. This can help them understand and retain the content and skills being taught.

Game rules and interactivity stand as important components in health game design. Game rules ensure alignment with educational objectives, and the inherent challenge-reward system in these games drives players to continue, thereby continuously learning and adopting healthier behaviors [5,13]. Defining game rules or challenges and delivering feedback can increase users' self-concept, efficacy, knowledge skills, communication, and social support, resulting in better health behaviors for self-care and adherence, lowering health costs, and establishing a stronger health system [18].

Health education game users are drawn to characters that resemble them, experiencing validation when such characters are featured in media [36]. Young role models, especially those in media genres such as cartoons and video games, are particularly valued by these users [37]. For example, the motivation and design of the interactive health game series can focus on using positive role models to inspire and motivate players [37,38]. These role models are described as being successful in their adventures while also managing their health, which could help users, including children with chronic illnesses such as asthma or diabetes, feel more positive about their own abilities to manage their health and self-care [37,38].

Interestingly, health game players without specific medical conditions are often less certain about in-game decisions compared to their peers with those conditions [37]. Health education games allow players to try new things, fail, learn, and eventually win. Such games also motivate users to adopt a healthier lifestyle, adhere to medical advice when unwell, navigate life crises, and foster close social connections for support [39].

Regarding assessments and feedback mechanisms, health learners who receive personalized feedback and engage deeply with medical content tend to experience great benefit. This approach is especially effective in reaching younger individuals who might not typically consult other media or seek expert health advice [40]. Interactive health games not only foster communication and social support but also empower users to discuss their health with friends, family, and health care professionals. They also motivate users to actively seek out advice and support [37]. For instance, in a series of interactive health games, players accessed factual details about the causes, treatments, social contexts, and self-care options related to specific health topics [37].

Game-Based Learning Principles

One of the game-based learning principles that allows users to benefit from the game is the development of problem-solving skills [41], and educational games can assist users in developing these skills [41,42]. The modalities of game content representation should be adjusted to boost motivation and performance [43]. If learners cannot understand the app's content, no matter how rich and useful it is or how beautiful the design is, the app's entire instructional value is lost [44]. Learners can learn problem-solving, strategic and analytical thinking, decision-making, and other 21st century skills in narrative-centered learning environments [45].

Based on the constructivist learning theory, individuals gain deeper insights about the world through direct experiences and interactions [46,47]. Games offer a dynamic and interactive environment that aligns with this theory, enabling learners to actively explore, experiment, and tackle challenges [46,47]. The appeal of a game's narrative indicates its potential to captivate users [48]. The game creators should focus more on the content, storyline, and interaction components of the game to attract individual users when determining whether it will be successful or not [48].

The quality of a learning game is significantly influenced by the effectiveness of user feedback [49,50]. Numerous studies have shown that feedback enhances learning outcomes [51]. It provides learners with clarity on their strengths and areas that need improvement; it also serves as a motivational tool, encouraging continuous learning even within the gaming context [51].

Educational games can customize learning experiences by gauging a student's readiness, providing constructive feedback, and modifying the level of challenge [52]. It is essential for an educational game to have well-defined learning objectives that detail the desired skills and knowledge [53]. Game rules facilitate learning by allowing players to interact with their environment [54]. Achieving these objectives depends on adhering to specific rules, which may involve certain challenges or conditions that the learner must satisfy [43].

A learner's level of motivation can greatly influence their enthusiasm or indifference toward a task [55,56]. Moreover, there is substantial evidence suggesting that motivation enhances cognitive functions, particularly influencing what learners focus on and how they assimilate information [57-59].

Literature suggests that multiple factors influence the success of crowdfunding campaigns, especially those related to health education games [51,54]. These range from the trustworthiness of the developers and the quality of information presented to the design and content of the game itself. Although previous studies have shed light on the general principles of game-based learning and the dynamics of crowdfunding, there remains a gap in understanding how these principles specifically apply to health education games on platforms such as Kickstarter. Moreover, the user's perspective, which is crucial in determining the success of such campaigns, has not been thoroughly explored. We aim to bridge this knowledge gap by focusing on the user's perception and evaluating the critical components that resonate most with potential users, thereby influencing the success of health education game initiatives on crowdfunding platforms.

Objectives

This study aims to provide a comprehensive overview of 17 health education game projects launched on the crowdfunding platform Kickstarter and to understand user perceptions concerning the important factors that determine the success of such health education game crowdfunding initiatives. To achieve this, we conducted a user survey using a health education assessment rubric specifically designed to evaluate the key components contributing to the success of these projects on Kickstarter.

Methods

Data Collection for Health Education Games

A comprehensive keyword search using "Health, Education, Learning, Game" was conducted in August 2019 on Kickstarter, which identified 17 online health education game projects (Table 1). On the Kickstarter site, the system marked a project as "Successful" if it met or exceeded its financial goal within the time set by the creators. Conversely, projects that failed to meet their financial target within the designated period were labeled as "Unsuccessful" (Table 1).

Table . Descriptive data of health education game projects from the crowdfunding site Kickstarter. A project's success on Kickstarter was determined by its ability to achieve its financial goal within the set time frame.

Health education game	Pledge (US $)	Goal (US $)	Backer count, n	Country	Successful[a]
Playout: The Exercise Card Game [60]	11,011	10,000	224	United States	Yes
ACLS MegaCode Simulator for health care professionals [61]	328	1997	8	Canada	No
Blush by Renaissance [62]	5065	3500	80	Canada	Yes
Body Cycle Health Education App [63]	1778	20,000	41	United States	No
CHiLD - a psychological 2D RPG [64]	1199	554	92	Norway	Yes
Destiny's Sword for mental health [65]	30,930	30,000	209	Canada	Yes
Facing Dragons: a mixed-reality game to unlock your purpose [66]	3361	7104	34	Canada	No
Freestyle Jam Camp [67]	1145	500	18	United States	Yes
Mobile games to quantify symptoms of mental health disorders [68]	127	450,000	5	United States	No
PRESCRIPTION Playing Cards [69]	30,420	7500	178	Canada	Yes
Talk to Me visual novel: mental health [70]	4977	4460	146	United States	Yes
TEN: a card game designed to promote brain health [71]	1445	14,000	39	United States	No
The Chakra Collectable Coin [72]	1682	1300	41	United States	Yes
The Woosah Kit: a mental health first aid [73]	41	6236	3	United Kingdom	No
Tournesol Kids Game: activity cards to build resilience [74]	10,435	5000	140	United States	Yes
Youth Run The World 5K [75]	7370	7000	74	United States	Yes
Zombied: gamify health and fitness activities [76]	12	37,217	2	United Kingdom	No

[a]"Yes" refers to "Successful" projects that met or exceeded their financial goal, whereas "No" refers to "Unsuccessful" projects that did not.

Ethical Considerations

Before commencing this study, the researchers obtained approval from the Institutional Review Board of the University of South Florida (001588). The participants provided informed consent, with the option to withdraw at any time without penalty. The Institutional Review Board approval sufficiently covered the secondary use of data. The study guaranteed that all collected data were either anonymized or deidentified to protect personal information, with stringent protective measures in place for any data that could not be fully anonymized. The study was voluntary, without any compensation for participation.

Online Survey Design

We use the Qualtrics online survey platform (Qualtrics) to create an online survey based on health education game assessment rubrics derived from the literature. This survey allowed participants to evaluate and rank crowdfunding health education games on the Kickstarter website. The survey incorporated 8 dimensions—each essential for the evaluation of health education games. These dimensions, along with their definitions and cited literature, are presented in Table 2.

Table . Crowdfunding health education game evaluation dimensions and definitions.

Dimensions	Definition	Related literature
Skill building	The game's ability to progressively impart and reinforce health-related skills to players, ensuring that learning is continuous and effective throughout the game's duration.	[42,77]
Content organization	The clarity, structure, and logical flow of the game's health education content, ensuring that it is presented in a manner that is both comprehensible and engaging for players.	[35,53,78]
Narrative	The clarity and continuity of the game's storyline in relation to health education, ensuring that players experience a coherent sense of progression and purpose as they navigate through the game's content.	[25,48]
Interactivity	The game's ability to facilitate effective interactions, the completion of health-related tasks, and active participation through user-driven inputs and actions.	[35,77,79]
Assessment and feedback	The game's capability to immediately evaluate and communicate a player's progression and provide timely and relevant feedback.	[35,80-82]
Game rules	The game provides clear, concise, and easily comprehensible rules to the players.	[35,52,83,84]
Learning objectives	The game delineates specific, measurable outcomes that players are anticipated to achieve upon its completion.	[35,85-87]
Motivation	The game's elements are intriguing and appealing enough to prompt user participation and action.	[88-90]

Before the main survey was launched, a pilot test of the survey instrument was conducted with 7 undergraduate students majoring in health science. This pilot test aimed to assess the validity and understandability of the survey questions. The participants were asked to read through the survey and provide feedback on its clarity and relevance. Based on their comments, necessary revisions were made to the questions to enhance the overall quality of the survey.

In the final version of the survey, participants rated the dimensions on a 3-point Likert scale. The scoring system for these dimensions ranged from 0 to 2, with the following interpretations: 0="Does not meet expectations" or "Poor," 1="Meets expectations" or "Fair," and 2="Exceeds expectations" or "Good." Participants could also select "Unable to decide" or "Not applicable" if they felt unable to make a judgment on a particular dimension. Additionally, an open-ended question was incorporated: "Do you have any comments or concerns (accuracy of terms, comprehensiveness, clarity of questions, etc) for this question sets?" This allowed participants to provide further feedback on the survey questions.

In November 2019, undergraduate students majoring in health science were invited to participate in the online survey. Those who agreed to participate were provided with a standardized set of questions, accompanied by comprehensive instructions and definitions for the 8 evaluation dimensions, as detailed in Table 2. Each student was then randomly assigned 1 specific crowdfunding health education game from a pool of 17 games, referenced in Table 1. Their task was to evaluate their assigned game based on these 8 dimensions. Ultimately, 75 undergraduate students were recruited as participants.

Data Analysis

We used STATA 15 software (StataCorp) for statistical analyses. We used several data analysis approaches to understand the results.

Descriptive Statistical Analysis

This method provides a summary of the main aspects of the data, offering a simple overview of the data. By calculating the percentage of ranking types and the mean scores of the dimensions, we can gain insights into the general behavior and preferences of the survey participants.

Exploratory Factor Analysis

Exploratory factor analysis is used to reduce the data's dimensionality and identify the underlying relationships between the measured variables [91]. It was used to group the 8 dimensions into meaningful categories, helping to decipher any latent structures within the data set. This ensured that we could identify which sets of dimensions tended to co-occur or were rated similarly by participants.

Wilcoxon-Mann-Whitney Test

The Wilcoxon-Mann-Whitney test [92] is a nonparametric statistical hypothesis test used to compare 2 unrelated samples. This test was used to determine if there were any significant differences in the rankings given by participants to different game dimensions.

Multivariate Analysis

The aim of this study extends beyond merely understanding the dimensions. It also seeks to predict the success of crowdfunding health education games based on these dimensions. We used logistic regression with a binary variable—success of the crowdfunding project—for prediction [91]. This model can determine the odds of a game being successful based on the rankings of its dimensions, offering insights into which dimensions are the most influential predictors of success.

By using these methods, the study ensured a comprehensive analysis of the data—from understanding the basic patterns and deciphering underlying component structures to finally being able to predict the success of crowdfunding health education games based on their dimensions.

Results

A list of health education games launched on Kickstarter is presented in Table 1. This table enumerates 17 distinct health education games originating from various countries, namely the United States, Canada, Norway, and the United Kingdom. Some projects have exceeded their goals by a large margin, whereas others have fallen substantially short. The diversity of the sample provides a comprehensive foundation for our study. This diversity enabled an exploration into users' perceptions regarding educational game assessment rubrics. Such an investigation can discern potential factors that could influence the success trajectory of health education games on crowdfunding platforms such as Kickstarter.

Table 2 focuses on the various dimensions relevant to the design and evaluation of games. These dimensions were based on established literature, highlighting their credibility and validity. When assessing potential predictors of crowdfunding success based on feedback from 75 survey participants, certain dimensions stood out as being more important (Table 3).

Table . Ranking of the 8 assessed dimensions for crowdfunding health education games (n=75).

Dimensions	Score, mean (SD)[a]
Skill building	1.77 (0.54)
Content organization	1.7 (0.52)
Narrative	1.51 (0.69)
Interactivity	1.51 (0.75)
Assessment and feedback	1.49 (0.69)
Game rules	1.47 (0.71)
Learning objectives	1.39 (0.64)
Motivation	1.29 (0.59)

[a]Scoring system: 0=”Poor,” 1=”Fair,” and 2=”Good.”

Skill building was ranked first, followed by content organization and then narrative. Skill building holds the top rank due to its emphasis on continuous learning and engagement, ensuring that players progressively acquire and refine their skills throughout the game (Table 3). The importance of content organization is highlighted by its role in enhancing user experience; a well-organized game offers clear navigation, allowing players to immerse themselves fully (Table 3). Narrative further enhances the gaming experience by introducing an engaging storyline that lends context and purpose, enriching the gameplay. Interactivity is important for keeping players engaged. It gives them a sense of belonging and influence within the game world. Yet, intriguingly, motivation ranks the lowest among these dimensions, even though its presence ensures that games are compelling enough to retain players' interest and drive continuous participation (Table 3). Although skill building and content organization seem to be the areas where these games excel, motivation appears to be a challenging area for many developers.

To identify the assessment structure for campaign initiatives' quality reflected by 75 survey respondents' rankings, the study conducted an exploratory factor analysis using principal-components analysis as the extraction method and varimax with Kaiser normalization as the rotation method (Table 4). The cutoff size for criterion loadings was set to 0.45 [59]. Both the Bartlett (χ^2=68.26, P<.001) and measure of sampling adequacy (0.57) tests for the sample pointed to a significant level of correlation among the dimensions.

Table . Factor components for the 8 dimensions in crowdfunding health education games. Principal-components analysis served as the extraction method, and varimax with Kaiser normalization served as the rotation method.

Dimensions	Component		
	1	2	3
Game rules	−0.050	−0.057	0.843[a]
Learning objectives	0.253	0.727[a]	−0.025
Narrative	0.181	0.629[a]	−0.335
Motivation	0.665[a]	0.102	−0.021
Interactivity	0.489	0.185	0.506[a]
Skill building	−0.136	0.727[a]	0.266
Assessment and feedback	0.883[a]	−0.036	−0.019
Content organization	0.490[a]	0.370	0.033

[a]Values above the cutoff size for criterion loadings (0.45).

The exploratory factor analysis indicated that these 8 dimensions can be grouped into 3 components: game content (content organization, motivation, and assessment and feedback), instruction (learning objectives, narrative, and skill building), and game design (game rules and interactivity; Table 4). The game content–related components suggests that a well-organized game with clear feedback mechanisms can effectively motivate players. The instruction-related components reflect the instructional journey of the player, from understanding the objectives and engaging with narrative to building skills. The game design–related components are fundamental to the gameplay experience, ensuring that players are not just passive observers but active participants.

To review the perception gaps among these dimensions for successful or unsuccessful crowdfunding campaigns, group-based comparison was conducted between these dimensional means. Table 5 showed the gaps between successful and unsuccessful games in dimension ratings. Among them, motivation, interactivity, game rules, and learning objectives demonstrated larger difference gaps in decreasing order, and these were followed by assessment and feedback, skill building, narrative, and content organization.

Table . Wilcoxon-Mann-Whitney test of the 8 assessments based on successful or unsuccessful crowdfunding of health education games.

Dimensions and categories		Answers, n	Score, mean (SD)	*U* statistic	*P* value
Content organization				0.28	.78
	Success	53	1.68 (0.55)		
	Unsuccessful	15	1.67 (0.49)		
Interactivity				2.05	.04[a]
	Success	53	1.57 (0.72)		
	Unsuccessful	15	1.13 (0.83)		
Skill building				0.94	.35
	Success	53	1.79 (0.49)		
	Unsuccessful	15	1.60 (0.74)		
Learning objectives				2.03	.04[a]
	Success	51	1.43 (0.64)		
	Unsuccessful	15	1.07 (0.59)		
Narrative				.09	.37
	Success	53	1.53 (0.70)		
	Unsuccessful	14	1.36 (0.74)		
Motivation				2.91	.004[a]
	Success	53	1.38 (0.56)		
	Unsuccessful	15	0.87 (0.52)		
Game rules				2.14	.03[a]
	Success	53	1.55 (0.70)		
	Unsuccessful	15	1.13 (0.74)		
Assessment and feedback				1	.32
	Success	53	1.49 (0.64)		
	Unsuccessful	14	1.29 (0.73)		

[a]Significant level $P<.05$.

The Wilcoxon-Mann-Whitney test comparing distributions of successful and unsuccessful games showed that motivation (P=.004), game rules (P=.03), learning objectives (P=.04), and interactivity (P=.04) showed statistically significant difference among these 2 groups (Table 5). These dimensions showed clear distinctions between successful and unsuccessful games, suggesting that these dimensions might be crucial for the success of such games. On the other hand, dimensions such as content organization and skill building, while important, did not show a significant difference between the 2 categories of games. This could mean that both successful and unsuccessful games have well implemented these dimensions, but they might not be the distinguishing factors for success. The multivariate analysis showed that learning objectives and motivation were 2 significant dimensions associated with successful health education game crowdfunding campaigns (Table 6). This suggests that these 2 dimensions might be especially important for the success of health-related games.

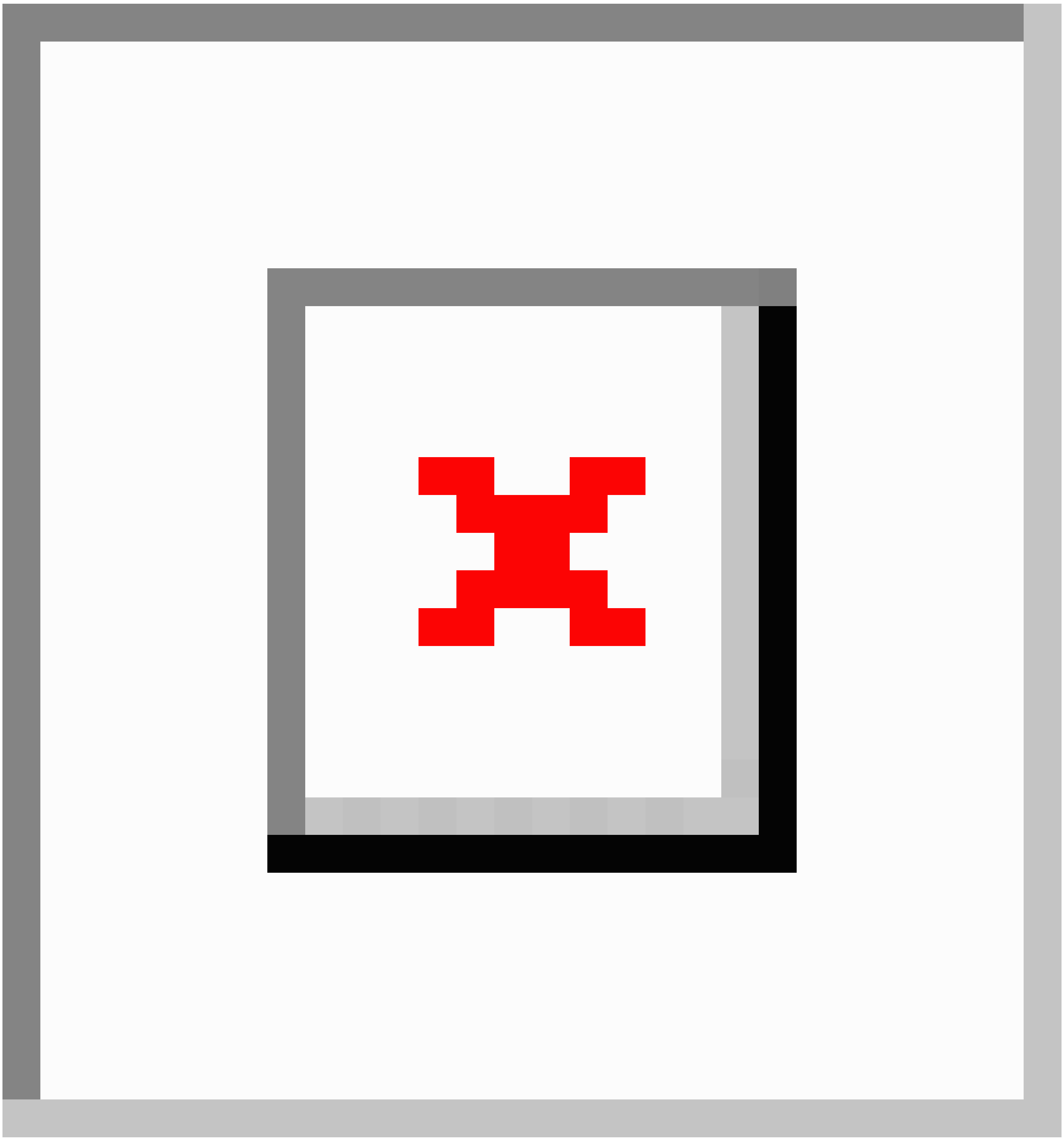
Table . Multivariate logistic regression predicting the success of the health educational games.

Dimensions	Odds ratio (95% CI)	*P* value
Game rules	3.24 (0.68-18.40)	.13
Learning objectives	3.55 (1.42-14.38)	.02[a]
Narrative	1.57 (0.37-6.61)	.54
Content organization	0.07 (0.01-1.55)	.09
Motivation	3.05 (1.46-9.36)	.03[a]
Interactivity	1.70 (0.46-6.22)	.42
Skill building	1.31 (0.21-8.18)	.77
Assessment and feedback	1.38 (0.21-8.88)	.73

[a]Significant level $P<.05$.

Figure 1 presents an empirical framework that outlines the key components underpinning the success of health education game crowdfunding. The model highlights the balance between foundational structural components, such as game rules and content organization, and experiential elements that enhance the player's immersion and engagement, such as motivation and narrative. A successful educational game should seamlessly integrate all these facets. This not only ensures the delivery of educational content but also fosters an environment where players are intrinsically driven to remain engaged and continue their learning journey within the game.

Figure 1. Framework for successful crowdfunding of health education games.



Discussion

Principal Findings

The crowdfunding landscape for health education games is diverse, with success determined by a myriad of factors beyond just a funding goal. Factors such as the clarity of the project's purpose, its presentation, and its marketing likely play a substantial role in attracting users [35,41]. It is also important to have a reasonable and attainable goal, as this might increase the likelihood of a project's success.

Crowdfunding backers, especially on platforms such as Kickstarter, often support projects that offer value beyond just entertainment. Skill building in games implies that players will acquire new abilities or knowledge, making them both fun and beneficial. This dual-purpose might appeal to game players who see an opportunity for a return on investment, not just in potential product rewards but also in personal or societal skill development.

The ranking of these dimensions sheds light on the preferences and priorities of both backers and players. It is possible that backers perceive tangible attributes such as skill building and content organization as immediate indicators of game quality and potential success. These elements can be readily demonstrated in promotional materials, making them more attractive to potential backers. On the other hand, motivation, being more abstract and subjective, might be harder to convey and measure, leading to its lower ranking. It is essential for game developers to recognize these perceptions and strike a balance in their design, ensuring a comprehensive and engaging game experience that appeals to a broad audience.

For skill building, it is essential for players to acquire and build skills as they progress in the game. This ensures continuous learning and engagement. Well-structured game content helps players navigate and understand the game better, thus enhancing their experience. An engaging storyline provides context and purpose, making gameplay more meaningful. Player interactivity is vital for player engagement. Players should feel that they are part of the game world and can influence it. Immediate feedback helps players understand their progression and areas of improvement. Clear rules ensure that players can easily understand how to play the games, leading to smoother game experiences. For health education games, it is important to have clear learning outcomes that guide the game design. The game must be engaging enough to keep players interested and motivated to continue.

The multivariate analysis identified learning objectives and motivation as the 2 significant predictors of a health education game's crowdfunding success, as detailed in Table 6. This indicates the emphasis users place on clear educational outcomes and the motivation to engage with the game. Users prioritize games that offer clear educational outcomes and that effectively motivate players to engage. The significance of learning objectives suggests that backers might prioritize games that have a clear educational goal, ensuring that players gain tangible knowledge or skills. Motivation, on the other hand, ensures that players remain engaged and committed to the game's objectives. When combined, these dimensions can lead to a game that not only educates but does so in a compelling manner, maximizing player retention and learning outcomes.

Limitations and Future Work

The study has some limitations due to the examination of user perception, which is based on a small number of user responses in a small number of crowdfunding campaigns. The study examined subjective opinions across 8 evaluation dimensions, but the reasons for crowdfunding's effectiveness in health education games require further investigation. In addition, we surveyed participants as potential backers. A more comprehensive approach would involve surveying actual backers, those who make real investments, to discern any differences in perceptions. This could provide a richer understanding of the dynamics at play. The impact of quality on the campaign content and media aspects, as well as user indicators of motivation and interactivity, was investigated in this study. Through crowdfunding, health education games improve engagements, learning components, and cultural adaptability for user engagement [8-10].

Conclusion

Crowdfunding for health education games presents a unique opportunity to bridge the gap between game developers and potential users. There has been little research that has provided empirical evidence for evaluating user perspectives on crowdfunding health education games. Further empirical evaluations are clearly beneficial to providing a rigorous validation of gamification's effectiveness in eHealth. This research conducted an exploratory study and identified 3 major components that matter for health game crowdfunding success. These components are related to game design, instruction, and game content. Interestingly, motivation and assessment and feedback were grouped into game content categories, not into game design categories. This indicates that the proposals for health-related crowdfunding education games are comprehensive, encompassing content that is engaging, interesting, and attractive, with solid assessment and feedback components. Among them, given the nature of health subjects, entrepreneurs and educators should pay more attention to game development factors such as motivation, interactivity, and game rules, so that the health or scientific subjects can be easily infused in the gaming process. Making health games look playful and attractive enables users to easily grasp basic health knowledge during the gaming process [93]. Interestingly, there is little difference in content organization between successful and unsuccessful games, which indicates that even if the game content is easy to follow, it is still not enough. Backers and potential funders or users mostly agree with the health content itself, but they care more about the game development components, using these dimensions to assess the crowdfunding game proposal and determine if these game designs are acceptable and make logical sense.

Our findings recognize the importance of aligning game design with user preferences. The success of health education games on crowdfunding platforms relies on a combination of clear educational objectives, effective player engagement mechanism, and well-structured game content. The study highlights the significance of learning objectives and motivation as key determinants of crowdfunding success for health education games. Game developers aiming for success in this domain should prioritize these dimensions, thus ensuring that their games offer a clear educational outcome.

Acknowledgments

We would like to extend our sincere gratitude to the 3 anonymous reviewers for their invaluable and insightful feedback on our manuscript. Their consistent comments significantly improved our manuscript.

Conflicts of Interest

None declared.

References

1. Bartolome NA, Zorrilla AM, Zapirain BG. Can game-based therapies be trusted? is game-based education effective? a systematic review of the serious games for health and education. Presented at: 2011 16th International Conference on Computer Games (CGAMES); Jul 27-30, 2011; Louisville, KY p. 275-282. [doi: 10.1109/CGAMES.2011.6000353]
2. Baranowski T, Buday R, Thompson D, Lyons EJ, Lu AS, Baranowski J. Developing games for health behavior change: getting started. Games Health J 2013 Aug;2(4):183-190. [doi: 10.1089/g4h.2013.0048] [Medline: 24443708]
3. Kickstarter. URL: https://www.kickstarter.com/ [accessed 2024-02-02]
4. Davis AJ, Parker HM, Gallagher R. Gamified applications for secondary prevention in patients with high cardiovascular disease risk: a systematic review of effectiveness and acceptability. J Clin Nurs 2021 Oct;30(19-20):3001-3010. [doi: 10.1111/jocn.15808] [Medline: 33872436]
5. Amresh A, Salla R, Sinha M, Birr R. Design, implementation and evaluation of a game-based intervention targeting Latino children for improving obesity outcomes. Presented at: 2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH); May 11-13, 2016; Orlando, FL p. 1-8. [doi: 10.1109/SeGAH.2016.7586280]
6. Bassani G, Marinelli N, Vismara S. Crowdfunding in healthcare. J Technol Transf 2019 Aug 1;44(4):1290-1310. [doi: 10.1007/s10961-018-9663-7]
7. Besancenot D, Vranceanu R. A global game model of medical project crowdfunding. International Game Theory Review 2020 Sep;22(03):1950016. [doi: 10.1142/S0219198919500166]
8. Gonzales AL, Kwon EY, Lynch T, Fritz N. Better everyone should know our business than we lose our house”: costs and benefits of medical crowdfunding for support, privacy, and identity. New Media & Society 2018 Feb;20(2):641-658. [doi: 10.1177/1461444816667723]
9. Choy K, Schlagwein D. Crowdsourcing for a better world: on the relation between IT affordances and donor motivations in charitable crowdfunding. Information Technology & People 2016 Mar 7;29(1):221-247. [doi: 10.1108/ITP-09-2014-0215]
10. Kim JG, Hong H, Karahalios K. Understanding identity presentation in medical crowdfunding. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2018:1-12. [doi: 10.1145/3173574.3173708]
11. Kenworthy NJ. Crowdfunding and global health disparities: an exploratory conceptual and empirical analysis. Global Health 2019 Nov 28;15(Suppl 1):71. [doi: 10.1186/s12992-019-0519-1] [Medline: 31775896]
12. Kenworthy N. Like a grinding stone: how crowdfunding platforms create, perpetuate, and value health inequities. Med Anthropol Q 2021 Sep;35(3):327-345. [doi: 10.1111/maq.12639] [Medline: 33711178]
13. Eukel HN, Frenzel JE, Cernusca D. Educational gaming for pharmacy students - design and evaluation of a diabetes-themed escape room. Am J Pharm Educ 2017 Sep;81(7):6265. [doi: 10.5688/ajpe8176265] [Medline: 29109566]
14. Ortiz RA, Witte S, Gouw A, et al. Engaging a community for rare genetic disease: best practices and education from individual crowdfunding campaigns. Interact J Med Res 2018 Feb 5;7(1):e3. [doi: 10.2196/ijmr.7176] [Medline: 29402763]
15. Rose S, Wentzel D, Hopp C, Kaminski J. Launching for success: the effects of psychological distance and mental simulation on funding decisions and crowdfunding performance. J Bus Ventur 2021 Nov;36(6):106021. [doi: 10.1016/j.jbusvent.2020.106021]
16. Lukk M, Schneiderhan E, Soares J. Worthy? crowdfunding the Canadian health care and education sectors. Can Rev Sociol 2018 Aug;55(3):404-424. [doi: 10.1111/cars.12210] [Medline: 29984886]
17. Olszewski AE, Wolbrink TA. Serious gaming in medical education: a proposed structured framework for game development. Simul Healthc 2017 Aug;12(4):240-253. [doi: 10.1097/SIH.0000000000000212] [Medline: 28027076]
18. Ryu S, Park J, Kim K, Kim YG. Reward versus altruistic motivations in reward-based crowdfunding. International Journal of Electronic Commerce 2020 Mar 1;24(2):159-183. [doi: 10.1080/10864415.2020.1715531]
19. Mollick E. The dynamics of crowdfunding: an exploratory study. J Bus Ventur 2014 Jan;29(1):1-16. [doi: 10.1016/j.jbusvent.2013.06.005]
20. Viotto da Cruz J. Beyond financing: crowdfunding as an informational mechanism. J Bus Ventur 2018 May;33(3):371-393. [doi: 10.1016/j.jbusvent.2018.02.001]
21. Antonenko PD, Lee BR, Kleinheksel AJ. Trends in the crowdfunding of educational technology startups. TechTrends 2014 Nov;58(6):36-41. [doi: 10.1007/s11528-014-0801-2]
22. Cha J. Crowdfunding for video games: factors that influence the success of and capital pledged for campaigns. International Journal on Media Management 2017 Jun 30;19(3):240-259. [doi: 10.1080/14241277.2017.1331236]
23. Tyni H. Double duty: crowdfunding and the evolving game production network. Games Cult 2020 Mar;15(2):114-137. [doi: 10.1177/1555412017748108]
24. Smith AN. The backer–developer connection: exploring crowdfunding’s influence on video game production. New Media & Society 2015 Feb;17(2):198-214. [doi: 10.1177/1461444814558910]
25. Yuan X, Wang L, Yin X, Wang H. How text sentiment moderates the impact of motivational cues on crowdfunding campaigns. Financ Innov 2021 Jun 18;7(1):46. [doi: 10.1186/s40854-021-00258-w]
26. Hoegen A, Steininger DM, Veit D. How do investors decide? an interdisciplinary review of decision-making in crowdfunding. Electron Markets 2018 Aug;28(3):339-365. [doi: 10.1007/s12525-017-0269-y]
27. Macht SA, Weatherston J. The benefits of online crowdfunding for fund‐seeking business ventures. Strategic Change 2014 Feb 24;23(1-2):1-14. [doi: 10.1002/jsc.1955]

28. Gerber EM, Hui J. Crowdfunding: motivations and deterrents for participation. ACM Trans Comput-Hum Interact 2013 Dec 1;20(6):1-32. [doi: 10.1145/2530540]
29. Zhou MJ, Lu B, Fan WP, Wang GA. Project description and crowdfunding success: an exploratory study. Inf Syst Front 2018;20(2):259-274. [doi: 10.1007/s10796-016-9723-1] [Medline: 29755287]
30. Wang N, Li Q, Liang H, Ye T, Ge S. Understanding the importance of interaction between creators and backers in crowdfunding success. Electron Commer Res Appl 2018;27:106-117. [doi: 10.1016/j.elerap.2017.12.004]
31. Woods C, Yu H, Huang H. Predicting the success of entrepreneurial campaigns in crowdfunding: a spatio-temporal approach. J Innov Entrep 2020 Jul 1;9(1):13. [doi: 10.1186/s13731-020-00122-8]
32. Henkemans OAB, Bierman BPB, Janssen J, et al. Design and evaluation of a personal robot playing a self-management education game with children with diabetes type 1. Int J Hum Comput Studies 2017 Oct;106:63-76. [doi: 10.1016/j.ijhcs.2017.06.001]
33. Jacobson D, Jacobson J, Leong T, Lourenco S, Mancl L, Chi DL. Evaluating child toothbrushing behavior changes associated with a mobile game app: a single arm pre/post pilot study. Pediatr Dent 2019 Jul 15;41(4):299-303. [Medline: 31439090]
34. Begg M. Leveraging game-informed healthcare education. Med Teach 2008;30(2):155-158. [doi: 10.1080/01421590701874041] [Medline: 18464139]
35. Annetta LA, Lamb R, Stone M. Assessing serious educational games. In: Annetta L, Bronack SC, editors. Serious Educational Game Assessment: SensePublishers; 2011:75-93. [doi: 10.1007/978-94-6091-329-7_5]
36. Orji R, Vassileva J, Mandryk RL. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. User Model User-Adap Inter 2014 Dec;24(5):453-498. [doi: 10.1007/s11257-014-9149-8]
37. Lieberman DA. Management of chronic pediatric diseases with interactive health games: theory and research findings. J Ambul Care Manage 2001 Jan;24(1):26-38. [doi: 10.1097/00004479-200101000-00004] [Medline: 11189794]
38. Rewolinski JA, Kelemen A, Liang Y. Type I diabetes self-management with game-based interventions for pediatric and adolescent patients. Comput Inform Nurs 2020 Jun 25;39(2):78-88. [doi: 10.1097/CIN.0000000000000646] [Medline: 32590404]
39. Dos Santos H, Bredehoft MD, Gonzalez FM, Montgomery S. Exercise video games and exercise self-efficacy in children. Glob Pediatr Health 2016 Apr 26;3:2333794X16644139. [doi: 10.1177/2333794X16644139] [Medline: 27336015]
40. Sung HY, Hwang GJ, Yen YF. Development of a contextual decision-making game for improving students' learning performance in a health education course. Computers & Education 2015 Mar;82:179-190. [doi: 10.1016/j.compedu.2014.11.012]
41. Thompson D, Baranowski T, Buday R, et al. Serious video games for health how behavioral science guided the development of a serious video game. Simul Gaming 2010 Aug 1;41(4):587-606. [doi: 10.1177/1046878108328087] [Medline: 20711522]
42. Amory A, Seagram R. Educational game models: conceptualization and evaluation. South African Journal of Higher Education 2003 Mar 31;17(2):206-217. [doi: 10.4314/sajhe.v17i2.25314]
43. Garris R, Ahlers R, Driskell JE. Games, motivation, and learning: a research and practice model. In: Jentsch F, Curtis M, editors. Simulation in Aviation Training: Routledge; 2017:475-501. [doi: 10.4324/9781315243092]
44. Reeves TC, Harmon SW. Systematic evaluation procedures for instructional hypermedia/multimedia. Presented at: Annual Meeting of the American Educational Research Association; Apr 14, 1993; Atlanta, GA.
45. Gee JP. What video games have to teach us about learning and literacy. Computers in Entertainment 2003 Oct 1;1(1):1-20. [doi: 10.1145/950566.950595]
46. Duffy TM, Jonassen DH. Constructivism and the Technology of Instruction: A Conversation: Routledge; 1992. [doi: 10.4324/9780203461976]
47. Huba ME, Freed JE. Learner-Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning: Allyn & Bacon; 2000.
48. Armstrong MB, Landers RN. An evaluation of gamified training: using narrative to improve reactions and learning. Simulation & Gaming 2017 Aug;48(4):513-538. [doi: 10.1177/1046878117703749]
49. Scheeler MC, Ruhl KL, McAfee JK. Providing performance feedback to teachers: a review. Teach Educ Spec Educ 2004 Oct;27(4):396-407. [doi: 10.1177/088840640402700407]
50. Codding RS, Feinberg AB, Dunn EK, Pace GM. Effects of immediate performance feedback on implementation of behavior support plans. J Appl Behav Anal 2005;38(2):205-219. [doi: 10.1901/jaba.2005.98-04] [Medline: 16033167]
51. Marzano RJ, Pickering D, Pollock JE. Classroom Instruction That Works: Research-Based Strategies for Increasing Student Achievement: ASCD; 2001.
52. Cagiltay NE, Ozcelik E, Ozcelik NS. The effect of competition on learning in games. Computers & Education 2015 Sep;87:35-41. [doi: 10.1016/j.compedu.2015.04.001]
53. Sanchez E. Key criteria for game design: a framework. In: Baldissin N, Bettiol S, Magrin SS, Nonino F, editors. Busines Game-Based Learning in Management Education: MEET Project; 2011.
54. Siang AC, Rao RK. Theories of learning: a computer game perspective. Presented at: Fifth International Symposium on Multimedia Software Engineering, 2003; Dec 10-12, 2003; Taichung, Taiwan p. 239-245. [doi: 10.1109/MMSE.2003.1254447]

55. Csikszentmihalyi M, Nakamura J. The dynamics of intrinsic motivation: a study of adolescents. In: Flow and the Foundations of Positive Psychology: Springer; 2014:175-197. [doi: 10.1007/978-94-017-9088-8_12]
56. Pintrich PR, Smith DAF, Garcia T, Mckeachie WJ. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). Educ Psychol Meas 1993 Sep;53(3):801-813. [doi: 10.1177/0013164493053003024]
57. Eccles J, Wigfield A. Teacher expectations and student motivation. In: Teacher Expectancies: Erlbaum; 1985:185-226.
58. Pintrich PR, Schunk DH. Motivation in Education: Theory, Research, and Applications: Prentice Hall; 2002.
59. Pugh KJ, Bergin DA. Motivational influences on transfer. Educ Psychol 2006 Sep;41(3):147-160. [doi: 10.1207/s15326985ep4103_2]
60. Kovell E, Fisher S, Mann A. Playout: The Exercise Card Game - fitness made fun. Kickstarter. 2020 Dec 7. URL: https://www.kickstarter.com/projects/1209102011/playout-the-game-the-exercise-card-game [accessed 2024-02-02]
61. ACLS MegaCode Simulator for health care professionals. Kickstarter. 2017 Jun 8. URL: https://www.kickstarter.com/projects/2097311518/acls-megacode-simulator-for-health-care-profession [accessed 2024-02-02]
62. Renaissance. Blush. Kickstarter. 2016 Jul 29. URL: https://www.kickstarter.com/projects/renaissancebp/blush [accessed 2024-02-02]
63. Body Cycle Health Education App [© patent pending]. Kickstarter. 2018 Jul 3. URL: https://www.kickstarter.com/projects/renaissancebp/blush [accessed 2024-02-02]
64. Team n z. CHiLD - a psychological 2D RPG. Kickstarter. 2020 Dec 23. URL: https://www.kickstarter.com/projects/teamnoiz/child-a-psychological-2d-rpg [accessed 2024-02-02]
65. 2Dogs Games. Destiny's Sword - an inclusive, uniquely social online game. Kickstarter. 2023 Aug 2. URL: https://www.kickstarter.com/projects/2dogsgames/destinys-sword-an-inclusive-uniquely-social-mmo [accessed 2024-02-02]
66. Facing Dragons: a mixed-reality game to unlock your purpose (canceled). Kickstarter. 2020 May 2. URL: https://www.kickstarter.com/projects/facingdragons/facing-dragons-a-mixed-reality-game-to-unlock-your-purpose [accessed 2024-02-02]
67. Ugalde T. Freesyle Jam Camp - teaches kids (of all ages) how to play frisbee (disc). Kickstarter. 2010 Jan 23. URL: https://www.kickstarter.com/projects/76144658/freesyle-jam-camp-teaches-kids-of-all-ages-how [accessed 2024-02-02]
68. Mobile games to quantify symptoms of mental health disorders. Kickstarter. 2018 Sep 6. URL: https://www.kickstarter.com/projects/eschlicht/gaming-for-good-games-to-diagnose-mental-health-di [accessed 2024-02-02]
69. Ha F. PRESCRIPTION Playing Cards. Kickstarter. 2020 Aug 6. URL: https://www.kickstarter.com/projects/felixha00/prescription-playing-cards [accessed 2024-02-02]
70. Lina. Talk to Me visual novel. Kickstarter. 2022 Dec 7. URL: https://www.kickstarter.com/projects/836732297/talk-to-me-visual-novel [accessed 2024-02-02]
71. TEN. Kickstarter. 2015 Sep 9. URL: https://www.kickstarter.com/projects/472932995/ten [accessed 2024-02-02]
72. Fickles the Clown. The Chakra Collectable Coin. Kickstarter. 2019 Jun 18. URL: https://www.kickstarter.com/projects/fickles/the-lucky-magic-chakra-health-coin [accessed 2024-02-02]
73. The Woosah Kit (canceled). Kickstarter. 2019 Jul 2. URL: https://www.kickstarter.com/projects/30445683/the-woosah-kit [accessed 2024-02-02]
74. Tournesol Kids. The Tournesol Kids Game: activity cards to build resilience. Kickstarter. 2019 Feb 21. URL: https://www.kickstarter.com/projects/tournesolkids/the-tournesol-kids-game-activity-cards-to-build-re [accessed 2024-02-02]
75. Watson R. Youth Run The World 5K. Kickstarter. 2017 Dec 1. URL: https://www.kickstarter.com/projects/334650164/youth-run-the-world-5k [accessed 2024-02-02]
76. Zombied. Kickstarter. 2017 Apr 11. URL: https://www.kickstarter.com/projects/649944397/zombied [accessed 2024-02-02]
77. Lieberman DA, Fisk MC, Biely E. Digital games for young children ages three to six: from research to design. Comput Sch 2009 Nov 30;26(4):299-313. [doi: 10.1080/07380560903360178]
78. Ortiz P, Khin Khin E. Traditional and new media's influence on suicidal behavior and contagion. Behav Sci Law 2018 Mar;36(2):245-256. [doi: 10.1002/bsl.2338] [Medline: 29659071]
79. Hermans RCJ, van den Broek N, Nederkoorn C, Otten R, Ruiter ELM, Johnson-Glenberg MC. Feed the alien! the effects of a nutrition instruction game on children's nutritional knowledge and food intake. Games Health J 2018 Jun;7(3):164-174. [doi: 10.1089/g4h.2017.0055] [Medline: 29634366]
80. Staiano AE. Learning by playing: video gaming in education—a cheat sheet for games for health designers. Games Health J 2014 Oct;3(5):319-321. [doi: 10.1089/g4h.2014.0069] [Medline: 26192487]
81. Lieberman DA. Using digital games to promote health behavior change. In: Sundar S, editor. The Handbook of the Psychology of Communication Technology 2015:507-527. [doi: 10.1002/9781118426456.ch23]
82. Chaudy Y, Connolly T. Specification and evaluation of an assessment engine for educational games: integrating learning analytics and providing an assessment authoring tool. Entertain Comput 2019 May;30:100294. [doi: 10.1016/j.entcom.2019.100294]
83. Radhakrishnan K, Toprac P, O'Hair M, et al. Interactive digital e-health game for heart failure self-management: a feasibility study. Games Health J 2016 Dec;5(6):366-374. [doi: 10.1089/g4h.2016.0038] [Medline: 27976955]
84. Sajjad S, Abdullah AH, Sharif M, Mohsin S. Psychotherapy through video game to target illness related problematic behaviors of children with brain tumor. Curr Med Imaging 2014;10(1):62-72. [doi: 10.2174/1573405610666140313004302]

85. Mueller S, Soriano D, Boscor A, et al. MANTRA: development and localization of a mobile educational health game targeting low literacy players in low and middle income countries. BMC Public Health 2020 Jul 28;20(1):1171. [doi: 10.1186/s12889-020-09246-8] [Medline: 32723317]
86. Akl EA, Sackett K, Pretorius R, et al. Educational games for health professionals. Cochrane Database Syst Rev 2008 Jan 23(1):CD006411. [doi: 10.1002/14651858.CD006411.pub2] [Medline: 18254103]
87. Abraham O, LeMay S, Bittner S, Thakur T, Stafford H, Brown R. Investigating serious games that incorporate medication use for patients: systematic literature review. JMIR Serious Games 2020 Apr 29;8(2):e16096. [doi: 10.2196/16096] [Medline: 32347811]
88. Parisod H, Pakarinen A, Axelin A, Löyttyniemi E, Smed J, Salanterä S. Feasibility of mobile health game "Fume" in supporting tobacco-related health literacy among early adolescents: a three-armed cluster randomized design. Int J Med Inform 2018 May;113:26-37. [doi: 10.1016/j.ijmedinf.2018.02.013] [Medline: 29602430]
89. Wols A, Poppelaars M, Lichtwarck-Aschoff A, Granic I. The role of motivation to change and mindsets in a game promoted for mental health. Entertain Comput 2020 Aug;35:100371. [doi: 10.1016/j.entcom.2020.100371]
90. Poppelaars M, Lichtwarck-Aschoff A, Kleinjan M, Granic I. The impact of explicit mental health messages in video games on players' motivation and affect. Comput Hum Behav 2018 Jun;83:16-23. [doi: 10.1016/j.chb.2018.01.019]
91. Hair J, Black B, Babin B, Anderson R, Tatham R. Multivariate Data Analysis: Prentice-Hall; 2005.
92. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Statist 1947 Mar;18(1):50-60. [doi: 10.1214/aoms/1177730491]
93. Grimes A, Kantroo V, Grinter RE. Let's play! mobile health games for adults. In: UbiComp '10: Proceedings of the 12th ACM International Conference on Ubiquitous Computing: Association for Computing Machinery; 2010:241-250. [doi: 10.1145/1864349.1864370]

Edited by G Eysenbach; submitted 15.05.22; peer-reviewed by E Braad, K Nalin, V Ochs; revised version received 23.10.23; accepted 14.11.23; published 27.02.24.

Please cite as:
Huang H, Yu H, Li W
Assessing the Importance of Content Versus Design for Successful Crowdfunding of Health Education Games: Online Survey Study
JMIR Serious Games 2024;12:e39587
URL: https://games.jmir.org/2024/1/e39587
doi:10.2196/39587

Original Paper

Psychometric Properties of the Metacognitions About Online Gaming Scale in the Chinese Population and Its Relationship With Internet Gaming Disorder: Cross-Sectional Study

Shuhong Lin[1,2*], MM; Xinxin Chen[1,2*], MM; Linxiang Tan[3], MD; Zhenjiang Liao[1,2], MM; Yifan Li[1,2], MM; Ying Tang[1,2], BM; Qiuping Huang[4], MD; Hongxian Shen[1,2], MD

[1]Department of Psychiatry, Second Xiangya Hospital of Central South University, Changsha, China
[2]National Clinical Research Center for Mental Disorders, Changsha, China
[3]Education Center for Mental Health, Central South University, Changsha, China
[4]School of Humanities and Management, Hunan University of Chinese Medicine, Changsha, China
[*]these authors contributed equally

Corresponding Author:
Hongxian Shen, MD
Department of Psychiatry
Second Xiangya Hospital of Central South University
139 Renmin Road
Changsha, 410000
China
Phone: 86 13875970393
Email: shenhx2018@csu.edu.cn

Abstract

Background: Metacognitions about online gaming have been shown to be correlated with Internet Gaming Disorder (IGD). Knowledge of metacognitions about online gaming can help to understand IGD. The Metacognitions about Online Gaming Scale (MOGS) is a reliable and valid tool to measure specific metacognitions about online gaming in both adults and adolescents, which is lacking in China.

Objective: This study was conducted to assess the psychometric properties of the Chinese version of the MOGS (C-MOGS) and its relationship with IGD in the Chinese population.

Methods: A total of 772 Chinese individuals (age: mean 21.70, SD 8.81 years; age range: 13-57 years; 458/772, 59.3% male) completed a web-based questionnaire survey, including the C-MOGS and a battery of validated scales measuring IGD, gaming motives, depression, and anxiety.

Results: Through exploratory and confirmatory factor analyses, the 3-factor structure was confirmed to have adequate model fit and internal consistency reliability (Cronbach α≥.799, Guttman split-half coefficients≥0.754). Concurrent validity of the C-MOGS was supported by its correlations with IGD (*P*<.001), gaming motives (*P*<.001), depression (*P*<.001), and anxiety (*P*<.001). Furthermore, the incremental validity analysis showed that the C-MOGS predicted 13% of the variance in IGD while controlling for gender, age, weekly gaming hours, gaming motives, depression, and anxiety.

Conclusions: This study provides evidence that the psychometric properties of the C-MOGS are appropriate and emphasizes its positive association with IGD. The C-MOGS is a reliable and valid instrument for mental health workers to assess metacognitions about online gaming in the Chinese population.

(JMIR Serious Games 2024;12:e45985) doi:10.2196/45985

KEYWORDS

metacognition; metacognitions about online gaming; Internet Gaming Disorder; psychometric properties; Chinese

Introduction

Metacognition refers to the awareness of one's own thoughts and behaviors, as well as the ability to monitor and alter behavior. It encompasses any cognitive process that receives information from and exerts a controlling influence on another cognitive process [1-4]. More specifically, it comprises metacognitive knowledge and metacognitive regulation. Metacognitive knowledge refers to information and beliefs about one's cognitive processes, while metacognitive regulation pertains to skills to regulate thoughts, including planning, supervision, and regulation [5]. Metacognition contributes to effective decision-making across a variety of contexts [4]. For instance, it facilitates the smooth operation of ongoing thought and behavior by helping us recognize our errors [6], regulate the deployment of executive function [7], and detect lapses of attention [8]. Originating from cognitive psychology, metacognition has been linked to psychological disturbances [9,10].

In recent years, studies have highlighted the potential role of metacognitions in the development of addictive behaviors, such as problematic gaming behavior [11,12]. However, due to the lack of suitable research instruments, conducting further investigations in China has been challenging. To address this issue, this study aimed to evaluate the validity of the Metacognitions about Online Gaming Scale (MOGS) [13] among the Chinese population and its association with gaming behavior.

According to the self-regulatory executive function model, metacognitions play a critical role in the occurrence and development of psychological dysfunction [14]. In this model, psychological dysfunction is activated and perpetuated by a fixed thinking pattern called cognitive attentional syndrome (CAS), which comprises several maladaptive coping strategies (eg, rumination, threat-monitoring, and avoidance). The CAS is driven and maintained by maladaptive metacognitions [15]. Maladaptive metacognitions mistakenly regard the CAS as an effective coping style, resulting in a vicious cycle of ineffective self-regulation [16]. Over the last 40 years, metacognitions have been associated with several mental and psychological problems [17], such as obsessive-compulsive disorder, schizophrenia, addiction, anxiety, and depression [18-20].

In the domain of addictive behaviors, metacognitions are divided into 2 subtypes: positive and negative [21]. The former refers to the beliefs that engaging in specific addictive behaviors is a strategy of affective and cognitive self-regulation, such as "Drinking helps me think more clearly" and "Gambling can improve my mood" [22,23]. The latter refers to the concerns about the uncontrollability and danger of thoughts or engagement with addictive behaviors. For example, "Drinking will interfere with my thought" and "Once I start thinking about drinking, I cannot stop" [24]. Previous studies have shown that positive metacognitions can motivate addictive behaviors in the early stage, while negative metacognitions contribute to their perpetuation by activating negative emotional states as a reinforcement [11,21]. In recent years, metacognition has been correlated with many addictive behaviors, such as problematic alcohol use [25-27], nicotine dependence [28,29], gambling disorder [30-32], problematic Internet use [33-35], problematic social media use [36-38], and Internet Gaming Disorder (IGD; problematic online gaming) [39,40].

As an addictive behavior, IGD was first included in the research appendix section of the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) in 2013 [41], then it was officially included in the addiction disease unit of the *Eleventh Revision of the International Classification of Diseases (ICD-11)* in 2018 [42]. Its core characteristics include losing control while gaming, prioritizing gaming over other interests, and causing functional damage in daily life. Excessive online gaming results in various problems, such as sacrificing real-life relationships, sleep, work, and education, leading to brain damage [43-46]. According to a recent review, the global prevalence of IGD was 3.05%, and it was higher among Asians (5.08%) than Europeans (2.72%) [47]. In China, the prevalence ranges from 3.5% to 17%, which is higher than the global average level [48-50].

In order to effectively prevent and treat this disorder, extensive research has been conducted to investigate its etiology. These studies have revealed a significant association between IGD and various psychological factors, including negative affect, gaming motives, and maladaptive cognition [51]. The Interaction of Person-Affect-Cognition-Execution (I-PACE) model proposed by Brand et al [52,53] suggests that the initiation of addictive behaviors arise from the integration of emotional and cognitive responses to internal or external stimuli along with specific motivations. Motives are sets of knowledge that represent the emotional preferences expressed in our thoughts and concepts. Gaming motives could be considered as stimulating factors of gaming behavior, which may play an important role in the development of IGD [51]. Furthermore, Spada et al [21] posited that the development and persistence of addictive behaviors, including IGD, are strongly influenced by particular metacognitions about addictive behaviors.

According to previous studies, metacognitions have been associated with IGD [38,39]. However, these studies mainly focused on generic metacognitions (eg, beliefs about worry, cognitive monitoring, the need for thought suppression). To assess specific metacognitions about online gaming, Spada and Caselli [13] developed a self-rating instrument called the Metacognitions about Online Gaming Scale (MOGS). In the original validation of the MOGS, an exploratory factor analysis (EFA) was performed with 225 adults in Study 1 that suggested a 2-factor solution: Negative Metacognitions about Online Gaming (N-MOG; 6 items) and Positive Metacognitions about Online Gaming (P-MOG; 6 items) [13]. The N-MOG assesses negative metacognitions about the uncontrollability and danger of thoughts on gaming. The P-MOG measures positive metacognitions in which online gaming helps individuals regulate affect and thought. In Study 2, the confirmatory factor analysis (CFA) with another sample of 348 individuals further divided N-MOG into 2 factors and built a 3-factor structure: Negative Metacognitions about the Uncontrollability of Online Gaming (N-MOGU), Negative Metacognitions about the Dangers of Online Gaming (N-MOGD), and P-MOG [13]. All 3 factors reported adequate internal reliability (Cronbach $\alpha \geq .79$).

While exploring predictive validity, the study showed that MOGS was positively related to gaming hours and Internet addiction [13]. Overall, these findings demonstrated the reliability and validity of the MOGS.

To extend the utility of the MOGS to adolescent populations from other countries, Akbari et al [54] translated it into Persian and evaluated its psychometric properties among 769 Iranian adolescents. The results showed that the 3-factor structure had appropriate construct validity and internal consistency (Cronbach α≥.79). Furthermore, metacognitions about online gaming were able to independently predict problematic gaming behavior while controlling for personality traits, gaming motives, gaming-related cognitions, and negative affect [54]. Another study investigating the association between IGD and social anxiety reported that metacognitions about online gaming were significantly correlated with IGD and mediated the latter's relationship with social anxiety [55].

These studies indicated an association between specific metacognitions about online gaming and IGD. Further exploration could be beneficial for the treatment and prevention of IGD, especially in countries with a higher prevalence, such as China. However, it is difficult to conduct relevant research in China because of the lack of instruments used to evaluate specific metacognitions. Therefore, the primary objective of this study was to translate the MOGS into Chinese and validate its psychometric properties among Chinese adolescent and adult gamers using online convenience sampling. Additionally, the study aimed to investigate the unique influence of metacognitions about online gaming on IGD while considering variables such as anxiety, depression, and motivation. The hypothesis was that, within the Chinese population, positive and negative metacognitions about online gaming would serve as independent risk factors for IGD, distinct from other contributing factors.

Methods

Participants

We recruited all individuals online through convenience sampling in June 2021. The inclusion criteria were as follows: (1) age ≥13 years, (2) Chinese speakers who could understand the questionnaires, (3) consent to participate (adolescents with parental consent), and (4) played games at least one hour every week (excluding online gambling) in the last 12 months.

In total, 996 individuals participated in this survey. We excluded 88 individuals whose answer was "No" to the item "Are your answers to this questionnaire true and reliable?", 37 who gave the same answers to more than 50% of the questions and whose time spent on the questionnaire was less than the mean minus 3 SD, and 99 who were younger than 13 years old. The final sample included 772 participants.

Ethical Considerations

Before starting the anonymous online investigation, participants were informed about the purpose and rights of the study and signed an online informed consent form. Those younger than 18 years needed to inform their guardians and obtain consent before filling out the questionnaire. The ethics committee of the Second Xiangya Hospital of Central South University approved this study (protocol code 2020004; dated March 1, 2020).

Measures

Basic Information

Basic information included sociodemographic information and Internet gaming characteristics. The former included gender, age, employment, years of education, and family structure (eg, single-child family). For the latter, participants reported their average time spent gaming (weekly gaming hours), gaming devices (a multiple-choice question), the number of long-term gaming partners, and self-evaluation of gaming addiction.

Metacognitions About Online Gaming

Metacognitions about online gaming were measured using the MOGS [13], which contains 12 items rated on a 4-point Likert scale (1=Do not agree to 4=Agree very much). The MOGS comprises the following 3 factors: (1) N-MOGU (3 items, such as "Once I start online gaming I cannot stop"), (2) N-MOGD (3 items, such as "Online gaming makes me lose control"), and (3) P-MOG (6 items, such as "Online gaming stops me from worrying"). A higher score indicates a higher degree of specific metacognition about online gaming.

IGD Symptoms

The severity of IGD symptoms was assessed using the Internet Gaming Disorder Scale-Short Form (IGDS9-SF) [56,57]. The IGDS9-SF is a 9-item scale developed from the core symptoms of IGD proposed by the DSM-5 and assesses gaming activities and their adverse effects in the past 12 months. All items are rated on a 5-point Likert scale (1=never to 5=very often). The scores range from 9 to 45. Higher scores represent more severe IGD symptoms. With adequate reliability (Cronbach α≥.9), the Chinese version of the IGDS9-SF was used in our research [58,59]. The Cronbach α was .90 in this study.

Gaming Motives

We assessed gaming motives using the Motives for Online Gaming Questionnaire (MOGQ) [60]. It includes 27 items comprising the following 7 motivational dimensions (all rated on a 4-point Likert scale): escape, skill development, recreation, competition, coping, fantasy, and social. The Chinese version of the MOGQ has excellent reliability (Cronbach α≥.83) and validity [61]. Higher scores reflect stronger motives for online gaming. In this study, the Cronbach α was .95 for the total scale and ranged from 0.84 to 0.89 for each subscale.

Depression and Anxiety

The Patient Health Questionnaire-9 (PHQ-9) [62] was used to measure depressive symptoms. It is a diagnostic screening tool that monitors the severity of depression over the last 2 weeks. All items are scored on a 4-point Likert scale (0=Not at all to 3=Nearly every day). The scores range from 0 to 27. Higher scores denote worse depressive symptoms. The Chinese version of the PHQ-9 [63] has suitable reliability (Cronbach α=.85). The Cronbach α was .89 for this study.

The Generalized Anxiety Disorder-7 (GAD-7) [64] was used to measure anxiety symptoms. It is a self-rated scale that assesses the severity of anxiety symptoms over the last 2 weeks. All items are scored on a 4-point Likert scale (0=Not at all to 3=Nearly every day). The scores range from 0 to 21. Higher scores represent worse anxiety symptoms. The Chinese version of the GAD-7 [65] was used, with appropriate internal consistency (Cronbach α=.90) and validity. The Cronbach α was .92 for this study.

Procedures

The MOGS was translated into Chinese by 2 professional translators using a standard translation and back-translation method [66]. For some controversial items (eg, "Online gaming makes me lose control," "Online gaming makes my worries more bearable"), we consulted the author of the original scale. Considering the original scale, 2 bilingual psychologists revised the translated version and checked its face validity. A pilot study was conducted with 5 adults and 5 adolescents to test the understandability. Based on their feedback, some descriptions of items were modified, and the final Chinese version of the MOGS (C-MOGS) was created.

We conducted this survey online using Questionnaire Star, a professional online survey platform. By reading recruitment advertisements posted on social networking sites (eg, WeChat, Weibo, and other webcast platforms), individuals could open the questionnaire link. On the first page, participants could read the objectives and content of this research and confirm their participation (minors, those younger than 18 years, had to obtain the consent of their guardian). Each IP address can only be used once to avoid repeated participation. After the questionnaire was submitted, all the data were sent to the researcher's account, and only the researcher could view the data.

Data Analyses

Data analyses were conducted using SPSS version 25.0 (IBM Corp) and AMOS version 24.0 (IBM Corp). First, basic statistical analyses (eg, descriptive analysis, independent samples *t* tests, chi-square tests) were performed on sociodemographic variables and Internet gaming characteristics. To analyze the construct of the C-MOGS, the total sample was randomly split into 2 subsamples. Sample-1 (n1=390; 229/390, 58.7% male; age: mean 22.25, SD 9.05 years) was used for EFA, and sample-2 (n2=382; 229/382, 59.9% male; age: mean 21.14, SD 8.52 years) was used for CFA. Except for EFA and CFA, all analyses were conducted on data from the entire sample. Independent samples *t* tests and chi-square tests showed that there were no significant differences between the 2 subsamples regarding age (t_{770}=–1.754, *P*=.08), gender (χ^2_1=0.121, *P*=.73), average weekly gaming time (t_{770}=0.775, *P*=.44), gaming devices (χ^2_1<1.012, *P*=.31), long-term gaming partner (χ^2_3=4.046, *P*=.26), and other demographic variables (*P*=.33-.88). An EFA with principal component analysis (PCA) and varimax-rotation method was conducted on the C-MOGS items. To validate the models derived from the EFA, a CFA was completed with AMOS 24.0 using the maximum likelihood method. The model fit was appraised using multiple fit indexes, including chi-squared:degree of freedom ratio (χ^2/df<5), goodness-of-fit index (GFI>0.90), Tucker-Lewis Index (TLI>0.90), comparative fit index (CFI>0.90), standardized root of the mean square residual (SRMR<0.08), root mean square error of approximation (<0.05=close fit; <0.08=acceptable fit; <0.1=mediocre fit) [67]. The reliability of the C-MOGS was examined by assessing the internal consistency of the scale and subscale. Acceptable values for the Cronbach α and Guttman split-half coefficients are >0.70, while values >0.80 are considered good [68]. Finally, to test the concurrent and incremental validity, correlation analyses and hierarchical multiple regression analyses were conducted between the C-MOGS and Internet gaming characteristics (gaming time, IGD, gaming motives) as well as anxiety and depression.

Results

Sample Characteristics

In this study, analysis was conducted on data from 772 participants (458 men, 59.3%) aged between 13 years and 57 years (age: mean 21.70, SD 8.81 years; participants aged 13-17 years: 281/772, 36.4%). The majority of them were students (555/772, 71.9%). Smartphones were the most popular device for gaming (705/772, 91.3%). Of the sample, 69.4% (536/772) had one or more long-term gaming partners. Participants spent an average of 13.43 (SD 10.88) hours every week playing games. More details on the sample characteristics are shown in Table 1.

Table 1. Sociodemographic and Internet gaming characteristics of the sample (n=772).

Characteristic	Participants' results
Gender, n (%)	
Male	458 (59.3)
Female	314 (40.7)
Age (years), mean (SD)	21.70 (8.81)
Employment , n (%)	
Student	555 (71.9)
Full-time employee	186 (24.1)
Part-time employee	13 (1.7)
Unemployed	18 (2.3)
Length of education (years), n (%)	
≤12	339 (43.9)
>12	433 (56.1)
Single child, n (%)	
Yes	289 (37.4)
No	483 (62.6)
Gaming devices, n (%)	
Smartphone	705 (91.3)
Computer	278 (36.1)
Tablet	125 (16.2)
Game console	51 (6.6)
Long-term gaming partners, n (%)	
None	236 (30.6)
≥1 and <3	198 (25.6)
≥3 and <6	159 (20.6)
≥6	179 (23.2)
Self-evaluation of gaming addiction, n (%)	
Yes	129 (16.7)
No idea	203 (26.3)
No	440 (57.0)
Weekly gaming time (hours), mean (SD)	13.43 (10.88)

Factorial Structure of the C-MOGS

EFA

To identify the potential factorial structure of the C-MOGS, an EFA was performed on data from sample-1 (n=390). First, we used the Kaiser-Meyer-Olkin (KMO) and Bartlett tests of sphericity to ensure that the sample was suitable for EFA. The KMO value was 0.894, and the Bartlett test of sphericity was significant (χ^2_{66}=3159.742, *P*<.001), confirming the data were sufficient.

The initial analysis extracted 2 factors using the criteria of an eigenvalue>1 and factor loading>0.40. The 2-factor solution (eigenvalues of 5.573 and 2.597) accounted for 68.08% of the total variance, and the loading of all the items was >0.4 (0.646-0.918; Table 2). Factor 1 included items 1 through 6, referred to as the N-MOG; factor 2 included items 7 through 12, which described the P-MOG.

Additionally, according to the dimension of the original scale [13], we also conducted a PCA by setting 3 factors to be extracted. The 3-factor solution (eigenvalues of 5.573, 2.597, and 0.797) explained 74.73% of the total variance (37.44%, 19.31%, and 17.98%, respectively). Item-factor loadings are presented in Table 2. The factors were as follows: factor 1 (items 1, 2, and 3) referred to the N-MOGU; factor 2 (items 4, 5, and 6) was related to the N-MOGD; and factor 3 (items 7, 8, 9, 10, 11, and 12) was related to the P-MOG [54].

Table 2. Item-factor loadings of the Chinese version of the Metacognitions about Online Gaming Scale (C-MOGS) based on exploratory factor analyses (sample-1, n=390).

Items	2-factor model			3-factor model			
	F1[a,b]	F2[b,c]	Communality	F1[b,d]	F2[b,e]	F3[b,c]	Communality
(1) I continue to play despite I think it would be better to stop	0.646	0.215	0.463	0.809	0.110	0.178	0.698
(2) I have no control over how much time I play	0.830	0.177	0.721	0.808	0.370	0.151	0.813
(3) Once I start online gaming, I cannot stop	0.808	0.178	0.684	0.743	0.403	0.157	0.739
(4) Online gaming makes me lose control	0.746	0.108	0.569	0.213	0.840	0.134	0.768
(5) Thoughts about online gaming interfere with my functioning	0.743	0.007	0.552	0.222	0.825	0.032	0.731
(6) Thoughts about online gaming are becoming an obsession	0.776	0.165	0.630	0.474	0.624	0.169	0.643
(7) Online gaming makes my worries more bearable	0.275	0.734	0.614	0.190	0.204	0.735	0.618
(8) Online gaming reduces my negative feelings	0.131	0.882	0.796	0.087	0.104	0.885	0.801
(9) Online gaming helps me to control my negative thoughts	0.093	0.918	0.852	0.125	0.014	0.915	0.852
(10) Online gaming stops me from worrying	0.116	0.849	0.734	0.083	0.087	0.850	0.737
(11) Online gaming reduces my anxious feelings	0.107	0.917	0.853	0.081	0.076	0.919	0.856
(12) Online gaming distracts my mind from problems	0.222	0.808	0.703	0.256	0.064	0.800	0.710

[a]Negative Metacognitions about Online Gaming (N-MOG).

[b]Factor loadings present the factor matrix values.

[c]Positive Metacognitions about Online Gaming (P-MOG).

[d]Negative Metacognitions about the Uncontrollability of Online Gaming (N-MOGU).

[e]Negative Metacognitions about the Dangers of Online Gaming (N-MOGD).

Confirmatory Factor Analysis

To further evaluate the structural validity of the C-MOGS, we conducted a CFA on sample-2 (n=382) using AMOS 25.0. We compared the goodness of model fit between the 2 aforementioned models. We first tested the 2-factor model, which had a substandard fit in some indexes: χ^2/df=3.962 and root mean squared error of approximation (RMSEA)=0.880. In comparison, the 3-factor model showed an adequate model fit: χ^2/df=3.477, GFI=0.929, CFI=0.958, TLI=0.945, SRMR=0.065, RMSEA=0.081 (Table 3). The correlations between P-MOG, N-MOGU, and N-MOGD were moderate (r=0.389 and 0.377, respectively) and were relatively strong between N-MOGU and N-MOGD (r=0.905).

Due to the high correlation between the 2 negative metacognitive factors, we also created a bifactor model (Figure 1), in which N-MOGU and N-MOGD loaded on a second-order factor (N-MOG) and P-MOG was a first-order factor. In this model, the goodness of model fit was the same as that of the 3-factor model (Table 3), and the correlation between N-MOG and P-MOG was moderate (r=0.396).

Table 3. Model fit indices of the confirmatory factor analyses for the Chinese version of the Metacognitions about Online Gaming Scale (C-MOGS; Sample 2, n=382).

Model	χ^2 (df)	χ^2/df	GFI[a]	CFI[b]	TLI[c]	SRMR[d]	RMSEA[e]
2-factor model	205.517 (52)	3.962	0.917	0.948	0.935	0.067	0.880
3-factor model	173.867 (50)	3.477	0.929	0.958	0.945	0.065	0.810
Bifactor model	173.867 (50)	3.477	0.929	0.958	0.945	0.065	0.810

[a]GFI: goodness-of-fit index.

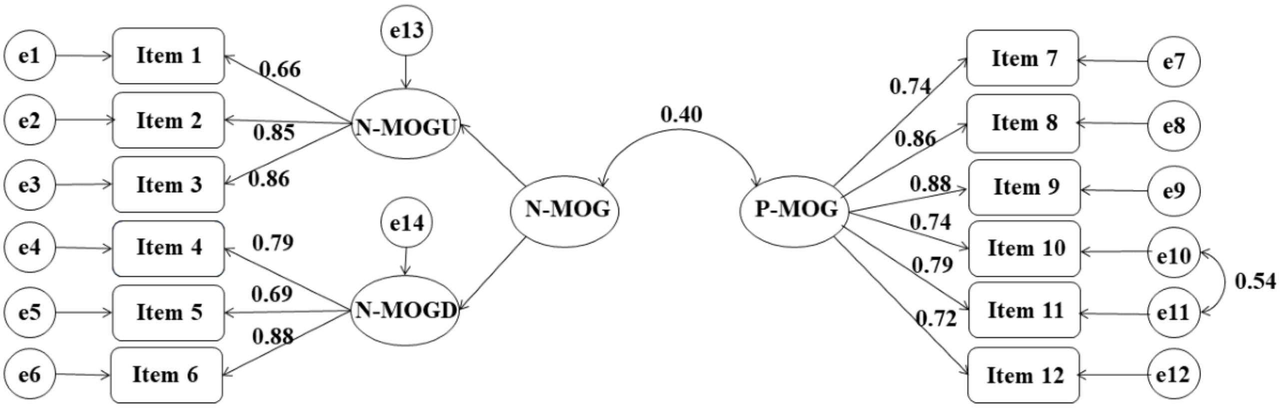
[b]CFI: comparative fit index.

[c]TLI: Tucker-Lewis Index.

[d]SRMR: standardized root of the mean square residual.

[e]RMSEA: root mean square error of approximation.

Figure 1. The bifactor model of the Chinese version of the Metacognitions about Online Gaming Scale (C-MOGS), showing the latent factors as ovals, the 12 items in the C-MOGS as rectangles, the error terms (e1-e14) as circles, and the standardized factor loading above the arrows. N-MOG: Negative Metacognitions about Online Gaming; N-MOGD: Negative Metacognitions about the Dangers of Online Gaming; N-MOGU: Negative Metacognitions about the Uncontrollability of Online Gaming; P-MOG: Positive Metacognitions about Online Gaming.



Reliability

The Cronbach α coefficient and split-half reliability were calculated for the scale and its subscales in the total sample (n=772). The α coefficient for the total scale was .894, and it was .823 for the N-MOGU, .799 for the N-MOGD, and .925 for the P-MOG. No item deletion improved the internal consistency. The Guttman split-half coefficient of the overall scale was 0.942, and for each dimension, the coefficients were 0.776, 0.754, and 0.841. These findings confirmed that the C-MOGS and its subscales exhibit adequate internal consistency.

Moreover, we calculated the correlation coefficient between each item and its relative factor scores. The results showed that the item-total correlations for all items were high (r≥0.551).

Concurrent Validity

We further analyzed the correlation between the 3 factors of the C-MOGS and IGD, gaming motives, anxiety, and depression to test the concurrent validity. Based on the Shapiro-Wilk test, these variables did not follow a normal distribution (all *Ps*<.05). Therefore, Spearman correlation analysis was chosen to explore the relationships between the variables. Table 4 shows the descriptive statistics (median and IQR), and Table 5 shows the correlations between the variables. Each factor of the C-MOGS showed positive correlations with the IGDS9-SF, weekly gaming hours, every dimension of the MOGQ, the PHQ-9, and the GAD-7 (r=0.153 to 0.759, all *Ps*<.01). Moreover, the correlation matrix showed positive correlations between the IGDS9-SF and the other variables (r=0.352 to 0.700, all *Ps*<.01).

Table 4. Descriptive statistics for the variables (n=772).

Variables	Median (IQR)	Range
IGDS9-SF[a]	17 (10)	9-45
WGH[b]	9 (11)	1-69
Motives for Online Gaming Questionnaire (MOGQ)		
Social	7 (6)	4-20
Escape	7 (5)	4-20
Competition	7 (6)	4-20
Coping	9 (6)	4-20
Skill	7 (6)	4-20
Fantasy	6 (5)	4-20
Recreation	9 (6)	3-15
PHQ-9[c]	6 (7)	0-27
GAD-7[d]	4 (6)	0-21
N-MOGU[e]	4 (3)	3-12
N-MOGD[f]	4 (2)	3-12
P-MOG[g]	11 (7)	6-24

[a]IGDS9-SF: Internet Gaming Disorder Scale-Short Form.
[b]WGH: weekly gaming hours (average time).
[c]PHQ-9: Patient Health Questionnaire-9.
[d]GAD-7: Generalized Anxiety Disorder-7.
[e]N-MOGU: Metacognitions about the Uncontrollability of Online Gaming.
[f]N-MOGD: Negative Metacognitions about the Dangers of Online Gaming.
[g]P-MOG: Positive Metacognitions about Online Gaming.

Table 5. Spearman correlation coefficients among the variables (n=772).

Variables	IGDS9-SF[a]	WGH[b]	MOGQ[c]							PHQ-9[d]	GAD-7[e]	N-MOGU[f]	N-MOGD[g]	P-MOG[h]
			Social	Escape	Competition	Coping	Skill	Fantasy	Recreation					
IGDS9-SF														
Correlation	1	0.501	0.352	0.589	0.494	0.542	0.380	0.488	0.416	0.466	0.420	0.700	0.587	0.511
P value	—[i]	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
WGH														
Correlation	0.501	1	0.196	0.326	0.288	0.353	0.190	0.290	0.386	0.253	0.193	0.364	0.280	0.312
P value	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
MOGQ: Social														
Correlation	0.352	0.196	1	0.473	0.521	0.544	0.606	0.523	0.378	0.108	0.118	0.246	0.174	0.398
P value	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	.003	.001	<.001	<.001	<.001
MOGQ: Escape														
Correlation	0.589	0.326	0.473	1	0.503	0.786	0.563	0.658	0.458	0.412	0.412	0.435	0.352	0.674
P value	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
MOGQ: Competition														
Correlation	0.494	0.288	0.521	0.503	1	0.544	0.615	0.552	0.467	0.188	0.196	0.372	0.316	0.412
P value	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
MOGQ: Coping														
Correlation	0.542	0.353	0.544	0.786	0.544	1	0.684	0.636	0.617	0.284	0.297	0.396	0.245	0.759
P value	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
MOGQ: Skill														
Correlation	0.380	0.190	0.606	0.563	0.615	0.684	1	0.589	0.403	0.121	0.151	0.249	0.153	0.531
P value	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	.001	<.001	<.001	<.001	<.001
MOGQ: Fantasy														
Correlation	0.488	0.290	0.523	0.658	0.552	0.636	0.589	1	0.442	0.289	0.295	0.370	0.322	0.476
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001
MOGQ: Recreation														
Correlation	0.416	0.386	0.378	0.458	0.467	0.617	0.403	0.442	1	0.234	0.226	0.348	0.157	0.430
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001
PHQ-9														
Correlation	0.466	0.253	0.108	0.412	0.188	0.284	0.121	0.289	0.234	1	0.789	0.411	0.411	0.270
P value	<.001	<.001	.003	<.001	<.001	<.001	.001	<.001	<.001	—	<.001	<.001	<.001	<.001
GAD-7														
Correlation	0.420	0.193	0.118	0.412	0.196	0.297	0.151	0.295	0.226	0.789	1	0.383	0.376	0.280

Variables	IGDS9-SF[a]	WGH[b]	MOGQ[c]							PHQ-9[d]	GAD-7[e]	N-MOGU[f]	N-MOGD[g]	P-MOG[h]
			Social	Escape	Competition	Coping	Skill	Fantasy	Recreation					
P value	<.001	<.001	.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001
N-MOGU														
Correlation	0.700	0.364	0.246	0.435	0.372	0.396	0.249	0.370	0.348	0.411	0.383	1	0.530	0.359
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001
N-MOGD														
Correlation	0.587	0.280	0.174	0.352	0.316	0.245	0.153	0.322	0.157	0.411	0.376	0.530	1	0.257
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001
P-MOG														
Correlation	0.511	0.312	0.398	0.674	0.412	0.759	0.531	0.476	0.430	0.270	0.280	0.359	0.257	1
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—

[a]IGDS9-SF: Internet Gaming Disorder Scale-Short Form.
[b]WGH: weekly gaming hours (average time).
[c]Motives for Online Gaming Questionnaire.
[d]PHQ-9: Patient Health Questionnaire-9.
[e]GAD-7: Generalized Anxiety Disorder-7.
[f]N-MOGU: Metacognitions about the Uncontrollability of Online Gaming.
[g]N-MOGD: Negative Metacognitions about the Dangers of Online Gaming.
[h]P-MOG: Positive Metacognitions about Online Gaming.
[i]Not applicable.

Incremental Validity

We conducted a hierarchical multiple linear regression analysis to identify the incremental effect of metacognitions about online gaming on IGD. The IGDS9-SF was the outcome variable, and the 3 factors of the C-MOGS were predictor variables, along with other variables related to the IGDS9-SF (gender, weekly gaming hours, the 7 factors of the MOGQ, and the total PHQ-9 and GAD-7 scores). Each variable was input in the following order: step 1: age and gender (0=female, 1=male); step 2: weekly gaming hours; step 3: the 7 factors of the MOGQ; step 4: GAD-7, PHQ-9; step 5: the 3 dimensions of the C-MOGS.

The Durbin-Watson statistic showed that the observed values were independent of each other (D-W=2.077). All tolerance values were above 0.1 (0.180-0.878), indicating no multicollinearity. The results are presented in Table 6. The 3 factors of the C-MOGS accounted for 13.0% of the variance in the IGDS9-SF (*P*<.001). In step 5, the final model indicated that gender, weekly gaming hours, the PHQ-9 score, the MOGQ-Escape score, the MOGQ-Competition score, and the factors of the C-MOGS were significant positive predictors of the IGDS9-SF (R^2=0.729, *P*<.001, adjusted R^2=0.724), and the most important predictor was the N-MOGU (β=0.326, *P*<.001).

Table 6. Hierarchical multiple regression analyses with the Internet Gaming Disorder Scale-Short Form (IGDS9-SF) as the outcome variable and the Chinese version of the Metacognitions about Online Gaming Scale (C-MOGS) factors as predictor variables, together with gender, weekly gaming hours, motives related to online gaming, depression, and anxiety (n=772).

Variable	Step 1[a]			Step 2[b]			Step 3[c]			Step 4[d]			Step 5[e]		
	β	T	*P* value	β	T	*P* value	β	T	*P* value	β	T	*P* value	β	T	*P* value
Age	–0.056	–1.551	.12	–0.110	–3.633	<.001	–0.068	–2.640	.008	–0.050	–2.044	.04	0.004	0.206	.84
Gender	0.204	5.694	<.001	0.110	3.603	<.001	0.084	3.170	.002	0.114	4.588	<.001	0.076	3.712	<.001
WGH[f]	—[g]	—	—	0.540	17.971	<.001	0.315	11.137	<.001	0.277	10.309	<.001	0.155	6.705	<.001
Motives for Online Gaming Questionnaire															
Social	—	—	—	—	—	—	–0.019	–0.567	.57	0.008	0.241	.81	0.010	0.387	.70
Escape	—	—	—	—	—	—	0.381	8.805	<.001	0.265	6.328	<.001	0.132	3.743	<.001
Competition	—	—	—	—	—	—	0.163	4.740	<.001	0.149	4.632	<.001	0.066	2.458	.01
Coping	—	—	—	—	—	—	0.022	0.425	.67	0.016	0.333	.74	–0.022	–0.493	.62
Skill	—	—	—	—	—	—	–0.106	–2.607	.009	–0.058	–1.529	.13	0.024	0.761	.45
Fantasy	—	—	—	—	—	—	0.098	2.601	.009	0.068	1.916	.056	0.020	0.688	.49
Recreation	—	—	—	—	—	—	0.017	0.515	.61	0.009	0.299	.77	0.016	0.618	.54
PHQ-9[h]	—	—	—	—	—	—	—	—	—	0.224	5.553	<.001	0.104	3.091	.002
GAD-7[i]	—	—	—	—	—	—	—	—	—	0.073	1.847	.07	0.025	0.784	.43
N-MOGU[j]	—	—	—	—	—	—	—	—	—	—	—	—	0.326	11.286	<.001
N-MOGD[k]	—	—	—	—	—	—	—	—	—	—	—	—	0.206	7.248	<.001
P-MOG[l]	—	—	—	—	—	—	—	—	—	—	—	—	0.089	2.969	.003

[a]R^2=0.049; adjusted R^2=0.047; ΔR^2=0.049; *P*<.001.
[b]R^2=0.331; adjusted R^2=0.328; ΔR^2=0.281; *P*<.001.
[c]R^2=0.539; adjusted R^2=0.533; ΔR^2=0.208; *P*<.001.
[d]R^2=0.599; adjusted R^2=0.593; ΔR^2=0.061; *P*<.001.
[e]R^2=0.729; adjusted R^2=0.724; ΔR^2=0.130; *P*<.001.
[f]WGH: weekly gaming hours (average time).
[g]Not applicable.
[h]PHQ-9: Patient Health Questionnaire-9.
[i]GAD-7: Generalized Anxiety Disorder-7.
[j]N-MOGU: Metacognitions about the Uncontrollability of Online Gaming.
[k]N-MOGD: Negative Metacognitions about the Dangers of Online Gaming.
[l]P-MOG: Positive Metacognitions about Online Gaming.

Discussion

Principal Findings

To investigate the psychometric properties of the Chinese MOGS and its association with IGD, this study translated and tested it in China for the first time. In general, the results suggested that the C-MOGS could potentially serve as a valid and reliable tool to assess specific metacognitions about online gaming and it may have the capacity to predict IGD independently.

First, factor analyses were used to explore the structural validity of the scale. The EFA suggested a 2-factor solution (N-MOG and P-MOG), which was consistent with the first assumption of the original scale [13]. By setting 3 factors to be extracted, the EFA also obtained the same 3-factor solution as the final version of the original scale (N-MOGU, N-MOGD, and P-MOG) [13,54]. Through CFA, the 3-factor model was later proved to have the best data fit. Moreover, we attempted to build a bifactor model that included a first-order factor (P-MOG) and a second-order factor (N-MOG: N-MOGU and N-MOGD). This model had the same goodness of model fit as the 3-factor structure. To maintain consistency with the original scale, the 3-factor structure is recommended for measuring specific online gaming metacognitions in the Chinese population. For studies that compare N-MOG and P-MOG, the bifactor model can be considered.

The 3-factor structure of the C-MOGS demonstrated adequate internal consistency, with Cronbach α coefficients ranging from .799 to .925 for each factor and the full scale, along with Guttman split-half coefficients ranging from 0.754 to 0.942. The current findings also provide evidence for the concurrent and incremental validity of the C-MOGS. Each subscale was significantly positively correlated with IGD, weekly gaming hours, gaming motives, depression, and anxiety. Moreover, the C-MOGS accounted for 13.0% of the variance in IGD while controlling for other variables. These findings highlight the utility of the C-MOGS as a reliable and valid tool to assess metacognitions about online gaming among the Chinese population.

Furthermore, this study explored the effects of metacognitions about online gaming, gaming motives, anxiety, and depression on IGD using hierarchical multiple linear regression analysis. After adding metacognitions about online gaming to the regression equations, the final model accounted for 72.9% of the variance in IGD. In addition to metacognitions, gender, weekly gaming hours, escapism motives, competition motives, and depression significantly predicted IGD, suggesting that these factors collectively contribute to the development and maintenance of IGD [69-71]. Importantly, the inclusion of metacognitions led to a reduction in the standardized regression coefficients of these variables, and the predictive effect of anxiety on IGD became nonsignificant. This indicates that metacognition may partially mediate or explain the impact of these factors on IGD. This finding is consistent with previous research, suggesting that metacognitions about online gaming may mediate the influence of other psychological factors, such as psychological dependence, anxiety, and depression, on IGD [55,72-75]. These results indicate that specific metacognitions about online gaming are important predictors of IGD, which is consistent with previous studies [13,54,55]. However, the mechanisms underlying the role of metacognitions in IGD seem to be interrelated with other psychological factors, which remains inconclusive.

In a hypothesized model, metacognitions about online gaming may promote problematic gaming engagement by increasing gaming time and disrupting normal emotion and cognition [76]. Consistent with this view, our study found that people with more metacognitions about online gaming would spend more time playing games and feel more anxious and depressed. P-MOG increases gaming time by promoting online gaming as a self-regulation method for emotion and cognition [13,21]. N-MOGU will maintain problematic gaming engagement by destroying one's confidence in self-control, while N-MOGD can induce negative reinforcement and compulsive gaming engagement by triggering negative emotions such as anxiety and depression [21,76]. Furthermore, gaming motives may be an intermediate factor, as our study found: Gaming motives were simultaneously significantly correlated with MOGS and IGD. Dysfunctional metacognition activates maladaptive coping strategies and motivation, which causes negative emotions to persist and eventually leads to IGD [77]. Moreover, other studies have different views. For example, metacognitions have a mediating effect on the association between emotional dysregulation and problematic Internet use [78], and online gaming thought suppression and impulsiveness mediate the relationship between metacognition and IGD [79]. Therefore, the association between metacognition and IGD cannot be summarized by simple causality. Other psychological variables, such as motives, coping style, impulsiveness, and emotional regulation, should be considered in future research.

Since maladaptive metacognitions are an important predictor of IGD, interventions specifically addressing maladaptive metacognitions, such as metacognition therapy (MCT), may be beneficial for the prevention and treatment of IGD. MCT, an intervention aimed at modifying dysfunctional metacognition, is effective for treating psychiatric and psychological diseases such as anxiety, depression, and schizophrenia [80-84]. Although MCT is not widely used in the treatment of addictive behaviors, researchers are attempting to prove its efficacy [11]. In some pilot studies, MCT was used to effectively treat alcohol abuse and gambling disorder [85,86]. However, the specific efficacy of MCT for treating IGD needs to be further verified in clinical research. This study provides evidence for the potential value of MCT in the clinical treatment of IGD and offers an effective tool for conducting MCT for IGD specifically in the Chinese population.

Limitations

Although this study has the advantages of a large sample size with people of different ages, it has several limitations that should be considered. First, this study adopted convenience sampling instead of random sampling, and only gamers were included. Therefore, it does not sufficiently represent all Chinese people. Second, collecting data using an online self-report questionnaire may increase the probability of participants giving false answers. However, this procedure is reported to be as reliable as pencil-and-paper surveys [87], which is likely to reduce social desirability and increase levels of honesty [88]. Third, this study lacked test-retest reliability of the C-MOGS; further research is required to test its stability. Finally, as a cross-sectional study, we could not infer the causality of the studied variables. Thus, longitudinal research is needed to further explore the relationship between metacognition and IGD.

Conclusion

In summary, this study offers some evidence that supports the satisfactory psychometric properties of the C-MOGS and highlights the possibility of metacognition as an independent risk factor in gaming behavior. It may be a useful and prospective tool for exploring psychological mechanisms of IGD and helping health professionals identify risky gamers (eg, individuals with more metacognitions about online gaming, specifically negative metacognitions about the uncontrollability of online gaming). Additionally, MCT may be beneficial for the prevention and treatment of IGD. This study may support more attention for metacognitive beliefs in addictive behaviors.

Acknowledgments

The authors would like to thank all the study participants. This work was supported by the Chinese National Programs for Brain Science and Brain-like Intelligence Technology (number 2021ZD0202102), Hunan Provincial Social Science Achievement Evaluation Committee (number XSP24YBZ185), the Natural Science Foundation of Hunan Province (number 2023JJ40476), Hunan Provincial Innovation Foundation for Postgraduate (number QL20230053), and Postgraduate Innovative Project of Central South University (number 2023XQLH041).

Conflicts of Interest

None declared.

References

1. Flavell JH. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist 1979;34(10):906-911. [doi: 10.1037//0003-066x.34.10.906]
2. Metcalfe J. Metacognitive judgments and control of study. Curr Dir Psychol Sci 2009 Jun 01;18(3):159-163 [FREE Full text] [doi: 10.1111/j.1467-8721.2009.01628.x] [Medline: 19750138]
3. Fernandez-Duque D, Baird JA, Posner MI. Executive attention and metacognitive regulation. Conscious Cogn 2000 Jun;9(2 Pt 1):288-307. [doi: 10.1006/ccog.2000.0447] [Medline: 10924249]
4. Heyes C, Bang D, Shea N, Frith CD, Fleming SM. Knowing ourselves together: the cultural origins of metacognition. Trends Cogn Sci 2020 May;24(5):349-362 [FREE Full text] [doi: 10.1016/j.tics.2020.02.007] [Medline: 32298621]
5. Schraw G, Moshman D. Metacognitive theories. Educ Psychol Rev 1995 Dec;7(4):351-371. [doi: 10.1007/bf02212307]
6. Rabbitt PM. Error correction time without external error signals. Nature 1966 Oct 22;212(5060):438-438. [doi: 10.1038/212438a0] [Medline: 5970176]
7. Bryce D, Whitebread D, Szűcs D. The relationships among executive functions, metacognitive skills and educational achievement in 5 and 7 year-old children. Metacognition Learning 2014 Jul 9;10(2):181-198. [doi: 10.1007/s11409-014-9120-4]
8. Adam KCS, Vogel EK. Confident failures: Lapses of working memory reveal a metacognitive blind spot. Atten Percept Psychophys 2017 Jul 3;79(5):1506-1523 [FREE Full text] [doi: 10.3758/s13414-017-1331-8] [Medline: 28470554]
9. Wells A, Matthews G. Attention and Emotion: A Clinical Perspective. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 1994.
10. Myers EJ, Abel DB, Mickens JL, Russell MT, Rand KL, Salyers MP, et al. Meta-analysis of the relationship between metacognition and disorganized symptoms in psychosis. Schizophr Res 2024 Feb;264:178-187. [doi: 10.1016/j.schres.2023.12.009] [Medline: 38154360]
11. Hamonniere T, Varescon I. Metacognitive beliefs in addictive behaviours: A systematic review. Addict Behav 2018 Oct;85:51-63. [doi: 10.1016/j.addbeh.2018.05.018] [Medline: 29852356]
12. Casale S, Musicò A, Spada MM. A systematic review of metacognitions in Internet Gaming Disorder and problematic Internet, smartphone and social networking sites use. Clin Psychol Psychother 2021 Nov 04;28(6):1494-1508 [FREE Full text] [doi: 10.1002/cpp.2588] [Medline: 33783920]
13. Spada MM, Caselli G. The Metacognitions about Online Gaming Scale: Development and psychometric properties. Addict Behav 2017 Jan;64:281-286. [doi: 10.1016/j.addbeh.2015.07.007] [Medline: 26210288]
14. Wells A, Matthews G. Modelling cognition in emotional disorder: the S-REF model. Behav Res Ther 1996;34(11-12):881-888. [doi: 10.1016/s0005-7967(96)00050-2] [Medline: 8990539]
15. Wells A. Emotional disorders and metacognition: innovative cognitive therapy. Psychiatric Ment Health Nurs 2002 Nov 23;9(2):246-247. [doi: 10.1046/m.1365-2850.2002.00006.x]
16. Wells A. The Self-Regulatory Executive Function (S-REF) Model. In: Emotional Disorders and Metacognition: Innovative Cognitive Therapy. Hoboken, NJ: John Wiley & Sons, Ltd; 2002.
17. Sun X, Zhu C, So SHW. Dysfunctional metacognition across psychopathologies: A meta-analytic review. Eur Psychiatry 2017 Sep;45:139-153. [doi: 10.1016/j.eurpsy.2017.05.029] [Medline: 28763680]
18. Hoven M, Lebreton M, Engelmann JB, Denys D, Luigjes J, van Holst RJ. Abnormalities of confidence in psychiatry: an overview and future perspectives. Transl Psychiatry 2019 Oct 21;9(1):268 [FREE Full text] [doi: 10.1038/s41398-019-0602-7] [Medline: 31636252]
19. Rosello R, Martinez-Raga J, Tomas JM, Mira A, Cortese S. Cognitive and behavioral profiles in children with autism spectrum disorder with and without attention-deficit/hyperactivity disorder. Child Adolesc Ment Health 2023 May 19;28(2):269-276. [doi: 10.1111/camh.12562] [Medline: 35441444]
20. Lungu PF, Lungu C, Ciobîcă A, Balmus IM, Bolo A, Dobrin R, et al. Metacognition in schizophrenia spectrum disorders-current methods and approaches. Brain Sci 2023 Jun 28;13(7):1004 [FREE Full text] [doi: 10.3390/brainsci13071004] [Medline: 37508936]
21. Spada MM, Caselli G, Nikčević AV, Wells A. Metacognition in addictive behaviors. Addict Behav 2015 May;44:9-15 [FREE Full text] [doi: 10.1016/j.addbeh.2014.08.002] [Medline: 25182375]

22. Spada MM, Wells A. Metacognitions about alcohol use in problem drinkers. Clin Psychology and Psychoth 2006 Feb 21;13(2):138-143. [doi: 10.1002/cpp.478]
23. Lindberg A, Fernie BA, Spada MM. Metacognitions in problem gambling. J Gambl Stud 2011 Mar;27(1):73-81. [doi: 10.1007/s10899-010-9193-1] [Medline: 20429027]
24. Gierski F, Spada MM, Fois E, Picard A, Naassila M, Van der Linden M. Positive and negative metacognitions about alcohol use among university students: Psychometric properties of the PAMS and NAMS French versions. Drug Alcohol Depend 2015 Aug 01;153:78-85. [doi: 10.1016/j.drugalcdep.2015.06.003] [Medline: 26094183]
25. Dragan WL, Domozych W, Czerski PM, Dragan M. Positive metacognitions about alcohol mediate the relationship between variability and problematic drinking in a sample of young women. Neuropsychiatr Dis Treat 2018;14:2681-2688 [FREE Full text] [doi: 10.2147/NDT.S169514] [Medline: 30349266]
26. Khosravani V, Zandifar A, Sharifi Bastan F, Kolubinski DC, Amirinezhad A. Psychometric properties of the Persian versions of the Positive Alcohol Metacognitions Scale (Persian-PAMS) and the Negative Alcohol Metacognitions Scale (Persian-NAMS) in alcohol-dependent individuals. Addict Behav 2020 Feb;101:106113. [doi: 10.1016/j.addbeh.2019.106113] [Medline: 31648141]
27. Laghi F, Pompili S, Bianchi D, Lonigro A, Baiocco R. Dysfunctional metacognition processes as risk factors for drunkorexia during adolescence. J Addict Dis 2020;38(3):291-300. [doi: 10.1080/10550887.2020.1756709] [Medline: 32368953]
28. Alma L, Spada MM, Fernie BA, Yilmaz-Samanci AE, Caselli G, Nikčević AV. Metacognitions in smoking: Evidence from a cross-cultural validation of the metacognitions about smoking questionnaire in a Turkish sample. Psychiatry Res 2018 Jan;259:160-168 [FREE Full text] [doi: 10.1016/j.psychres.2017.10.015] [Medline: 29053986]
29. Nikčević AV, Alma L, Marino C, Kolubinski D, Yılmaz-Samancı AE, Caselli G, et al. Modelling the contribution of negative affect, outcome expectancies and metacognitions to cigarette use and nicotine dependence. Addict Behav 2017 Nov;74:82-89. [doi: 10.1016/j.addbeh.2017.06.002] [Medline: 28599165]
30. Caselli G, Fernie B, Canfora F, Mascolo C, Ferrari A, Antonioni M, et al. The Metacognitions about Gambling Questionnaire: Development and psychometric properties. Psychiatry Res 2018 Mar;261:367-374 [FREE Full text] [doi: 10.1016/j.psychres.2018.01.018] [Medline: 29353763]
31. Rogier G, Beomonte Zobel S, Morganti W, Ponzoni S, Velotti P. Metacognition in gambling disorder: A systematic review and meta-analysis. Addict Behav 2021 Jan;112:106600. [doi: 10.1016/j.addbeh.2020.106600] [Medline: 32861990]
32. Hoven M, de Boer NS, Goudriaan AE, Denys D, Lebreton M, van Holst RJ, et al. Metacognition and the effect of incentive motivation in two compulsive disorders: Gambling disorder and obsessive-compulsive disorder. Psychiatry Clin Neurosci 2022 Sep 29;76(9):437-449 [FREE Full text] [doi: 10.1111/pcn.13434] [Medline: 35674699]
33. Hamidi F, Ghasedi J. Cognitive and metacognitive impairments of drug addicted, internet addicted and normal individuals in youth ages: a comparative study. Int J High Risk Behav Addict 2020 Feb 02;9(1):1. [doi: 10.5812/ijhrba.95400]
34. Seyed Hashemi SG, Hosseinnezhad S, Dini S, Griffiths MD, Lin C, Pakpour AH. The mediating effect of the cyberchondria and anxiety sensitivity in the association between problematic internet use, metacognition beliefs, and fear of COVID-19 among Iranian online population. Heliyon 2020 Oct;6(10):e05135 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e05135] [Medline: 33072911]
35. Akbari M, Bahadori MH, Mohammadkhani S, Kolubinski DC, Nikčević AV, Spada MM. A discriminant analysis model of psychosocial predictors of problematic Internet use and cannabis use disorder in university students. Addict Behav Rep 2021 Dec;14:100354 [FREE Full text] [doi: 10.1016/j.abrep.2021.100354] [Medline: 34141856]
36. Balıkçı K, Aydın O, Sönmez İ, Kalo B, Ünal-Aydın P. The relationship between dysfunctional metacognitive beliefs and problematic social networking sites use. Scand J Psychol 2020 Oct;61(5):593-598. [doi: 10.1111/sjop.12634] [Medline: 32145032]
37. Casale S, Caponi L, Fioravanti G. Metacognitions about problematic smartphone use: Development of a self-report measure. Addict Behav 2020 Oct;109:106484. [doi: 10.1016/j.addbeh.2020.106484] [Medline: 32497956]
38. Ünal-Aydın P, Obuća F, Aydın O, Spada MM. The role of metacognitions and emotion recognition in problematic SNS use among adolescents. J Affect Disord 2021 Mar 01;282:1-8. [doi: 10.1016/j.jad.2020.12.103] [Medline: 33387741]
39. Aydın O, Güçlü M, Ünal-Aydın P, Spada MM. Metacognitions and emotion recognition in Internet Gaming Disorder among adolescents. Addict Behav Rep 2020 Dec;12:100296 [FREE Full text] [doi: 10.1016/j.abrep.2020.100296] [Medline: 33364305]
40. Zhang MX, Lei LSM, Wells A, Dang L, Wu AMS. Validation of a Chinese version of the short form of Metacognitions Questionnaire (MCQ-30). J Affect Disord 2020 Dec 01;277:417-424. [doi: 10.1016/j.jad.2020.08.028] [Medline: 32866800]
41. Petry NM, Rehbein F, Gentile DA, Lemmens JS, Rumpf H, Mößle T, et al. An international consensus for assessing internet gaming disorder using the new DSM-5 approach. Addiction 2014 Sep;109(9):1399-1406. [doi: 10.1111/add.12457] [Medline: 24456155]
42. Rumpf H, Achab S, Billieux J, Bowden-Jones H, Carragher N, Demetrovics Z, et al. Including gaming disorder in the ICD-11: The need to do so from a clinical and public health perspective. J Behav Addict 2018 Sep 01;7(3):556-561 [FREE Full text] [doi: 10.1556/2006.7.2018.59] [Medline: 30010410]
43. Kuss D. Internet gaming addiction: current perspectives. PRBM 2013 Nov:125. [doi: 10.2147/prbm.s39476]

44. Zheng H, Hu Y, Wang Z, Wang M, Du X, Dong G. Meta-analyses of the functional neural alterations in subjects with Internet gaming disorder: Similarities and differences across different paradigms. Prog Neuropsychopharmacol Biol Psychiatry 2019 Aug 30;94:109656. [doi: 10.1016/j.pnpbp.2019.109656] [Medline: 31145927]
45. Deng X, Hu Y, Liu C, Li Q, Yang N, Zhang Q, et al. Psychological distress and aggression among adolescents with internet gaming disorder symptoms. Psychiatry Res 2024 Jan;331:115624 [FREE Full text] [doi: 10.1016/j.psychres.2023.115624] [Medline: 38039647]
46. Ahmed GK, Abdalla AA, Mohamed AM, Mohamed LA, Shamaa HA. Relation between internet gaming addiction and comorbid psychiatric disorders and emotion avoidance among adolescents: A cross-sectional study. Psychiatry Res 2022 Jun;312:114584. [doi: 10.1016/j.psychres.2022.114584] [Medline: 35523030]
47. Stevens MW, Dorstyn D, Delfabbro PH, King DL. Global prevalence of gaming disorder: A systematic review and meta-analysis. Aust N Z J Psychiatry 2021 Jun;55(6):553-568. [doi: 10.1177/0004867420962851] [Medline: 33028074]
48. Mak K, Lai C, Watanabe H, Kim D, Bahar N, Ramos M, et al. Epidemiology of internet behaviors and addiction among adolescents in six Asian countries. Cyberpsychol Behav Soc Netw 2014 Nov;17(11):720-728. [doi: 10.1089/cyber.2014.0139] [Medline: 25405785]
49. Xiang Y, Jin Y, Zhang L, Li L, Ungvari GS, Ng CH, et al. An overview of the expert consensus on the prevention and treatment of gaming disorder in China (2019 edition). Neurosci Bull 2020 Jul;36(7):825-828 [FREE Full text] [doi: 10.1007/s12264-020-00475-w] [Medline: 32125603]
50. Liao Z, Huang Q, Huang S, Tan L, Shao T, Fang T, et al. Prevalence of internet gaming disorder and its association with personality traits and gaming characteristics among Chinese adolescent gamers. Front Psychiatry 2020;11:598585 [FREE Full text] [doi: 10.3389/fpsyt.2020.598585] [Medline: 33312143]
51. Ji Y, Yin MXC, Zhang AY, Wong DFK. Risk and protective factors of Internet gaming disorder among Chinese people: A meta-analysis. Aust N Z J Psychiatry 2022 Apr 10;56(4):332-346. [doi: 10.1177/00048674211025703] [Medline: 34250835]
52. Brand M, Young KS, Laier C, Wölfling K, Potenza MN. Integrating psychological and neurobiological considerations regarding the development and maintenance of specific Internet-use disorders: An Interaction of Person-Affect-Cognition-Execution (I-PACE) model. Neurosci Biobehav Rev 2016 Dec;71:252-266 [FREE Full text] [doi: 10.1016/j.neubiorev.2016.08.033] [Medline: 27590829]
53. Brand M, Wegmann E, Stark R, Müller A, Wölfling K, Robbins TW, et al. The Interaction of Person-Affect-Cognition-Execution (I-PACE) model for addictive behaviors: Update, generalization to addictive behaviors beyond internet-use disorders, and specification of the process character of addictive behaviors. Neurosci Biobehav Rev 2019 Sep;104:1-10 [FREE Full text] [doi: 10.1016/j.neubiorev.2019.06.032] [Medline: 31247240]
54. Akbari M, Bahadori MH, Bouruki Milan B, Caselli G, Spada MM. Metacognitions as a predictor of online gaming in adolescents: Psychometric properties of the metacognitions about online gaming scale among Iranian adolescents. Addict Behav 2021 Jul;118:106904. [doi: 10.1016/j.addbeh.2021.106904] [Medline: 33743493]
55. Marino C, Canale N, Vieno A, Caselli G, Scacchi L, Spada MM. Social anxiety and Internet gaming disorder: The role of motives and metacognitions. J Behav Addict 2020 Oct 12;9(3):617-628 [FREE Full text] [doi: 10.1556/2006.2020.00044] [Medline: 32750032]
56. Pontes HM, Griffiths MD. Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. Computers in Human Behavior 2015 Apr;45:137-143. [doi: 10.1016/j.chb.2014.12.006]
57. Poon LYJ, Tsang HWH, Chan TYJ, Man SWT, Ng LY, Wong YLE, et al. Psychometric properties of the Internet Gaming Disorder Scale-Short-Form (IGDS9-SF): systematic review. J Med Internet Res 2021 Oct 18;23(10):e26821 [FREE Full text] [doi: 10.2196/26821] [Medline: 34661543]
58. Yam C, Pakpour AH, Griffiths MD, Yau W, Lo CM, Ng JMT, et al. Psychometric testing of three Chinese online-related addictive behavior instruments among Hong Kong university students. Psychiatr Q 2019 Mar 16;90(1):117-128. [doi: 10.1007/s11126-018-9610-7] [Medline: 30328020]
59. Qin L, Cheng L, Hu M, Liu Q, Tong J, Hao W, et al. Clarification of the cut-off score for nine-item Internet Gaming Disorder Scale-Short Form (IGDS9-SF) in a Chinese context. Front Psychiatry 2020;11:470 [FREE Full text] [doi: 10.3389/fpsyt.2020.00470] [Medline: 32528331]
60. Demetrovics Z, Urbán R, Nagygyörgy K, Farkas J, Zilahy D, Mervó B, et al. Why do you play? The development of the motives for online gaming questionnaire (MOGQ). Behav Res Methods 2011 Sep;43(3):814-825. [doi: 10.3758/s13428-011-0091-y] [Medline: 21487899]
61. Wu AMS, Lai MHC, Yu S, Lau JTF, Lei M. Motives for online gaming questionnaire: Its psychometric properties and correlation with Internet gaming disorder symptoms among Chinese people. J Behav Addict 2017 Mar 01;6(1):11-20 [FREE Full text] [doi: 10.1556/2006.6.2017.007] [Medline: 28264590]
62. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
63. Zhang Y, Liang W, Chen Z, Zhang H, Zhang J, Weng X, et al. Validity and reliability of Patient Health Questionnaire-9 and Patient Health Questionnaire-2 to screen for depression among college students in China. Asia Pac Psychiatry 2013 Dec;5(4):268-275. [doi: 10.1111/appy.12103] [Medline: 24123859]

64. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
65. Zhou Y, Xu J, Rief W. Are comparisons of mental disorders between Chinese and German students possible? An examination of measurement invariance for the PHQ-15, PHQ-9 and GAD-7. BMC Psychiatry 2020 Oct 01;20(1):480 [FREE Full text] [doi: 10.1186/s12888-020-02859-8] [Medline: 33004042]
66. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
67. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
68. Cortina JM. What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology 1993 Feb;78(1):98-104. [doi: 10.1037/0021-9010.78.1.98]
69. Li Y, Tang Y, Huang S, Tan L, Huang Q, Chen X, et al. Role of gaming devices associated with Internet gaming disorder in China: cross-sectional study. JMIR Serious Games 2023 Jan 16;11:e40130 [FREE Full text] [doi: 10.2196/40130] [Medline: 36645700]
70. Sit HF, Chang CI, Yuan GF, Chen C, Cui L, Elhai JD, et al. Symptoms of internet gaming disorder and depression in Chinese adolescents: A network analysis. Psychiatry Res 2023 Apr;322:115097. [doi: 10.1016/j.psychres.2023.115097] [Medline: 36822036]
71. Wang H, Cheng C. The associations between gaming motivation and internet gaming disorder: systematic review and meta-analysis. JMIR Ment Health 2022 Feb 17;9(2):e23700 [FREE Full text] [doi: 10.2196/23700] [Medline: 35175204]
72. Dang L, Yang HM, Spada MM, Wu AMS. A three-wave longitudinal study on the underlying metacognitive mechanism between depression and Internet gaming disorder. J Behav Addict 2024 Jan 05:1. [doi: 10.1556/2006.2023.00072] [Medline: 38183452]
73. Efrati Y, Spada MM. "I have no control over how much time I play" the metacognitions about online gaming scale: Evidence from a cross-cultural validation among Israeli adolescents. Addict Behav 2023 Jun;141:107638. [doi: 10.1016/j.addbeh.2023.107638] [Medline: 36746106]
74. Lin S, Tan L, Chen X, Liao Z, Li Y, Tang Y, et al. Emotion dysregulation and Internet gaming disorder in young people: Mediating effects of negative affect and metacognitions. J Affect Disord 2023 Nov 15;341:104-111. [doi: 10.1016/j.jad.2023.08.077] [Medline: 37597784]
75. Gandolfi E, Soyturk I, Ferdig RE. Evaluating U.S. gamers' metacognitions about digital entertainment: Validation of Metacognition about Online Gaming Scale in the U.S. context. J Affect Disord 2021 Dec 01;295:954-959. [doi: 10.1016/j.jad.2021.08.133] [Medline: 34706468]
76. Caselli G, Marino C, Spada MM. Modelling online gaming metacognitions: the role of time spent gaming in predicting problematic Internet use. J Rat-Emo Cognitive-Behav Ther 2020 Jul 13;39(2):172-182. [doi: 10.1007/s10942-020-00365-0]
77. Marino C, Spada MM. Dysfunctional cognitions in online gaming and Internet gaming disorder: a narrative review and new classification. Curr Addict Rep 2017 Sep 01;4(3):308-316. [doi: 10.1007/s40429-017-0160-0]
78. Akbari M. Metacognitions or distress intolerance: The mediating role in the relationship between emotional dysregulation and problematic internet use. Addict Behav Rep 2017 Dec;6:128-133 [FREE Full text] [doi: 10.1016/j.abrep.2017.10.004] [Medline: 29450248]
79. Efrati Y, Kolubinski DC, Marino C, Spada MM. Modelling the contribution of metacognitions, impulsiveness, and thought suppression to behavioural addictions in adolescents. Int J Environ Res Public Health 2021 Apr 06;18(7):1 [FREE Full text] [doi: 10.3390/ijerph18073820] [Medline: 33917425]
80. Balzan RP, Mattiske JK, Delfabbro P, Liu D, Galletly C. Individualized metacognitive training (MCT+) reduces delusional symptoms in psychosis: a randomized clinical trial. Schizophr Bull 2019 Jan 01;45(1):27-36 [FREE Full text] [doi: 10.1093/schbul/sby152] [Medline: 30376124]
81. McEvoy PM. Metacognitive therapy for anxiety disorders: a review of recent advances and future research directions. Curr Psychiatry Rep 2019 Mar 18;21(5):29. [doi: 10.1007/s11920-019-1014-3] [Medline: 30880368]
82. Jelinek L, Van Quaquebeke N, Moritz S. Cognitive and metacognitive mechanisms of change in metacognitive training for depression. Sci Rep 2017 Jun 14;7(1):3449 [FREE Full text] [doi: 10.1038/s41598-017-03626-8] [Medline: 28615651]
83. Cheli S, Cavalletti V, Lysaker PH, Dimaggio G, Petrocchi N, Chiarello F, et al. A pilot randomized controlled trial comparing a novel compassion and metacognition approach for schizotypal personality disorder with a combination of cognitive therapy and psychopharmacological treatment. BMC Psychiatry 2023 Feb 20;23(1):113 [FREE Full text] [doi: 10.1186/s12888-023-04610-5] [Medline: 36803673]
84. Hotte-Meunier A, Penney D, Mendelson D, Thibaudeau E, Moritz S, Lepage M, et al. Effects of metacognitive training (MCT) on social cognition for schizophrenia spectrum and related psychotic disorders: a systematic review and meta-analysis. Psychol. Med 2023 Sep 29;54(5):914-920. [doi: 10.1017/s0033291723002611]
85. Caselli G, Martino F, Spada MM, Wells A. Metacognitive therapy for alcohol use disorder: a systematic case series. Front Psychol 2018;9:2619 [FREE Full text] [doi: 10.3389/fpsyg.2018.02619] [Medline: 30619014]

86. Gehlenborg J, Bücker L, Berthold M, Miegel F, Moritz S. Feasibility, acceptance, and safety of metacognitive training for problem and pathological gamblers (Gambling-MCT): a pilot study. J Gambl Stud 2021 Jun;37(2):663-687 [FREE Full text] [doi: 10.1007/s10899-020-09975-w] [Medline: 32955694]
87. Weigold A, Weigold IK, Russell EJ. Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. Psychol Methods 2013 Mar;18(1):53-70. [doi: 10.1037/a0031607] [Medline: 23477606]
88. Wood RTA, Griffiths MD, Eatough V. Online data collection from video game players: methodological issues. Cyberpsychol Behav 2004 Oct;7(5):511-518. [doi: 10.1089/cpb.2004.7.511] [Medline: 15667045]

Abbreviations

CAS: cognitive attentional syndrome
CFA: confirmatory factor analysis
CFI: comparative fit index
C-MOGS: Chinese version of the Metacognitions about Online Gaming Scale
EFA: exploratory factor analysis
GAD-7: Generalized Anxiety Disorder-7
GFI: goodness-of-fit index
***ICD-11*:** *Eleventh Revision of the International Classification of Diseases*
IGD: Internet Gaming Disorder
IGDS9-SF: Internet Gaming Disorder Scale-Short Form
I-PACE: Interaction of Person-Affect-Cognition-Execution
KMO: Kaiser-Meyer-Olkin
MCT: metacognition therapy
MOGQ: Motives for Online Gaming Questionnaire
MOGS: Metacognitions about Online Gaming Scale
N-MOG: Negative Metacognitions about Online Gaming
N-MOGD: Negative Metacognitions about the Dangers of Online Gaming
N-MOGU: Negative Metacognitions about the Uncontrollability of Online Gaming
PCA: principal component analysis
PHQ-9: Patient Health Questionnaire-9
P-MOG: Positive Metacognitions about Online Gaming
RMSEA: root mean squared error of approximation
SRMR: standardized root of the mean square residual
TLI: Tucker-Lewis Index

Edited by T Rashid Soron; submitted 25.01.23; peer-reviewed by H Zheng, Y Efrati; comments to author 04.08.23; revised version received 28.09.23; accepted 21.02.24; published 22.04.24.

Please cite as:
Lin S, Chen X, Tan L, Liao Z, Li Y, Tang Y, Huang Q, Shen H
Psychometric Properties of the Metacognitions About Online Gaming Scale in the Chinese Population and Its Relationship With Internet Gaming Disorder: Cross-Sectional Study
JMIR Serious Games 2024;12:e45985
URL: https://games.jmir.org/2024/1/e45985
doi:10.2196/45985
PMID:38648634

The Mediating Role of Problematic Use of Loot Boxes Between Internet Gaming Disorder and Online Gambling Disorder: Cross-Sectional Analytical Study

Joaquín González-Cabrera[1], PhD; Vanessa Caba-Machado[1], PhD; Adoración Díaz-López[1], PhD; Susana Jiménez-Murcia[2,3,4,5], PhD; Gemma Mestre-Bach[1], PhD; Juan M Machimbarrena[6], PhD

1
2
3
4
5
6

Corresponding Author:
Juan M Machimbarrena, PhD

Abstract

Background: The video game industry has introduced a new form of monetization through microtransactions. A controversial example has been the so-called "loot boxes" (LBs) as virtual objects, which are randomized and bought with legal money. In recent years, LBs have come to connect 2 distinct problem behaviors, namely internet gaming disorder (IGD) and online gambling disorder (OGD). Many association studies have been conducted on the 3 constructs, but few have delved into the relationship of problematic use of LBs (PU-LB) with IGD and OGD.

Objective: This study aims to explore the mediating role of the PU-LB between IGD and OGD.

Methods: This cross-sectional and analytical study used incidental sampling in 24 Spanish schools. The final sample consisted of 542 participants (male: n=523, 96.5%; age: range 11 - 30 y) who played video games, bought LBs, and had gambled online in the last 12 months. Participants then completed the Spanish versions of the Internet Gaming Disorder Scale–Short Form, Online Gambling Disorder Questionnaire, and PU-LB scale.

Results: IGD scores were found to be significantly associated with both PU-LB (*r*=0.473, *P*<.001) and OGD (*r*=0.209, *P*<.001). Moreover, PU-LB was significantly associated with OGD (*r*=0.351, *P*<.001). The structural equation model results indicated that IGD had no significant direct effect on OGD (*P*=.903). However, the indirect effect of IGD on OGD through PU-LB was significant (*P*<.001). Therefore, PU-LB fully mediated the relationship between IGD and OGD. Furthermore, these results were found in the subsamples of both minors (<18 y) and young adults (≥18 y).

Conclusions: It is suggested that there is a mediation effect of problematic LB use between internet gambling and online gambling problems in both minors and young adults. This has potential practical implications by providing more evidence on how LBs have become a hinge feature between 2 clinically relevant and independent issues. In this regard, adequate industry self-regulation is needed, and effective legislation for the protection of minors is necessary.

(JMIR Serious Games 2024;12:e57304) doi:10.2196/57304

KEYWORDS

loot boxes; loot box; gaming; gambling; problematic; video games; game; games; addict; addiction; addictions; addictive; internet; virtual object; virtual objects; gamification; IGD; OGD; monetize; monetization; reward; rewards; incentive; incentives; internet gaming disorder; online gambling disorder

Introduction

Video games are a form of interactive entertainment that has gained enormous popularity around the world. Video games can be seen even as a form of cultural expression because they reflect the creativity, values, beliefs, and experiences of both game developers and gamers alike [1].

According to a report by DFC Intelligence, 40% of the global population plays video games, which represents an estimated 3.1 billion people playing video games as of 2022 [2]. In 2022, the global video game industry reached a market size of approximately US $246 billion [3]. According to the Spanish Video Game Association [4], in 2022, a total of 18.2 million people in Spain were recorded as gamers (there will be more than 48 million inhabitants in 2023), with 53% being male and 47% female. According to the report, Spaniards play video games for 7.42 hours per week (8% less than in the previous report), and turnover has risen to 2012 million euros (an increase of 12% compared with the previous year). Furthermore, the age profile of Spanish video game players is mainly young, with 79% between 6 and 11 years old, 84% between 11 and 14 years old, and 71% between 15 and 24 years old [4]. In addition to the increase in video game sales, one of the factors behind this economic success lies in the incorporation of in-game purchases, which represent an increasingly large revenue stream for the industry [5,6]. In this regard, in 2023, a national study in Spain showed that 17.7% of adolescents between 14 and 18 years of age had gambled online. Of these, more than 50% claim to have gambled in the context of video games [7].

There is an ever-increasing number of video games that offer microtransactions (ie, the payment of a stipulated price for a specific, well-known skin or perk) [8]. Each video game features its own type of microtransactions, some of which are only of an aesthetic nature, while others may influence the dynamics of the game. Within the microtransactions, there is a special modality that in recent years has attracted the attention of researchers for its possible relation to random reward mechanisms: loot boxes (LBs), which are also called crates, cases, or chests. The acquisition of LBs involves the purchase of a virtual object (which can be acquired in various manifestations such as boxes, slot machines, chests, or in the form of animals), which is randomized and paid for with legal tender (this can be obtained from a prepaid card, a credit card payment or by prepurchasing currency from a game or environment to buy LBs) [9]. The fact that this virtual object is the product of a random reward is what has made it similar to gambling, as both share a random reward mechanism [9-11]. Paradigmatic examples of LBs in video games may be the case analyzed by Lemmens [12] on FIFA Ultimate Team and those analyzed by Xiao et al [13], with some of the top downloaded games for Android, such as Game of Thrones: Conquest or Pokemon GO, including LBs.

On the one hand, some primary studies have associated LB purchases with clinical problems with video games (usually assessed with questionnaires that follow the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* [*DSM-5*] criteria for internet gaming disorder [IGD]). Supporting evidence for this link has been found by authors in cross-sectional studies [6,14,15], but there is hardly any longitudinal evidence [16]. On the other hand, there is more evidence of a direct relationship between the purchase of LBs and problem gambling [5,10,11,15,17-24]. However, only a few studies have related the purchase of LBs to clinical problems of online gambling (following the *DSM-5* [25] and *International Classification of Diseases, 11th Revision* [*ICD-11*] criteria) [26], and almost all these have been cross-sectional studies. One of them is the research with minors and adults by González-Cabrera et al [27] carried out in Spain with a large sample of more than 6500 participants, where 3 out of every 10 participants purchased LBs in the last 12 months. Although limited, there is also evidence of the association of these problems over time in minors, with over 50% of LB purchasers still buying after 6 months [16]. Studies have also already been conducted as adults linking LBs to gambling over time [28].

Overall, we still need to answer the possible hypotheses raised by Spicer et al [29] about LBs and gambling: either (1) users who gamble in other environments buy more LBs; (2) buyers of LBs are more likely to start gambling, through the "gateway effect"; or (3) there is a complex and dynamic relationship between both behaviors, where gambling is known to interact with other risky behaviors. In line with the latter suggestion, LBs have been related to gaming and gambling problems in minors and mostly adults [19,27,29,30]. Nevertheless, the possible mediating role of LBs between both clinical problems has been much less addressed. LBs have been a novel and relatively massive phenomenon over the past 5 years (approximately). This phenomenon has become the hinge that can bridge 2 clinical and nosological entities that appear separately in the diagnostic manuals [25] and that have not been modified in the current revision of the *DSM-5-TR* [31]. As such, a pioneering study was conducted with adults in which the mediation of microtransaction engagement between gaming and gambling was analyzed [32]. This mediating construct was assessed using the Risky Loot-Box Index (RLI) [17], which captures cognitive concern about LB use, impulsive use, and chasing losses, but it has limitations in terms of validity and reliability. In addition, this study did not use a clinical assessment tool for online gambling (and did not consider research with minors).Despite these limitations, the results were interesting as they did not achieve a complete mediation (ie, the direct relationship between video game problems and betting was also significant). Overall, the results indicate that participants with IGD were more likely to purchase microtransactions and to report more gambling-related problems. It is also possible that there have been significant changes in the consumption of LBs since the 2020 release (King et al [32]), as this business model has become increasingly common and has grown in recent years [13].

This study, based on the study by King et al [32], included 2 clinical measures of online gaming and gambling problems and an instrument with adequate validity indicators of problematic LB use. In addition, a large sample of adults and minors (often less addressed in the literature) was included, with the latter requiring special safeguards regarding the LB phenomenon as covert gambling [9]. The authors posit the hypothesis that there is no direct relationship between clinical problems with video games and online gambling, unless there is a problematic use of LBs (PU-LB) that mediates this relationship and therefore generates a significant indirect effect between the 2 clinical variables in the model. Thus, the aim was to perform a mediation of the PU-LB scale between IGD and online gambling disorder (OGD) in a sample of minors and young adults.

Methods

Design and Recruitment

This study used a cross-sectional design. The sampling was incidental and was carried out in 24 nonuniversity educational centers in 8 Spanish regions (ie, Asturias, Cantabria, Castilla la Mancha, Castilla y Leon, Comunidad de Madrid, Comunidad

Foral de Navarra, Comunidad Valenciana, and País Vasco). The educational stages covered ranged from compulsory secondary education (approximately between 11 and 16 y of age) to baccalaureate education (approximately between 16 and 18 y of age), with the addition of vocational training (where the ages range between 15 and 30 y old). This sample is part of a larger study on internet risks in adolescence. The final sample consisted of participants (both younger and older than 18 years) who had answered “yes” to the following 3 questions: have you played video games in the last 12 months, have you bought any video game LBs with money in the last 12 months, and have you gambled online in any type of game in the last 12 months?

Assessment Instruments

Participants were initially asked sociodemographic questions (gender, age, study center, and province).

IGD was assessed with the Spanish version of the Internet Gaming Disorder Scale–Short Form (IGD9-SF) [33-35]. This scale consists of 9 items based on the *DSM-5* criteria for IGD (eg, “Have you deceived any of your family, therapists, or friends about the time you spend gaming?”) [25]. The scale response options range from 0 (never) to 4 (very often). The total score ranges between 0 and 36, with greater scores suggesting higher symptom-severity of disordered gaming. In terms of internal reliability, the Cronbach α coefficient and ω coefficient in the present sample were 0.85 and 0.86, respectively.

As for OGD, this construct was evaluated with the Online Gambling Disorder Questionnaire (OGD-Q) [36]. This scale consists of 11 items that assess OGD in adolescence (eg, “Do you feel nervous, irritated, or angry when trying to reduce or stop gambling online?”). The scale response options range from 0 (never) to 4 (every day), with greater scores suggesting higher symptom-severity of disordered gambling. The total score ranges between 0 and 45. The scale Cronbach α coefficient and ω coefficient were 0.92 and 0.92, respectively.

The PU-LB scale [27], consists of 18 items assessing the potentially problematic nature of engaging in LB purchasing behavior (eg, “Loot boxes have caused problems in my life (either social, economic, family, school, or work. etc” or “I usually buy loot boxes to feel better or happier”). The scale response options range from 0 (strongly disagree) to 5 (strongly agree), with total scores ranging from 0 to 90, where higher scores suggest a more PU-LB. In relation to its internal reliability, the Cronbach α coefficient and ω coefficient in the present sample were 0.92 and 0.87, respectively.

Procedure

The survey was conducted online through the Survey Monkey platform on either a mobile device or computer. The participants were given access to and supervised by their teachers. The evaluation was carried out in educational centers (during the school timetable). The researchers previously trained the teachers in data collection. In addition, participant detection mechanisms were enabled, such as those suggested by Niessen et al [37], which included maximum response time “longstring” and “person-fit statistics.” The time needed to complete the questionnaires ranged between 5 and 15 minutes, depending on students’ age and reading comprehension.

Ethical Considerations

The study was conducted with the authorization of all the participants in the investigation and with the consent of the school directors, students, and families. Students and families’ collaboration was voluntary, anonymous, and disinterested. The project was approved by the Research Ethics Committee of International University of La Rioja (Spain) (PI007-2020 y PI001/2021), and the Juvenile Prosecutor’s Office was informed. The study received consent from all participants and school principals. Consent forms were sent to parents or guardians of participants younger than 18 years, and the purpose of the study was explained. About 0.8% of the participants did not want to respond to the questionnaire, while less than 1% of parents or guardians refused participation. Participants older than 18 years provided informed consent when completing the survey. Although there were no formal exclusion criteria, except for refusal to participate by parents or guardians for the overall sample, to be included in this study, participants had to answer affirmatively to a dichotomous question (yes or no) on whether they had played video games in the last 12 months, whether they had gambled online in the last 12 months, and whether they had bought LBs in the last 12 months. Only those who answered “yes” to each of these questions were assessed and included in the study.

Statistical Analysis

The SPSS (version 26; IBM Corp) program was used to (1) explore and screen all data through descriptive statistics; (2) test for reliability by Cronbach α, ω coefficient, and normality through skewness and kurtosis; and (3) explore the relationships between variables through bivariate correlations. The absolute values of skewness and kurtosis are normal when they are below ±3 for skewness and ±10 for kurtosis [38].

MPLUS (version 8.0; Muthen & Muthen) [39] was used to test (1) the factor structure of the PU-LB through confirmatory factor analysis (CFA) and (2) the relationships between IGD, PU-LB, and OGD through structural equation modeling analysis. The Maximum Likelihood Robust Estimator was used, and the fit of the model was estimated with the most reliable fit indices [40]: the Satorra-Bentler chi-square (S-Bχ^2), the root mean square error of approximation (RMSEA), the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the standardized root mean square residual (SRMR). A model was considered to adequately fit the data at values ≥.90 for the CFI and TLI, with values above .95 preferred, and values ≤.08 for the RMSEA and SRMR [38]. The significance of mediational paths was tested by means of bias-corrected bootstrapping with 5000 samples.

Results

Participants, Descriptive Statistics, Normality, and Reliability

The sample was composed of 542 participants (male: n=523, 96.5%) who played video games, purchased LBs, and gambled online in the last 12 months. The average age of the sample was 17.78 (SD 2.78) years (age: range 11 - 30 y), out of which 47.2% (n=256) were minors ($mean_{age}$ 15.3, SD 1.71 y) and the remaining 56.3% (n=306) were young adults older than 18 years ($mean_{age}$ 19.6, SD 1.88 y). There were also 107 (19.7%) students in compulsory secondary education, 36 (6.6%) in baccalaureate education, and 399 (73.7%) in vocational training.

Descriptives, segregated by age groups (minors and adults), for all variables, including the means, SDs, and Pearson bivariate correlations between the variables of the study, are presented in Table 1. Moreover, for the total sample (minors and adults), the descriptive statistics were as follows: for IGD, mean 8.216 (SD 0.301), skewness=1.023, and kurtosis=0.858; for PU-LB, mean 12.583 (SD 14.890), skewness=1.590, and kurtosis=2.794; and for OGD, mean 4.053 (SD 7.104), skewness=2.585, and kurtosis=7.385.

IGD scores were found to be significantly associated with both PU-LB (*r*=0.473, *P*<.001) and OGD (*r*=0.209, *P*<.001). Additionally, PU-LB was significantly associated with OGD (*r*=0.351, *P*<.001).

Table . Correlation matrix, descriptive statistics for internet gaming disorder (IGD), problematic use of loot boxes (PU-LB), and online gambling disorder (OGD). The results for the minors (n=236) are shown below the diagonal. The results for adults (n=306) are shown above the diagonal.

	Variables			Minors (n=256)			Young adults (n=306)		
	IGD	PU-LB	OGD	Questionnaire Score, mean (SD)	Skew[a]	Kurt[b]	Questionnaire Score, mean (SD)	Skew	Kurt
IGD	—[c]	0.477[d]	0.229[d]	8.22 (7.07)	1.02	1.23	8.2 (7.48)	1.025	0.64
PU-LB	0.484[d]	—	0.354[d]	15.33 (14.83)	0.96	0.12	10.46 (14.6)	2.203	6.02
OGD	0.179[d]	0.410[d]	—	3.3 (6.09)	3.19	12.18	4.62 (7.75)	2.282	5.33

[a]Skew: skewness.
[b]Kurt: kurtosis.
[c]Not applicable.
[d]All correlations were significant at *P*<.001.

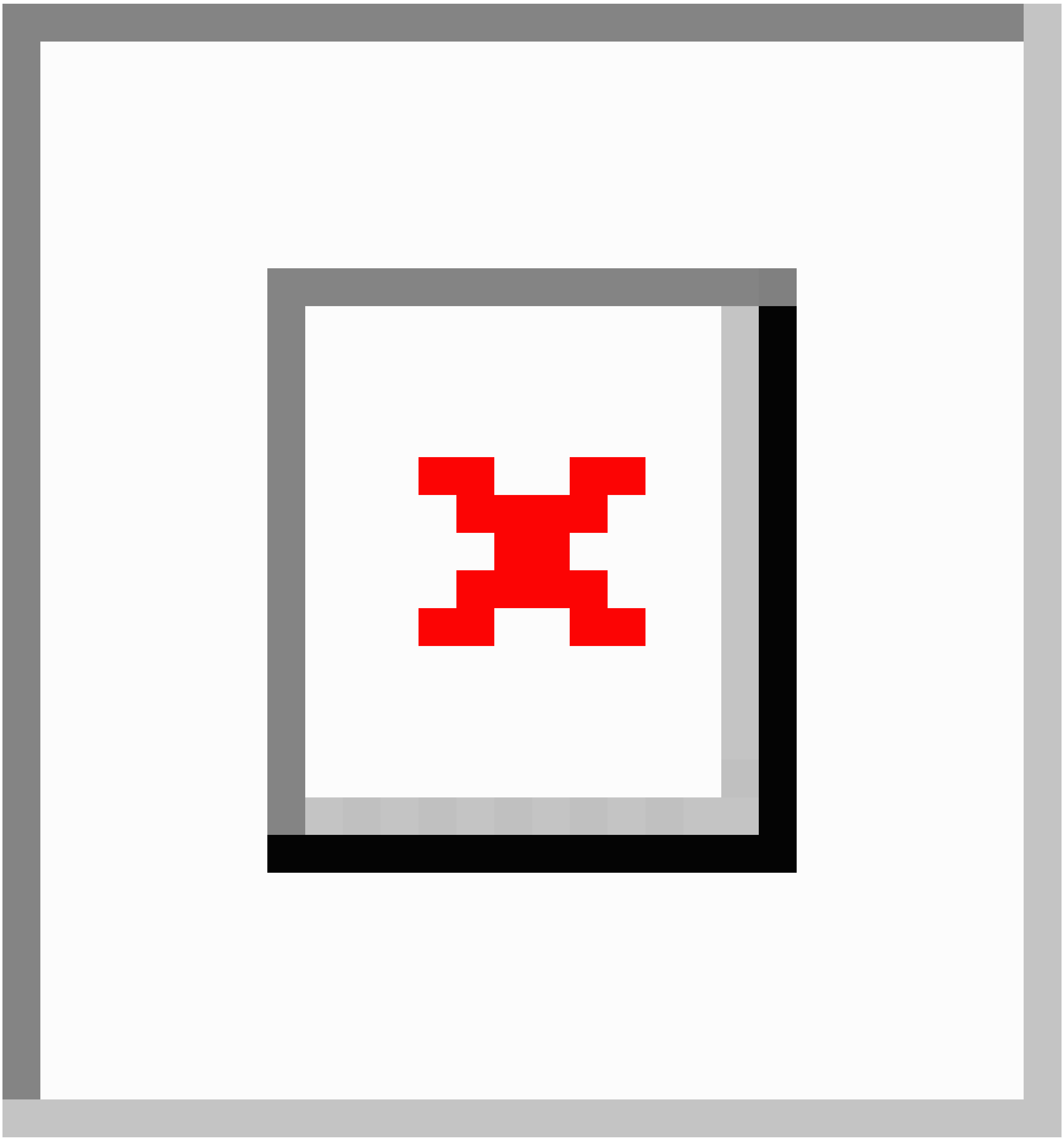
Structural Mediation Model

Before testing the structural mediation model, the CFA model of PU-LB (which had been validated through exploratory factor analysis but not through CFA) yielded some evidence for satisfactory fit (S-Bχ^2_{125}=300.207; RMSEA=.051, 90% CI .044-.058; CFI=.929; TLI=.913; SRMR=.054). The IGDS9-SF and OGD-Q had been previously validated and had shown good structural properties in previous studies [33,36].

Figure 1 displays standardized path coefficients for the structural equation modeling with the total sample and subsamples of minors and young adults. The model for the overall sample had an adequate fit (S-Bχ^2_{647}=1156.821; RMSEA=.038, 90% CI .035-.042; CFI=.924; TLI=.918; SRMR=.055). All items loading onto latent variable were significant (*P*<.001) and ranged from .45 to .86. Moreover, model for the subsample of young adults (*n*=306) (S-Bχ^2_{647}=1057.634; RMSEA=.046, 90% CI .041-.050; CFI=.910; TLI=.908; SRMR=.064), and minors (n=236) (S-Bχ^2_{647}=980.688; RMSEA=.047, 90% CI .041-.052; CFI=.892; TLI=.882; SRMR=.065), showed inconsistent indexes for CFI and TLI, and while the RMSEA and SRMR values remain acceptable, the values for the CFI fail to meet the cut-off (≥.90). However, Raykov [41] defended that CFI is a measure based on noncentrality and therefore, could be biased.

The results in the total sample indicated that IGD did not have a significant direct effect on OGD (β=.004, *P*=.903). However, the indirect effect of IGD on OGD by PU-LB was significant: (β=.223, 95% CI .131-.338, *P*<.001). Therefore, PU-LB fully mediated the relationship between IGD and OGD. In addition, these results were also found in the subsamples of minors (β=.327, 95% CI .177-.561, *P*<.001) and young adults (β=.187, 95% CI .072-.324, *P*<.001).

Figure 1. Structural equation model for IGD, PU-LB, and OGD with standardized factorial loadings. ***P*<.001; the values provided are in the format "total sample, (young adults), [minors]." IGD: internet gaming disorder; PU-LB: problematic use of loot boxes; OGD: online gambling disorder.



Discussion

The business model of many video game companies has been linked to LBs; this is worrying for many sectors of society, and this has generated wide interest in the academic and research context. Moreover, LBs appear to be a hinge that links 2 problem behaviors (IGD and OGD). In this regard, at an early stage in the study of LBs, King et al [32] conducted a mediation of microtransaction engagement between gaming and gambling in adults. There is evidence of an increase in the number of video games with LBs in recent years [13] and the difficulty for laws restricting the use of LBs to be effective [42]. The aim was of this study to perform a model that would allow us to analyze whether there would be a direct relationship between IGD and OGD in video game users who are LBs buyers and online gamblers in the last year or if, on the contrary, there would be an indirect effect thanks to the mediation of the PU-LB. The results suggest that because of the significant indirect effect (and the absence of direct effect), there is a total mediation effect. This reinforces the idea that the PU-LB may be a problem connecting 2 different problem behaviors (IGD and OGD) and that LBs, as a random reward mechanism, can be associated with gambling as a gateway or as a further gambling mechanism [29]. Thus, LBs seem to have spurred and connected 2 very pernicious issues to each other, adding a layer of complexity to the problem. Furthermore, this phenomenon is especially worrisome because minors are involved in these mechanisms (the model adjusts for the total sample and for minors and adults). From these data, important theoretical implications are derived. On the one hand, problems with gaming do not have a direct effect on gambling, and on the other hand, PU-LB generates a full mediation, justifying that hinge role. It should be noted that the entire study sample consisted of buyers of LBs, so it is not only necessary to buy them but also to present a problem with them. This is also related to the emphasis of the *ICD-11* [23] on the consequences that behavioral disorders should have. With the above, it is clear that industry mechanisms do not seem to be sufficient to regulate this process [43], nor are the governmental measures taken in some countries to curb the problem [44].

The study by King et al [32], as well as other studies, used the RLI to assess the risk of LB use [10,28,32]. In general, the RLI shows deficiencies in its psychometric validation process, as well as drawbacks for not covering other key aspects related to the problems that may arise due to the behavior of purchasing LBs (eg, impulsivity to buy more LBs, personal consequences, salience of play time or guilt, among others). For this reason, we used the PU-LB [16], which presents adequate validity and reliability indicators; its content includes validity indicators of general problems about LBs and specific indicators regarding the association among LBs, gaming, and gambling (mood regulation through the purchase or opening of LBs, postponing activities to get LBs, feeling the urge to buy them, thinking about the purchase activity or feeling bad about the time or money invested, etc). Furthermore, contributing to the pioneering work of King et al [32] is the use of clinical questionnaires (IGDS9-SF and OGD-Q) based on the diagnostic criteria of the *DSM-5* [25] and *ICD-11* [45] (including minors and adults). In this sense, there are no data to compare with, as this study has a singular focus.

This study has some relevant limitations. First, only self-report measures were used, which may generate response bias and social desirability bias. Second, there may be a retrospective recall bias, as participants were asked to think back to what they did in the last 12 months. Third, although the sample of participants was large and geographically dispersed, the sampling was not random, so it is not representative of the Spanish context. Fourth, there is an overrepresentation of male participants in the study, which is a common issue in many studies, since consumers of video games, gambling and LBs are mostly male. However, these data also indicate that boys require special attention regarding the LB problem (at least in the Spanish context). Fifth, the OGD-Q questionnaire is validated in Spanish adolescents (up to the age of 19 years) and not in adults. Although the reliability indicators are adequate, this may be a limitation of the study. Sixth, all parameters, apart from the OGD variable in the group of minors, exhibited skewness and kurtosis values indicative of a normal distribution. However, it is emphasized that the lack of normality does not pose a methodological obstacle, as a robust approach was used to effectively address the presence of nonnormal distributions in the statistical analysis. Finally, the fit indices of the model for minors are slightly below the thresholds considered good (particularly the CFI and TLI). However, this may be due to the nature of the constructs and the fact that online gambling is an illegal activity for minors; therefore, the data related to the OGD may have affected the model in general.

Given these potential limitations, future research should include longitudinal designs that take into account the variables used in this study and answer the questions posed by Spicer et al [29] on the "gateway effect" of LBs or the relationship between who opens and who purchases LBs, as opening is likely to focus more on gaming problems and purchasing is likely to focus more on gambling. As such, exploring the independent and additive effect of both roles may be an area for future research. However, this study has been able to address questions also raised by these authors in relation to the relationship between gaming and gambling problems. This study has potential practical implications by providing more evidence on how LBs have become a hinge feature between 2 clinically relevant and independent issues. While there has been very strict legislation on LBs in countries such as Belgium [42], it has not been effective because it has not been properly enforced. Despite this situation, there is still a need for politicians to properly regulate the LB framework and prevent minors from purchasing LBs. The key aspect is to enforce the law and provide resources to do so. In the Spanish context, there has been a draft bill in place since 2022 to regulate random reward mechanisms associated with interactive software products [46]. This draft bill has basic limitations such as a remarkably restrictive definition of a LB (article 3, point C), which, if approved, will be clearly insufficient. This is why, in the Spanish case, it is necessary to improve legislative efforts and add other actions. In this sense, it could help to make it a requirement for the video game industry to include information on LBs (within the framework of the Pan-European Gaming Information [PEGI] system), but this must be clear, specific, and comprehensible [43]. Finally, and perhaps most importantly, there is also a need for psychoeducational actions aimed primarily at preventing the purchase of LBs by minors. These actions should be evidence-based. Education of minors and their families is key and is the future course of action.

In conclusion, this study suggests that there is a mediation effect of problematic LB use between IGD and OGD in both minors and young adults.

Acknowledgments

This study was funded by Universidad Internacional de La Rioja (UNIR Research Plan [2022-2024]), Ministerio de Consumo (Dirección General de Ordenación del Juego-SUBV23/00022), and ITEI-UNIR (B23-006 & B24-001).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

JGC has received funding from the State Programme for R&D&I Oriented to the Challenges of Society (Programa Estatal de I+D+I Orientada a los retos de la Sociedad) and is currently receiving funding for research projects at the Universidad Internacional de La Rioja (Spain) and the Ministry of Consumer (Ministerio de Consumo) in Spain (in particular, for the study of loot boxes). He has also provided consultancy services for the company TICandBot (Logroño, Spain). All this funding did not involve any real conflict in the conduct of this research. The authors have no further competing interests to disclose.

References

1. Muriel D, Crawford G. Video Games as Culture: Considering the Role and Importance of Video Games in Contemporary Society: Routledge; 2018.
2. Crece el público gamer: un estudio revela sus ganancias y tendencias esenciales para este año. Forbes. 2022. URL: https://www.forbesargentina.com/innovacion/crece-publico-gamer-estudio-revela-sus-ganancias tendencias-esenciales-ano-n16215 [accessed 2022-05-25]
3. Evolución del valor de mercado de la industria del videojuego en el mundo entre 2022 y 2030. Videogame indusrty market evolution in the world between 2022 and 2030. Statista. URL: https://es.statista.com/estadisticas/598622/valor-de-mercado-del-videojuego-en-el-mundo/#:~:text=En%202022%2C%20la%20industria%20del,los%20680.000%20millones%20en%202030 [accessed 2024-09-15]
4. La industria del videojuego en España en 2022. The video game industry in Spain in 2022. Asociación Española de Videojuegos. 2024. URL: https://www.aevi.org.es/web/wp-content/uploads/2024/01/AEVI_INFOGRAPHIC-22-ENGLISH.pdf [accessed 2020-01-18]
5. Kristiansen S, Severin MC. Loot box engagement and problem gambling among adolescent gamers: findings from a national survey. Addict Behav 2020 Apr;103:106254. [doi: 10.1016/j.addbeh.2019.106254] [Medline: 31881408]
6. Li W, Mills D, Nower L. The relationship of loot box purchases to problem video gaming and problem gambling. Addict Behav 2019 Oct;97:27-34. [doi: 10.1016/j.addbeh.2019.05.016] [Medline: 31129456]
7. Observatorio Español de las Drogas y las Adicciones. Informe sobre: Adicciones comportamentales y otros trastornos adictivos 2023. Portal Plan Nacional sobre Drogas. 2023. URL: https://pnsd.sanidad.gob.es/profesionales/sistemasInformacion/sistemaInformacion/pdf/2023_Informe_Trastornos_Comportamentales.pdf [accessed 2024-09-13]
8. King DL, Delfabbro PH. Loot box limit-setting is not sufficient on its own to prevent players from overspending: a reply to Drummond, Sauer & Hall. Addiction 2019 Jul;114(7):1324-1325. [doi: 10.1111/add.14628] [Medline: 30955212]
9. Montiel I, Basterra-González A, Machimbarrena JM, Ortega-Barón J, González-Cabrera J. Loot box engagement: a scoping review of primary studies on prevalence and association with problematic gaming and gambling. PLoS One 2022;17(1):e0263177. [doi: 10.1371/journal.pone.0263177] [Medline: 35085370]
10. Drummond A, Sauer JD, Hall LC, Zendle D, Loudon MR. Author correction: why loot boxes could be regulated as gambling. Nat Hum Behav 2020 Aug 11;5(6):802-802. [doi: 10.1038/s41562-020-00933-5]
11. Zendle D, Cairns P. Video game loot boxes are linked to problem gambling: results of a large-scale survey. PLoS One 2018;13(11):e0206767. [doi: 10.1371/journal.pone.0206767] [Medline: 30462669]
12. Lemmens JS. Play or pay to win: loot boxes and gaming disorder in FIFA Ultimate Team. Tele Inform Rep 2022 Dec;8:100023. [doi: 10.1016/j.teler.2022.100023]
13. Xiao LY, Henderson LL, Newall PWS. Loot boxes are more prevalent in United Kingdom video games than previously considered: updating Zendle et al. (2020). Addiction 2022 Sep;117(9):2553-2555. [doi: 10.1111/add.15829] [Medline: 35129243]
14. Evren C, Evren B, Dalbudak E, Topcu M, Kutlu N. The relationship of loot box engagement to gender, severity of disordered gaming, using MMORPGs, and motives for online gaming. PBS 2021;11(1):25. [doi: 10.5455/PBS.20200915101516]
15. Zendle D, Meyer R, Cairns P, Waters S, Ballou N. The prevalence of loot boxes in mobile and desktop games. Addiction 2020 Sep;115(9):1768-1772. [doi: 10.1111/add.14973] [Medline: 31957942]
16. González-Cabrera J, Basterra-González A, Ortega-Barón J, et al. Loot box purchases and their relationship with internet gaming disorder and online gambling disorder in adolescents: a prospective study. Comput Human Behav 2023 Jun;143:107685. [doi: 10.1016/j.chb.2023.107685]
17. Brooks GA, Clark L. Associations between loot box use, problematic gaming and gambling, and gambling-related cognitions. Addict Behav 2019 Sep;96:26-34. [doi: 10.1016/j.addbeh.2019.04.009] [Medline: 31030176]
18. Close J, Spicer SG, Nicklin LL, Uther M, Lloyd J, Lloyd H. Secondary analysis of loot box data: are high-spending “whales” wealthy gamers or problem gamblers? Addict Behav 2021 Jun;117:106851. [doi: 10.1016/j.addbeh.2021.106851] [Medline: 33578105]
19. Garea SS, Drummond A, Sauer JD, Hall LC, Williams MN. Meta-analysis of the relationship between problem gambling, excessive gaming and loot box spending. Int Gambl Stud 2021 Sep 2;21(3):460-479. [doi: 10.1080/14459795.2021.1914705]

20. Macey J, Hamari J. eSports, skins and loot boxes: participants, practices and problematic behaviour associated with emergent forms of gambling. New Media Soc 2019 Jan;21(1):20-41. [doi: 10.1177/1461444818786216]
21. von Meduna M, Steinmetz F, Ante L, Reynolds J, Fiedler I. Loot boxes are gambling-like elements in video games with harmful potential: results from a large-scale population survey. Technol Soc 2020 Nov;63:101395. [doi: 10.1016/j.techsoc.2020.101395]
22. Wardle H, Zendle D. Loot boxes, gambling, and problem gambling among young people: results from a cross-sectional online survey. Cyberpsychol Behav Soc Netw 2021 Apr;24(4):267-274. [doi: 10.1089/cyber.2020.0299] [Medline: 33103911]
23. Zendle D, Meyer R, Over H. Adolescents and loot boxes: links with problem gambling and motivations for purchase. R Soc Open Sci 2019 Jun;6(6):190049. [doi: 10.1098/rsos.190049] [Medline: 31312481]
24. Zendle D, Cairns P. Correction: video game loot boxes are linked to problem gambling: results of a large-scale survey. PLoS One 2019;14(3):e0214167. [doi: 10.1371/journal.pone.0214167] [Medline: 30870502]
25. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5): American Psychiatric Publishing; 2013.
26. World Health Organization. International Classification of Diseases and Related Health Problems (11th Revision): World Health Organization; 2018.
27. González-Cabrera J, Basterra-González A, Montiel I, Calvete E, Pontes HM, Machimbarrena JM. Loot boxes in Spanish adolescents and young adults: relationship with internet gaming disorder and online gambling disorder. Comput Human Behav 2022 Jan;126:107012. [doi: 10.1016/j.chb.2021.107012]
28. Brooks GA, Clark L. The gamblers of the future? migration from loot boxes to gambling in a longitudinal study of young adults. Comput Human Behav 2023 Apr;141:107605. [doi: 10.1016/j.chb.2022.107605]
29. Spicer SG, Nicklin LL, Uther M, Lloyd J, Lloyd H, Close J. Loot boxes, problem gambling and problem video gaming: a systematic review and meta-synthesis. New Media Soc 2022 Apr;24(4):1001-1022. [doi: 10.1177/14614448211027175]
30. Gibson E, Griffiths MD, Calado F, Harris A. The relationship between videogame micro-transactions and problem gaming and gambling: a systematic review. Comput Human Behav 2022 Jun;131:107219. [doi: 10.1016/j.chb.2022.107219]
31. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR): American Psychiatric Association Publishing; 2022. [doi: 10.1176/appi.books.9780890425787]
32. King A, Wong-Padoongpatt G, Barrita A, Phung DT, Tong T. Risk factors of problem gaming and gambling in US emerging adult non-students: the role of loot boxes, microtransactions, and risk-taking. Issues Ment Health Nurs 2020 Dec;41(12):1063-1075. [doi: 10.1080/01612840.2020.1803461] [Medline: 32822251]
33. Beranuy M, Machimbarrena JM, Vega-Osés MA, et al. Spanish validation of the Internet Gaming Disorder Scale-Short Form (IGDS9-SF): prevalence and relationship with online gambling and quality of life. Int J Environ Res Public Health 2020 Feb 28;17(5):1562. [doi: 10.3390/ijerph17051562] [Medline: 32121280]
34. Pontes HM, Király O, Demetrovics Z, Griffiths MD. The conceptualisation and measurement of DSM-5 internet gaming disorder: the development of the IGD-20 Test. PLoS One 2014;9(10):e110137. [doi: 10.1371/journal.pone.0110137]
35. Pontes HM, Griffiths MD. Measuring DSM-5 internet gaming disorder: development and validation of a short psychometric scale. Comput Human Behav 2015 Apr;45:137-143. [doi: 10.1016/j.chb.2014.12.006]
36. González-Cabrera J, Machimbarrena JM, Beranuy M, Pérez-Rodríguez P, Fernández-González L, Calvete E. Design and measurement properties of the Online Gambling Disorder Questionnaire (OGD-Q) in Spanish adolescents. J Clin Med 2020 Jan 2;9(1):120. [doi: 10.3390/jcm9010120] [Medline: 31906512]
37. Niessen ASM, Meijer RR, Tendeiro JN. Detecting careless respondents in web-based questionnaires: which method to use? J Res Pers 2016 Aug;63:1-11. [doi: 10.1016/j.jrp.2016.04.010]
38. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 1980;88(3):588-606. [doi: 10.1037/0033-2909.88.3.588]
39. Muthén LK, Muthén BO. Mplus User's Guide, 8th edition: Muthén & Muthén; 1998.
40. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
41. Raykov T. Bias-corrected estimation of noncentrality parameters of covariance structure models. Struct Equ Modeling 2005 Jan;12(1):120-129. [doi: 10.1207/s15328007sem1201_6]
42. Xiao LY. Breaking ban: Belgium's ineffective gambling law regulation of video game loot boxes. Psychology (Irvine) 2023 Jan 17;9(1):57641. [doi: 10.1525/collabra.57641]
43. Garrett EP, Drummond A, Lowe-Calverley E, Sauer JD. Current loot box warnings are ineffective for informing consumers. Comput Human Behav 2023 Feb;139:107534. [doi: 10.1016/j.chb.2022.107534]
44. Xiao LY. Beneath the label: unsatisfactory compliance with ESRB, PEGI and IARC industry self-regulation requiring loot box presence warning labels by video game companies. R Soc Open Sci 2023 Mar;10(3):230270. [doi: 10.1098/rsos.230270] [Medline: 36998766]
45. ICD-11 for mortality and morbidity statistics. World Health Organization. 2018. URL: https://icd.who.int/browse11/l-m/en [accessed 2020-01-18]
46. Ministerio de Consumo. Anteproyecto de Ley por el que se regulan los mecanismos aleatorios de recompensa asociados a productos de software interactivo de ocio. Draft bill regulating random reward mechanisms associated with interactive

leisure software products. Ministerio de Derechos Sociales, Consumo y Agenda 2030. 2022. URL: https://www.consumo.gob.es/sites/consumo.gob.es/files/BORRADOR%20APL%20Y%20MAIN%20MECANISMOS%20ALEATORIOS%20RECOMPENSA%20010722.pdf [accessed 2024-09-05]

Abbreviations

CFA: confirmatory factor analysis
CFI: comparative fit index
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
***ICD-11*:** *International Classification of Diseases, 11th Revision*
IGD: internet gaming disorder
IGD9-SF: Internet Gaming Disorder Scale–Short Form
LB: loot box
OGD: online gambling disorder
OGD-Q: Online Gambling Disorder Questionnaire
PEGI: Pan-European Gaming Information
PU-LB: problematic use of loot boxes
RLI: Risky Loot-Box Index
RMSEA: root mean square error of approximation
S-Bχ^2**:** Satorra-Bentler chi-square
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by A Mavragani; submitted 12.02.24; peer-reviewed by A Sirola, L Xiao, L Monacis; revised version received 17.06.24; accepted 17.06.24; published 20.09.24.

Please cite as:
González-Cabrera J, Caba-Machado V, Díaz-López A, Jiménez-Murcia S, Mestre-Bach G, Machimbarrena JM
The Mediating Role of Problematic Use of Loot Boxes Between Internet Gaming Disorder and Online Gambling Disorder: Cross-Sectional Analytical Study
JMIR Serious Games 2024;12:e57304
URL: https://games.jmir.org/2024/1/e57304
doi:10.2196/57304

Original Paper

Association Between Internet Gaming Disorder and Suicidal Ideation Mediated by Psychosocial Resources and Psychosocial Problems Among Adolescent Internet Gamers in China: Cross-Sectional Study

Yanqiu Yu[1], PhD; Anise M S Wu[2,3], PhD; Vivian W I Fong[4], MSc; Jianxin Zhang[5], MD; Ji-bin Li[6], PhD; Joseph T F Lau[7,8], PhD

[1]Department of Preventive Medicine and Health Education, School of Public Health, Fudan University, Shanghai, China

[2]Department of Psychology, Faculty of Social Sciences, University of Macau, Macao, China

[3]Centre for Cognitive and Brain Sciences, Institute of Collaborative Innovation, University of Macau, Macao, China

[4]Center for Health Behaviours Research, Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China

[5]West China School of Public Health, Sichuan University, Chengdu, China

[6]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Department of Clinical Research, Sun Yat-sen University Cancer Center, Guangzhou, China

[7]Public Mental Health Center, School of Mental Health, Wenzhou Medical University, Wenzhou, China

[8]Zhejiang Provincial Clinical Research Center for Mental Disorders, The Affiliated Wenzhou Kangning Hospital, Wenzhou Medical University, Wenzhou, China

Corresponding Author:
Joseph T F Lau, PhD
Public Mental Health Center
School of Mental Health
Wenzhou Medical University
Ouhai District
Wenzhou, 325000
China
Phone: 86 57786689810
Email: jlau@cuhk.edu.hk

Abstract

Background: Adolescent internet gaming disorder (IGD) was associated with severe harm, including suicidal ideation. While suicidal ideation was predictive of completed suicides, further research is required to clarify the association between IGD and suicidal ideation among adolescents, as well as the mechanisms involved.

Objective: This study aimed to investigate the understudied association between IGD and suicidal ideation, as well as novel mechanisms associated with it, among Chinese adolescent internet gamers through psychosocial coping resources and psychosocial problems.

Methods: An anonymous, self-administered, cross-sectional survey was conducted among secondary school students who had played internet games in the past year in Guangzhou and Chengdu, China (from October 2019 to January 2020). In total, 1693 adolescent internet gamers were included in this study; the mean age was 13.48 (SD 0.80) years, and 60% (n=1016) were males. IGD was assessed by the 9-item Internet Gaming Disorder Checklist of the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]), while a single item assessed suicidal ideation: “Have you ever considered committing suicide in the past 12 months?” Univariate and multivariate logistic regression associations were conducted to test the significance and directions of the potential factors for suicidal ideation. The mediation mechanism was examined by structural equation modeling.

Results: Among all participants, the prevalence of IGD and suicidal ideation was 16.95% (287/1693) and 43.06% (729/1693), respectively. IGD cases were 2.42 times more likely than non-IGD cases to report suicidal ideation (adjusted odds ratio [OR] 2.42, 95% CI 1.73-3.37). Other significant factors of suicidal ideation included psychosocial coping resources (resilience and social support, both adjusted OR 0.97, 95% CI 0.96-0.98) and psychosocial problems (social anxiety: adjusted OR 1.07, 95% CI

1.05-1.09; loneliness, adjusted OR 1.13, 95% CI 1.10-1.16). The association between IGD and suicidal ideation was partially mediated by 3 indirect paths, including (1) the 2-step path that IGD reduced psychosocial coping resources, which in turn increased suicidal ideation; (2) the 2-step path that IGD increased psychosocial problems, which in turn increased suicidal ideation; and (3) the 3-step path that IGD reduced psychosocial coping resources which then increased psychosocial problems, which in turn increased suicidal ideation, with effect sizes of 10.7% (indirect effect/total effect: 0.016/0.15), 30.0% (0.05/0.15), and 13.3% (0.02/0.15), respectively. The direct path remained statistically significant.

Conclusions: IGD and suicidal ideation were alarmingly prevalent. Evidently and importantly, IGD was a significant risk factor for suicidal ideation. The association was partially explained by psychosocial coping resources of resilience and social support and psychosocial problems of social anxiety and loneliness. Longitudinal studies are needed to confirm the findings. Pilot randomized controlled trials are recommended to evaluate the effectiveness of interventions in reducing suicidal ideation by reducing IGD, improving psychosocial coping resources, and reducing psychosocial problems investigated in this study.

(JMIR Serious Games 2024;12:e48439) doi:10.2196/48439

KEYWORDS

internet gaming disorder; suicidal ideation; adolescents; mediation; structural equation modelling; resilience; loneliness; social support; social anxiety

Introduction

Adolescent suicide is the fourth leading cause of death among adolescents aged 15-19 years [1]. A review including studies conducted in 15 countries reported age-standardized suicide rates among male and female youths aged 15-29 years ranging from 2.4 to 51 per 100,000 and from 1.1 to 16.4 per 100,000, respectively [2]. A national report showed the prevalence of suicide ranging from 8 to 11.8 per 100,000 among adolescents aged 15-20 years in China [3,4]. The spectrum of suicide includes conditions ranging from suicidal ideation (having thoughts of ending one's own life), suicidal intent (having the specific intention of ending one's own life, often including planning or preparation), suicidal attempt (having attempted to end one's own life but not resulting in death), to death [5]. Suicidal ideation increases the risk of suicidal behaviors [6]. A better understanding of the risk factors of adolescent suicidal ideation and related mechanisms is warranted.

Addictive behaviors (eg, substance use and internet addiction) are well-documented risk factors for suicidal ideation [7,8]. Internet gaming disorder (IGD), which is a subtype of gaming disorder, has been included as a mental disorder in the *ICD-11* (*International Classification of Diseases* [*11th Revision*]) published by the World Health Organization in 2018 [9]. It was associated with harmful behavioral and psychological consequences, including suicidal ideation [10]. A review of 12 cross-sectional studies found moderate-to-strong positive associations between problematic gaming and suicidal ideation [11]. However, only 1 study targeted Chinese adolescents [12]. In addition, only 1 longitudinal study conducted among Swedish adults looked at such a relationship and found that behavioral addiction (including IGD) at age 25 years significantly predicted suicidal ideation at age 28 years [13]. Understanding the mechanisms between IGD and suicidal ideation would guide the development of effective interventions. To our knowledge, however, only 1 study conducted in China had investigated 1 such potential mechanism (ie, IGD was associated with increased insomnia symptoms that led to a higher risk of depression, which in turn elevated the risk of suicidal ideation) among adolescents [12]. Research is greatly warranted to fill this knowledge gap.

The psychosocial problems of social anxiety and loneliness are potential mediators between IGD and suicidal ideation. Loneliness refers to distressing feelings that arise when there is a lack of intimate or satisfactory social connections [14]. In contrast, social anxiety refers to the perceived fear of being watched and judged by others in social situations [15]. Loneliness increases the risk of social anxiety and vice versa [16]. Furthermore, they often come together and are positively associated with mental problems, including thoughts of self-harm and suicide [17-19]. In addition, there is further support for the above-proposed mediation. First, pathological internet use (including IGD) has been postulated to cause deficient face-to-face social interactions that would lead to social isolation [20,21]. Longitudinal studies have also reported that adolescent IGD predicted social anxiety and loneliness [22,23]. Second, according to the Interpersonal-Psychological Theory of Suicide, suicidal thoughts are determined by perceived burdensomeness to others or society and thwarted belongingness [24], which is potentially associated with social anxiety and loneliness [24,25].

It is plausible that psychosocial coping resources (eg, resilience and social support) will further mediate the association between IGD and psychosocial problems (eg, social anxiety and loneliness). In that case, a serial mediation between IGD and suicidal ideation, first by psychosocial coping resources and then by social anxiety and loneliness, would occur. Regarding psychosocial coping resources, resilience is a personal resource referring to the ability to withstand and bounce back from difficult life events [26]. Social support is an interpersonal resource referring to receiving assistance or comfort from interpersonal relationships within an individual's social network [27,28]. A couple of theories support the proposed mediation between IGD and psychosocial problems through psychosocial coping resources. First, potential functional impairments of IGD (eg, problems in personal life, social relationships, and academic performance) can be seen as strong stressors [29]. According to the Resource Deterioration Model, the presence of such stressors (functional impairments of IGD in this case) would diminish coping resources (resilience and social support in this case) [30,31]. Second, the Conversation of Resource theory postulates that losses in personal and interpersonal resources (those resulting from IGD in this case) would cause psychosocial problems (social anxiety and loneliness in this case) and suicidal ideation [32-35]. Thus, this study contended that IGD would reduce both personal coping resources (resilience) and interpersonal coping resources (social support), which would then increase social anxiety and loneliness, which would, in turn, increase suicidal ideation.

The proposed mechanisms of this study have further empirical support. First, a study conducted in Taiwan reported that the IGD group showed a higher level of stress and a lower level of resilience than the non-IGD group [36], while a longitudinal study found that IGD predicted a lower level of social support among Chinese university students [37]. Second, resilience and social support were negatively associated with social anxiety [38,39], loneliness [40,41], and suicidal ideation [42,43]. Third, social anxiety and loneliness were significantly associated with suicidal ideation [17,44]. Furthermore, extant literature has reported significant mediations between psychosocial resources and mental disorders through psychosocial problems. For instance, loneliness significantly mediated the association between perceived social support and depression among Chinese rural-to-urban migrants [45].

Given the background, this study investigated the prevalence of IGD and suicidal ideation among adolescent internet gamers in 2 Chinese cities. Factors of suicidal ideation were investigated, including (1) IGD, (2) a total of 2 types of psychosocial coping resources (resilience and social support), and (3) a total of 2 types of psychosocial problems (social anxiety and loneliness). The tested mediation model contains 3 mediation paths (indirect effects) between IGD and suicidal ideation that are (1) a 2-step path, in which IGD would reduce psychosocial resources (resilience and social support), which would, in turn, increase suicidal ideation; (2) a 2-step path, where IGD would increase psychosocial problems (social anxiety and loneliness), which would, in turn, increase suicidal ideation; and (3) a 3-step serial mediation path postulating that IGD would reduce psychosocial resources (resilience and social support), which would then increase psychosocial problems (social anxiety and loneliness), which would, in turn, increase suicidal ideation. In addition, the direct path from IGD to suicidal ideation was tested.

Methods

Participants and Data Collection

A cross-sectional survey was conducted among junior middle school students in 2 metropolises (Guangzhou and Chengdu) in China from October 2019 to January 2020. Guangzhou and Chengdu had population sizes of 15.3 and 16.3 million in 2019, respectively. All eighth-grade students from 4 Guangzhou schools and all seventh to ninth grade students from 3 Chengdu schools were conveniently selected and invited to participate in this survey. Notably, the Guangzhou sample has been used in 2 previous publications whose topics were completely different from this study; one investigated the impacts of the medicalization of IGD [46], and the other validated an assessment tool for potential resource losses related to gaming time reduction [47]. The data collection procedure was, hence, the same as those in these published studies and was briefly introduced here.

In the absence of teachers, students self-administered an anonymous, structured questionnaire using paper and pencil in the classroom. Well-trained fieldworkers briefed the students about the anonymous and voluntary nature and logistics of the study, that there would be no consequences of refusing or quitting this survey at any time if wished, and that the return of the completed questionnaires would imply informed consent. The questionnaire took 30 to 40 minutes to complete. The fieldworkers also assisted with inquiries from the students if needed and did the quality check when the students submitted the questionnaire.

There were 3039 completed questionnaires (a response rate of 99.09%, 3039/3067), among which 74 participants (2.44%) were removed from data analysis due to missing data in ≥20% of the questionnaire items (n=60, 1.97%) or any missing data in the key variables (eg, IGD and suicidal ideation; n=14, 0.46%). Out of the 2965 valid responses, 1272 (42.90%) participants had not played internet games in the past 12 months (nongamers) and were excluded from data analysis. The effective sample size of this study was hence 1693 (Guangzhou: 55.46%, n=939; Chengdu: 44.54%, n=754).

Ethical Considerations

This study involved informed consent from 3 parties. First, school consent was obtained from school principals before the investigation. Second, parents were informed about the survey, and the opt-out procedure was exercised. Third, students were briefed by the fieldworkers about the anonymous and voluntary nature of this study and that the submission of a completed questionnaire would indicate informed consent, such information was also printed on the cover page of the questionnaire; no written informed consent was obtained to avoid disclosure of personal identity. No incentive was provided to school coordinators, school staff, parents, and students. This project and the above-informed consent procedures were approved by the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong in 2019 (SBRE-18-430).

Measurements

Sociodemographics

Information was collected about age (years), sex, whether living with both parents or a single-parent family, and self-rated household financial situations relative to the participant's classmates (5 points: very poor, poor, moderate, good, or very good). Notably, the relative financial situation had much greater predictive validity for self-reported health and well-being than the absolute financial situation [48].

Suicidal Ideation

Furthermore, 1 item assessed, "Have you ever considered committing suicide in the past 12 months?" (yes or no were the response options). The single-item assessment of suicidal ideation has shown predictive validity for passive and active suicidal ideation as well as depression [49]. In addition, previous studies have used the same or similar single items to assess suicidal ideation [12,50,51].

IGD

It was assessed by the 9-item IGD Checklist of the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]). IGD was defined as the endorsement of at least 5 of the 9 types of *DSM-5* symptoms (preoccupation, withdrawal, tolerance, inability to control internet gaming, prioritization over other activities, continuation of internet gaming despite adverse consequences, deception of internet gaming time, avoidance, and significant loss due to internet gaming) in the past 12 months (yes or no response options) [52]. The Chinese version of the checklist has been validated among Chinese adolescents and yielded satisfactory psychometric properties [53]. The Cronbach α of the scale was 0.79 in this study.

Resilience

It was assessed by the 10-item Connor-Davidson Resilience Scale [26]; its Chinese version has been validated among adolescents and showed satisfactory psychometric properties [54]. A sample item is “I can adapt to change.” The items were rated on a 5-point Likert scale (0=never to 4=always). The Cronbach α of the scale was 0.93 in this study.

Social Support

It was assessed by 2 subscales of the Multidimensional Scale of Perceived Social Support, which assessed perceived social support from family members and friends [27]. Its Chinese version has been validated among adolescents in China and showed acceptable psychometric properties [55]. A sample item is “My family members/friends really try to help me.” The items were rated on a 7-point Likert scale (1=extremely disagree to 7=extremely agree). The Cronbach α of the scale was 0.93 in this study.

Social Anxiety

It was assessed by using the 9-item social anxiety subscale of the Multidimensional Anxiety Scale for Children [56]; the Chinese version has been validated among adolescents in China and showed good psychometric properties [57]. A sample item is “I’m afraid other people will think I’m stupid.” The items were rated on a 4-point Likert scale (0=disagree to 3=always agree). The Cronbach α of the scale was 0.92 in this study.

Loneliness

It was assessed by using the 8-item short-form of the UCLA (University of California, Los Angeles) Loneliness Scale; the Chinese version has been validated among adolescents in China, which showed acceptable psychometric properties [58]. A sample item is “I feel isolation from others.” The items were rated on a 4-point Likert scale (0=never to 3=always); higher scores indicated higher levels of loneliness. The Cronbach α of the scale was 0.78 in this study.

Probable Moderate or Above Depression

It was assessed by the 9-item Patient Health Questionnaire (PHQ-9), which is a multipurpose instrument for screening, diagnosing, and monitoring the severity of depression. Its Chinese version has been validated in adolescents and showed good psychometric properties [59,60]. A sample item is “Little interest or pleasure in doing things.” The items were rated on a 4-point Likert scale on the frequency of having potential symptoms of depression in the past 2 weeks (0=not at all to 3=nearly every day). Probable moderate or above depression was defined as a PHQ-9 score ≥10 in this study; it has been used in several previous studies [61,62]. The Cronbach α of the scale was 0.93 in this study.

Statistical Analysis

IGD status and probable moderate or above depression were used in binary forms. Mean (SD, range) and frequency (proportion) were used to statistically describe continuous and categorical variables, respectively. The Chi-square (χ^2) test and independent-sample *t* test were performed to examine between-group differences regarding categorical and continuous variables, respectively. Pearson correlation coefficients (*r*) were generated to assess correlations between variables. Univariate and multivariate logistic regression analyses (adjusted for sociodemographics and probable moderate or above depression) were conducted to test the significant factors (IGD and psychosocial factors) for suicidal ideation. Unadjusted odds ratios (ORs) and adjusted ORs and their respective 95% CIs were presented.

Structural equation modeling (SEM) was fit to test the underlying mechanisms between IGD and suicidal ideation through the 4 psychosocial variables. The estimator of weighted least square mean and variance was used in SEM. A total of 2 latent variables were created: (1) the latent variable of psychosocial resources was generated from the scale scores of both resilience and social support, and (2) the latent variable of psychosocial problems was generated from the scale scores of both social anxiety and loneliness. The goodness-of-fit of the SEM model was evaluated by comparative fit index (CFI) ≥0.90 and root-mean-square error of approximation (RMSEA) ≤0.08 [63,64]. Standardized path coefficients (β) were reported.

SPSS Statistics (version 23.0; IBM Corp) and Mplus (version 7.0; Muthen & Muthen) were used for statistical analyses; *P*<.05 was considered statistically significant.

Results

Descriptive Statistics: Sociodemographics

Among all participants, the mean age was 13.48 (SD 0.80; range 10-19) years. Over half of them were male (60%, 1016/1693); about one-fifth did not live with both parents (20.08%, 340/1693) or lived in a single-parent family (15.95%, 270/1693); 12.52% (212/1693) considered their household financial situation poorer or much poorer relative to their classmates (Table 1).

Table 1. Descriptive statistics.

	Overall (N=1693), n (%)	IGD[a]		
		Yes (n=287, 16.95%), n (%)	No (n=1406, 83.05%), n (%)	*P* of χ^2
Studied sites				<.001
Guangzhou	939 (55.46)	102 (35.54)	837 (59.53)	
Chengdu	754 (44.54)	185 (64.46)	569 (40.47)	
Sex				<.001
Female	668 (39.46)	80 (27.87)	588 (41.82)	
Male	1016 (60.0)	203 (70.73)	813 (57.82)	
Missing data	9 (0.54)	4 (1.39)	5 (0.36)	
Living with both parents				.001
Yes	1342 (79.27)	205 (71.43)	1137 (80.87)	
No	340 (20.08)	78 (27.18)	262 (18.63)	
Missing data	11 (0.65)	4 (1.39)	7 (0.50)	
Single-parent family				0.17
No	1410 (83.28)	229 (79.79)	1181 (84.00)	
Yes	270 (15.95)	53 (18.47)	217 (15.43)	
Missing data	13 (0.77)	5 (1.74)	8 (0.57)	
Self-rated household financial situation relative to classmates				<.001
Good or very good	437 (25.81)	78 (27.18)	359 (25.53)	
Average	1028 (60.72)	147 (51.22)	881 (62.66)	
Poor or very poor	212 (12.52)	57 (19.86)	155 (11.02)	
Missing data	16 (0.95)	5 (1.74)	11 (0.78)	
Probable moderate or above depression[b]				<.001
No	1172 (69.23)	115 (40.07)	1056 (75.11)	
Yes	521 (30.77)	172 (59.93)	349 (24.82)	
Suicidal ideation				<.001
No	964 (56.94)	90 (31.36)	874 (62.16)	
Yes	729 (43.06)	197 (68.64)	532 (37.84)	

[a]IGD: internet gaming disorder; IGD was defined as the endorsement of 5 or more out of the 9-item IGD checklist of the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]).

[b]Probable moderate or above depression was defined as the scale score of the Patient Health Questionnaire-9 score of ≥10.

Prevalence of IGD, Probable Moderate or Above Depression, and Suicidal Ideation

The results are shown in Table 1. Over two-fifths (43.06%, 729/1693) reported suicidal ideation in the past year, and 16.95% (287/1693) were classified as IGD cases. Compared with their non-IGD counterparts (mean 13.46, SD 0.79), those with IGD (mean 13.60, SD 0.81) were more likely to be older (*P*=.01), belonging to the Chengdu sample (185/287, 64.46% vs 569/1406, 40.47%), and male (203/287, 70.73% vs 813/1406, 57.82%), and self-reported poor or very poor relative household financial situation (57/287, 19.86% vs 155/1406, 11.02%). The prevalence of probable moderate or above depression (PHQ-9 score≥10) was 30.77% (521/1693); the mean score of PHQ-9 was 7.29 (SD 6.82, range 0-27). IGD was significantly and positively associated with both probable moderate or above depression (172/287, 59.93% vs 349/1406, 24.82%; *P*<.001) and suicidal ideation (197/287, 68.64% vs 532/1406, 37.84%; *P*<.001).

Levels of Psychosocial Resources and Psychosocial Problems

The mean scores of the resilience, social support, social anxiety, and loneliness scales were 23.07 (SD 8.73, range 0-40), 38.30 (SD 11.73, range 8-56), 12.70 (SD 7.14, range 0-27), and 15.78 (SD 5.04, range 8-32), respectively. The IGD group had significantly lower levels of psychosocial resources of resilience and social support and higher levels of psychosocial problems of social anxiety and loneliness than their counterparts (moderate to strong effect sizes; Cohen *d*=0.49-0.79; Table 2).

Table 2. Between-group differences in psychosocial resources and problems by IGD status (N=1693).

	Range	Overall, mean (SD)	IGD[a], mean (SD)	Non-IGD, mean (SD)	*P* value	Cohen *d*
Psychosocial resources						
Resilience	0-40	23.07 (8.74)	19.68 (7.84)	23.76 (8.75)	<.001	0.49
Social support	8-56	38.30 (11.73)	33.18 (11.63)	39.34 (11.48)	<.001	0.53
Psychosocial problems						
Social anxiety	0-27	12.70 (7.14)	15.80 (6.35)	12.08 (7.13)	<.001	0.55
Loneliness	8-32	15.78 (5.04)	18.85 (4.44)	15.16 (4.93)	<.001	0.79

[a]IGD: internet gaming disorder.

Correlations

[Table 3](#) shows the significant positive correlations between resilience and social support (psychosocial coping resources; *r*=0.50) and between loneliness and social anxiety (psychosocial problems; *r*=0.48). The 2 psychosocial resource variables were negatively correlated with the 2 psychosocial problems (*r*=–0.40 to –0.13).

Table 3. Correlation analysis (n=1693).

	Resilience	Social support	Social anxiety
Psychosocial resources			
Resilience	—[a]	—	—
Social support	0.50[b]	—	—
Psychosocial problems			
Social anxiety	−0.19[b]	−0.13[b]	—
Loneliness	−0.36[b]	−0.40[b]	0.48[b]

[a]Not applicable.

[b]*P*<.001.

Factors of Suicidal Ideation

Similar to the results of the univariate logistic regression analysis, multivariate logistic regression adjusting for the sociodemographics and probable moderate or above depression reported that increases in psychosocial coping resources of resilience (adjusted OR 0.97, 95% CI 0.96-0.98) and social support (adjusted OR 0.97, 95% CI 0.96-0.98) were associated with a lower risk of suicidal ideation while increases in IGD (adjusted OR 2.42, 95% CI 1.73-3.37) and psychosocial problems of social anxiety (adjusted OR 1.07, 95% CI 1.05-1.09) and loneliness (adjusted OR 1.13, 95% CI 1.10-1.16) were associated with a higher risk of suicidal ideation (Table 4).

Table 4. Univariate and multivariate logistic regression analysis on the individual associations between independent variables and suicidal ideation (n=1693). There are 10 individual models in total, 5 for univariate logistic regression and the other 5 for multivariate logistic regression.

	Unadjusted OR (95% CI)	*P* value	Adjusted OR[a] (95% CI)	*P* value
Psychosocial resources				
Resilience	0.95 (0.93-0.96)	<.001	0.97 (0.96-0.98)	<.001
Social support	0.95 (0.94-0.96)	<.001	0.97 (0.96-0.98)	<.001
Psychosocial problems				
Social anxiety	1.10 (1.09-1.12)	<.001	1.07 (1.05-1.09)	<.001
Loneliness	1.21 (1.18-1.24)	<.001	1.13 (1.10-1.16)	<.001
Internet gaming disorder				
No	Reference=1.0	—[b]	Reference=1.0	—
Yes	3.60 (2.74-4.72)	<.001	2.42 (1.73-3.37)	<.001

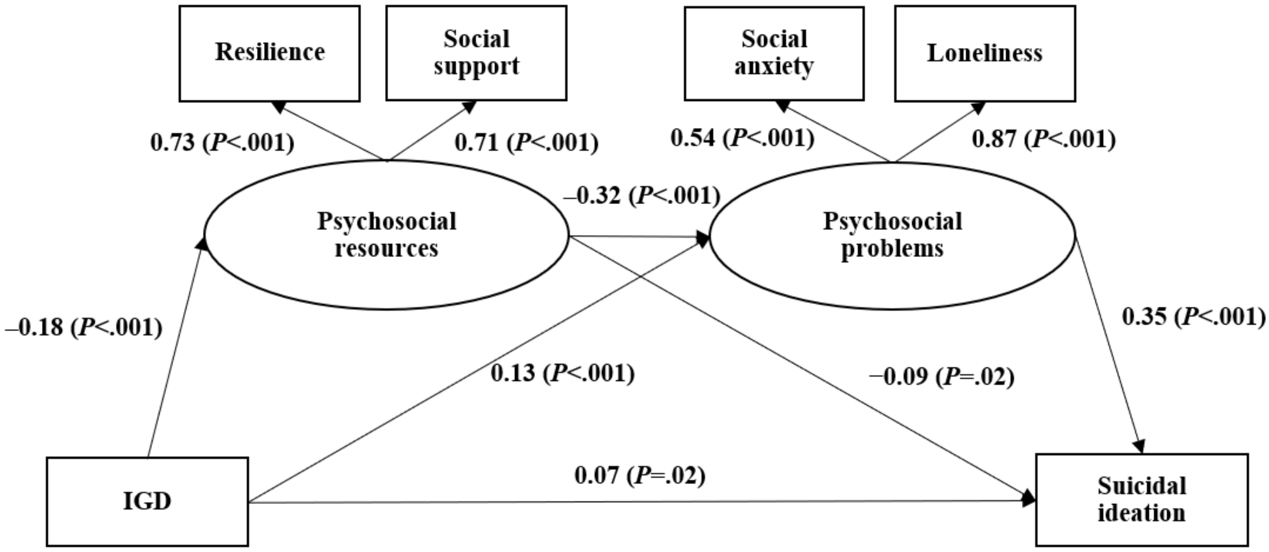
[a]The adjusted models were adjusted for background factors (ie, age, studied city, sex, living arrangement, single-parent family, and household financial situation relative to classmates) and probable moderate or above depression (Patient Health Questionnaire-9 score≥10).

[b]Not applicable.

Structural Equation Modeling

Figure 1 presents the results of SEM testing the proposed mediation effects of psychosocial coping resources and psychosocial problems between IGD and suicidal ideation. The model showed satisfactory goodness-of-fit (χ^2_{21}=176.21, *P*<.001; CFI=0.92; RMSEA=0.07), with factor loadings of the 2 latent variables ranging from 0.54 to 0.87 (all *P*<.001). The results showed that IGD had a significant direct effect on suicidal ideation (β=.07, *P*=.02). In addition, IGD was linked to suicidal ideation through 3 indirect paths: (1) a 2-step mediation path involving a negative association between IGD and the latent variable of psychosocial coping resources (β=–.18, *P*<.001), which was negatively associated with suicidal ideation (β=–.09, *P*=.02); (2) a 2-step mediation path involving a positive association between IGD and the latent variable of psychosocial problems (β=.13, *P*<.001), which was positively associated with suicidal ideation (β=.35, *P*<.001); and (3) a 3-step serial path involving a negative association between IGD and the latent variable of psychosocial coping resources (β=–.18, *P*<.001) and then a negative association between psychosocial coping resources and psychosocial problems (β=–.32, *P*<.001), which was in turn positively associated with suicidal ideation (β=.35, *P*<.001). Partial mediations were found as the direct path from IGD to suicidal ideation was of statistical significance. The corresponding mediation effect size of the above 3 indirect paths were 10.7% (indirect effect/total effect: 0.016/0.15), 30.0% (0.05/0.15), and 13.3% (β=.02/0.15, *P*<.001), respectively.

Figure 1. Structural equation modeling testing the mediation effect of psychosocial coping resources and psychosocial problems between internet gaming disorder (IGD) and suicidal ideation (standard coefficients were reported).



Discussion

Principal Findings

This survey was conducted among Chinese adolescents and found that IGD increased the risk of suicidal ideation. Both protective factors of psychosocial resources (resilience and social support) and risk factors of psychosocial problems (social anxiety and loneliness) for suicidal ideation were also identified. Furthermore, this study was novel to reveal the serial mediation mechanisms through psychosocial resources and psychosocial problems between IGD and suicidal ideation, which provide empirical evidence for future interventions targeting suicidal ideation as a negative consequence of IGD.

Comparison With Previous Work

This study observed the alarmingly high prevalence of suicidal ideation of 43.06% (729/1693) among junior middle school students in 2 Chinese cities, which was higher than that of studies previously conducted in adolescents of other Chinese cities (18.2%-27.2%) [12,65], Poland (24.7%) [51], and Japan (25.7%) [66] based on the same measurement of suicidal ideation. Notably, this study was conducted before the COVID-19 pandemic period. A study conducted on Canadian adolescents reported that the prevalence of suicidal ideation has tripled to 44% since the beginning of the COVID-19 pandemic [67]. The prevalence of adolescent suicidal ideation during the pandemic period in China might have also increased due to the suspension of classes and interruptions of social interactions and activities. As suicidal ideation predicts suicidal behavior, early detection, together with timely and evidence-based secondary interventions, are greatly warranted [68,69].

Corroborating previous cross-sectional and longitudinal findings [11,13], this study found a strong association between adolescent IGD and suicidal ideation. In addition, the high prevalence of IGD of 16.95% (287/1693) was reported among adolescent gamers in this study. Previous studies found a similarly high prevalence among adolescents in China [12,70], which was higher than the global prevalence estimate of 3.1% [71]. It was suggested by some researchers that Asian adolescents tend to have higher IGD prevalence than those in Western countries [72]. Such findings highlight the importance of preventing and treating adolescent IGD, which was associated with various negative consequences, including depression, sleep problems, and physical violence [10,23]. A review summarized several efficacious intervention strategies for the purpose, including psychotherapies such as mindfulness, gaming-specific cognitive behavioral therapy, basic counseling and support groups, family therapy, gaming abstinence, self-discovery camp, and residential camp and parent management [73].

A novel finding of this study is about the mediation mechanisms between IGD and suicidal ideation through psychosocial coping resources (resilience and social support) and psychosocial problems (social anxiety and loneliness). As the direct effect of IGD on suicidal ideation remained statistically significant, the association between IGD and suicidal ideation was partially mediated by 3 indirect paths. These findings add empirical evidence to support the theoretical postulations of the Interpersonal-Psychological Theory of Suicide [24], the Resource Deterioration Model [30,31], and the Conservation of Resource Theory [32-35] aforementioned in the Introduction section. In comparison, the 2-step indirect path that IGD increased psychosocial problems, which then increased suicidal ideation, explained a larger proportion of the association between IGD and suicidal ideation than the other 2-step path and the 3-step path. Psychosocial problems might thus have explained more harmful effects of IGD on suicidal ideation than losses in psychosocial resources. Notably, as these 3 indirect paths only showed an overall partial mediation effect, other unstudied plausible mediators may exist. Examples include health problems (eg, sleep quality), psychological well-being (eg, life satisfaction and quality of life), and stress, which were associated with both IGD and determinants of suicidal ideation [74,75].

Notably, the multivariate logistic regression analyses and SEM in this study were adjusted for depression in addition to sociodemographics. As this study aimed to understand the relationship between IGD and suicidal ideation, given the known associations between depression and IGD [10], suicidal ideation [43], and the 4 psychosocial mediators [45,76], the adjustment for depression could ensure that the relationship of interest would not have been confounded by depression. A similar approach has been used in the extant literature investigating the independent associations between both risk and protective factors and suicidal ideation [43,77].

This study has added empirical evidence to the knowledge that psychosocial coping resources of resilience and social support are protective factors, while psychosocial problems are risk factors for adolescent suicidal ideation. There are thus reasons to believe that modification of these 4 psychosocial variables would reduce suicidal ideation directly and indirectly by reducing the harmful impact of IGD on suicidal ideation. Enhancement of psychosocial resources would also reduce psychosocial problems and the impact of IGD on psychosocial problems. Furthermore, according to a meta-analysis, interventions focusing on psychoeducation and social cognitions were efficacious in reducing loneliness and social isolation. Related interventions included guided social group participation, psychological therapy sessions, self-management, and training enhancing social and emotional skills, social interaction, and social support [78]. Another meta-analysis reported that resilience-improving interventions based on a combination of cognitive behavioral therapy and mindfulness techniques were able to improve resilience and overall well-being (eg, lower levels of anxiety) [79]. Future interventions may consider these approaches.

Strengths and Limitations

This study was novel to investigate the harmful effect of IGD on suicidal ideation among Chinese adolescents and reveal the serial mediation mechanism potentially explaining this association through psychosocial resources and psychosocial problems. The results are implicative for future interventions. However, this study has other limitations. First, although using a single item to assess suicidal ideation has been widely used in the literature [67,68], such an assessment may be subject to misclassification (ie, false negative cases are those having suicidal ideation who were not screened out by the single-item question) in comparison with the multi-item assessment [68]. It may lead to an underestimation of the prevalence of suicidal ideation in this study and implies an alarmingly higher prevalence of suicidal ideation in Chinese adolescents. Second, social desirability bias may exist as the questions regarding IGD symptoms and suicidal ideation are sensitive and socially undesirable. Third, the cross-sectional study design precluded inference of temporal or causal relationships; future longitudinal and intervention studies are needed to confirm the findings. Fourth, the study population was conveniently selected from 7 schools in 2 Chinese cities; generalization of the results to other regions and populations in China and other countries should be made cautiously. Fifth, the relationships among IGD, psychosocial coping resources, and psychosocial problems may be bidirectional, as the literature also reported that resilience, social support, social anxiety, and loneliness were predictors of IGD [22,80-83]. Finally, some potential mediators between IGD and suicidal ideation (eg, sleep quality, psychological well-being, and stress) had not been investigated in this study.

Conclusions

In conclusion, this study observed alarmingly high levels of both IGD and suicidal ideation among adolescents in 2 Chinese metropolises. Very importantly and corroborating previous studies, a strong and positive association between IGD and suicidal ideation was found in this study, indicating that IGD may potentially lead to adolescent suicide. Adolescents and stakeholders (parents, teachers, and health workers) need to be made aware of the potential risk of IGD in elevating suicidal ideation. Future studies should also investigate factors moderating such a relationship. Furthermore, it is important and novel that this study revealed the mechanism of such an association, which included 3 indirect paths through psychosocial resources (resilience and social support) and psychosocial problems (social anxiety and loneliness) between IGD and suicidal ideation. Longitudinal and intervention studies are warranted to confirm these findings and explore other mechanisms. Tailor-made modifications of these psychosocial variables (especially psychosocial problems) may directly and indirectly reduce the harmful impacts of IGD on suicidal ideation.

Acknowledgments

This research was funded by the Shanghai Pujiang Programme (23PJC018) and the Young Scientists Fund of the National Natural Science Foundation of China (82304258). We would like to thank the participants for their contribution to this study.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JTFL and YY performed the conceptualization. YY and JTFL contributed to the methodology. JZ and JBL conducted the investigation. YY managed the software, conducted formal analysis, and performed data curation. JTFL conducted validation and managed the resources. YY, AMSW, VWIF, and JTFL wrote the original draft. YY and JTFL wrote, reviewed, and edited the paper. JTFL and YY performed supervision and managed funding acquisition.

Conflicts of Interest

None declared.

References

1. Adolescent mental health. WHO. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [accessed 2022-11-01]
2. Abraham Z, Sher L. Adolescent suicide as a global public health issue. Int J Adolesc Med Health 2019;31(4). [doi: 10.1515/ijamh-2017-0036] [Medline: 28686572]
3. China health and family planning yearbook. In: National Bureau of Statistics of China. Dongcheng, Beijing, China: Peking Union Medical College Press; 2016.
4. China health and family planning yearbook. In: National Bureau of Statistics of China. Dongcheng, Beijing, China: Peking Union Medical College Press; 2013.
5. Mazza JJ. In: Villarruel FA, Luster T, editors. Youth Suicidal Behavior: A Crisis in Need of Attention. Westport, CT: Praeger Publishers/Greenwood Publishing Group; 2006.
6. Rossom RC, Coleman KJ, Ahmedani BK, Beck A, Johnson E, Oliver M, et al. Suicidal ideation reported on the PHQ9 and risk of suicidal behavior across age groups. J Affect Disord 2017;215:77-84 [FREE Full text] [doi: 10.1016/j.jad.2017.03.037] [Medline: 28319695]
7. Poorolajal J, Haghtalab T, Farhadi M, Darvishi N. Substance use disorder and risk of suicidal ideation, suicide attempt and suicide death: a meta-analysis. J Public Health (Oxf) 2016;38(3):e282-e291. [doi: 10.1093/pubmed/fdv148] [Medline: 26503486]
8. Cheng YS, Tseng PT, Lin PY, Chen TY, Stubbs B, Carvalho AF, et al. Internet addiction and its relationship with suicidal behaviors: a meta-analysis of multinational observational studies. J Clin Psychiatry 2018;79(4):9291. [doi: 10.4088/JCP.17r11761] [Medline: 29877640]
9. International Classification of Diseases for Mortality and Morbidity Statistics. Geneva, Switzerland: World Health Organization; 2018.
10. Mihara S, Higuchi S. Cross-sectional and longitudinal epidemiological studies of Internet gaming disorder: a systematic review of the literature. Psychiatry Clin Neurosci 2017;71(7):425-444 [FREE Full text] [doi: 10.1111/pcn.12532] [Medline: 28436212]
11. Erevik EK, Landrø H, Mattson Å, Kristensen JH, Kaur P, Pallesen S. Problem gaming and suicidality: a systematic literature review. Addict Behav Rep 2022;15:100419 [FREE Full text] [doi: 10.1016/j.abrep.2022.100419] [Medline: 35340768]
12. Yu Y, Yang X, Wang S, Wang H, Chang R, Tsamlag L, et al. Serial multiple mediation of the association between internet gaming disorder and suicidal ideation by insomnia and depression in adolescents in Shanghai, China. BMC Psychiatry 2020;20(1):460 [FREE Full text] [doi: 10.1186/s12888-020-02870-z] [Medline: 32967648]
13. Wicki M, Andronicos M, Studer J, Marmet S, Gmel G. Behavioural addictions as risk factors for incidence and reoccurrence of suicide ideation and attempts in a prospective cohort study among young swiss men. Eur. Psychiatr 2021;64(S1):S176-S177. [doi: 10.1192/j.eurpsy.2021.469]
14. Weiss RS. Loneliness: The Experience of Emotional and Social Isolation. Cambridge, MA: MIT Press; 1973.
15. Bonetti L, Campbell MA, Gilmore L. The relationship of loneliness and social anxiety with children's and adolescents' online communication. Cyberpsychol Behav Soc Netw 2010;13(3):279-285 [FREE Full text] [doi: 10.1089/cyber.2009.0215] [Medline: 20557247]
16. Danneel S, Geukens F, Maes M, Bastin M, Bijttebier P, Colpin H, et al. Loneliness, social anxiety symptoms, and depressive symptoms in adolescence: longitudinal distinctiveness and correlated change. J Youth Adolesc 2020;49(11):2246-2264. [doi: 10.1007/s10964-020-01315-w] [Medline: 32918664]

17. Buckner JD, Lemke AW, Jeffries ER, Shah SM. Social anxiety and suicidal ideation: Test of the utility of the interpersonal-psychological theory of suicide. J Anxiety Disord 2017;45:60-63 [FREE Full text] [doi: 10.1016/j.janxdis.2016.11.010] [Medline: 27940416]
18. Badcock AC, Carrington-Jones P, Stritzke WGK, Page AC. An experimental investigation of the influence of loneliness on changes in belongingness and desire to escape. Suicide Life Threat Behav 2022;52(4):705-715 [FREE Full text] [doi: 10.1111/sltb.12854] [Medline: 35307857]
19. Hedley D, Uljarević M, Wilmot M, Richdale A, Dissanayake C. Understanding depression and thoughts of self-harm in autism: a potential mechanism involving loneliness. Research in Autism Spectrum Disorders 2018;46:1-7. [doi: 10.1016/j.rasd.2017.11.003]
20. Khatcherian E, Zullino D, De Leo D, Achab S. Feelings of loneliness: understanding the risk of suicidal ideation in adolescents with Internet addiction. a theoretical model to answer to a systematic literature review, without results. Int J Environ Res Public Health 2022;19(4):2012 [FREE Full text] [doi: 10.3390/ijerph19042012] [Medline: 35206200]
21. Kim HK, Davis KE. Toward a comprehensive theory of problematic Internet use: evaluating the role of self-esteem, anxiety, flow, and the self-rated importance of Internet activities. Computers in Human Behavior 2009;25(2):490-500. [doi: 10.1016/j.chb.2008.11.001]
22. Gentile D, Choo H, Liau A, Sim T, Li D, Fung D, et al. Pathological video game use among youths: a two-year longitudinal study. Pediatrics 2011;127(2):e319-e329. [doi: 10.1542/peds.2010-1353] [Medline: 21242221]
23. Lemmens JS, Valkenburg PM, Peter J. Psychosocial causes and consequences of pathological gaming. Comput Human Behav 2011;27(1):144-152. [doi: 10.1016/j.chb.2010.07.015]
24. Van Orden KA, Merrill KA, Joiner Jr TE. Interpersonal-psychological precursors to suicidal behavior: a theory of attempted and completed suicide. CPSR 2005;1(2):187-196. [doi: 10.2174/1573400054065541]
25. Davidson CL, Wingate LR, Grant DM, Judah MR, Mills AC. Interpersonal suicide risk and ideation: the influence of depression and social anxiety. Journal of Social and Clinical Psychology 2011;30(8):842-855. [doi: 10.1521/jscp.2011.30.8.842]
26. Connor KM, Davidson JR. Development of a new resilience scale: the connor-davidson resilience scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
27. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988;52(1):30-41. [doi: 10.1037/t02380-000]
28. Taylor SE, Stanton AL. Coping resources, coping processes, and mental health. Annu Rev Clin Psychol 2007;3:377-401. [doi: 10.1146/annurev.clinpsy.3.022806.091520] [Medline: 17716061]
29. Biggs A, Brough P, Drummond S. Lazarus and Folkman's psychological stress and coping theory. The handbook of stress and health: A guide to research and practice 2017:351-364. [doi: 10.1002/9781118993811.ch21]
30. Ensel WM, Lin N. The life stress paradigm and psychological distress. J Health Soc Behav 1991;32(4):321-341. [doi: 10.2307/2137101]
31. Hobbs T. Depression in the caregiving mothers of adult schizophrenics: a test of the resource deterioration model. Community Ment Health J 1997;33(5):387-399. [doi: 10.1023/a:1025070402153] [Medline: 9413666]
32. Hobfoll S. The Psychology and Philosophy of Stress, Culture, and Community. New York: Springer; 1998.
33. Hobfoll SE. Conservation of resources. A new attempt at conceptualizing stress. Am Psychol 1989;44(3):513-524. [doi: 10.1037//0003-066x.44.3.513] [Medline: 2648906]
34. Hobfoll S. The Ecology of Stress. New York, NY: Taylor & Francis; 1988.
35. Hobfoll SE. Conservation of resources theory: its implication for stress, health, and resilience. In: The Oxford Handbook of Stress, Health, and Coping. New York: Oxford University Press; 2011:147.
36. Yen J, Lin H, Chou W, Liu T, Ko C. Associations among resilience, stress, depression, and internet gaming disorder in young adults. Int J Environ Res Public Health 2019;16(17):3181. [doi: 10.3390/ijerph16173181] [Medline: 31480445]
37. Teng Z, Griffiths MD, Nie Q, Xiang G, Guo C. Parent-adolescent attachment and peer attachment associated with Internet gaming disorder: a longitudinal study of first-year undergraduate students. J Behav Addict 2020;9(1):116-128 [FREE Full text] [doi: 10.1556/2006.2020.00011] [Medline: 32359235]
38. Calsyn RJ, Winter JP, Burger GK. The relationship between social anxiety and social support in adolescents: a test of competing causal models. Adolescence 2005;40(157):103-113. [Medline: 15861620]
39. Ko CYA, Chang Y. Investigating the relationships among resilience, social anxiety, and procrastination in a sample of college students. Psychol Rep 2018;122(1):231-245. [doi: 10.1177/0033294118755111]
40. Chen Y, Hicks A, While AE. Loneliness and social support of older people in China: a systematic literature review. Health Soc Care Community 2014;22(2):113-123. [doi: 10.1111/hsc.12051] [Medline: 23714357]
41. Pakdaman S, Saadat SH, Shahyad S, Ahmadi K, Hosseinialhashemi M. The role of attachment styles and resilience on loneliness. International Journal of Medical Research & Health Sciences 2016;5(9):268-274 [FREE Full text]
42. George A, Moolman R. Resilience and suicide ideation: Coping mediator-moderator effects among adolescent learners. Journal of Psychology in Africa 2017;27(6):494-502. [doi: 10.1080/14330237.2017.1375214]
43. Fredrick SS, Demaray MK, Malecki CK, Dorio NB. Can social support buffer the association between depression and suicidal ideation in adolescent boys and girls? Psychology in the Schools 2018;55(5):490-505. [doi: 10.1002/pits.22125]

44. McClelland H, Evans JJ, Nowland R, Ferguson E, O'Connor RC. Loneliness as a predictor of suicidal ideation and behaviour: a systematic review and meta-analysis of prospective studies. J Affect Disord 2020;274:880-896. [doi: 10.1016/j.jad.2020.05.004] [Medline: 32664029]
45. Liang D, Teng M, Xu D. Impact of perceived social support on depression in Chinese rural-to-urban migrants: the mediating effects of loneliness and resilience. Journal Community Psychology 2019;47(7):1603-1613. [doi: 10.1002/jcop.22215]
46. Yu Y, Li JB, Lau JTF. Awareness and potential impacts of the medicalization of Internet gaming disorder: cross-sectional survey among adolescents in China. J Med Internet Res 2021;23(2):e22393 [FREE Full text] [doi: 10.2196/22393] [Medline: 33625362]
47. Yu Y, Li JB, Lau JT. Development and validation of the resource loss due to reduction in gaming time scale among adolescent internet gamers in China. Addict Behav 2023;141:107664. [doi: 10.1016/j.addbeh.2023.107664] [Medline: 36805816]
48. Brady D, Curran M, Carpiano RM. A test of the predictive validity of relative versus absolute income for self-reported health and well-being in the United States. Demogr Res 2023;48:775-808 [FREE Full text] [doi: 10.4054/demres.2023.48.26] [Medline: 37588006]
49. McClure K, Bell KA, Jacobucci R, Ammerman BA. Measurement invariance and response consistency of single-item assessments for suicidal thoughts and behaviors. Psychol Assess 2023;35(10):830-841. [doi: 10.1037/pas0001268] [Medline: 37668583]
50. Rotenstein LS, Ramos MA, Torre M, Segal JB, Peluso MJ, Guille C, et al. Prevalence of depression, depressive symptoms, and suicidal ideation among medical students: a systematic review and meta-analysis. JAMA 2016;316(21):2214-2236 [FREE Full text] [doi: 10.1001/jama.2016.17324] [Medline: 27923088]
51. Zygo M, Pawłowska B, Potembska E, Dreher P, Kapka-Skrzypczak L. Prevalence and selected risk factors of suicidal ideation, suicidal tendencies and suicide attempts in young people aged 13-19 years. Ann Agric Environ Med 2019;26(2):329-336 [FREE Full text] [doi: 10.26444/aaem/93817] [Medline: 31232067]
52. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Washington, DC: American Psychiatric Association; 2013.
53. Yang X, Jiang X, Wu AMS, Ma L, Cai Y, Wong KM, et al. Validation of the Internet gaming disorder symptoms checklist based on the fifth edition of the diagnostic and statistical manual of mental disorders in Chinese adolescents. Child Psychiatry Hum Dev 2023;54(1):26-33. [doi: 10.1007/s10578-021-01213-7] [Medline: 34338962]
54. Yu XN, Lau JT, Mak WW, Zhang J, Lui WW, Zhang J. Factor structure and psychometric properties of the connor-davidson resilience scale among Chinese adolescents. Compr Psychiatry 2011;52(2):218-224. [doi: 10.1016/j.comppsych.2010.05.010] [Medline: 21295229]
55. Chou KL. Assessing Chinese adolescents' social support: the multidimensional scale of perceived social support. Pers Individ Dif 2000;28(2):299-307. [doi: 10.1016/s0191-8869(99)00098-7]
56. March JS, Parker JD, Sullivan K, Stallings P, Conners CK. The multidimensional anxiety scale for children (MASC): factor structure, reliability, and validity. J Am Acad Child Adolesc Psychiatry 1997;36(4):554-565. [doi: 10.1097/00004583-199704000-00019] [Medline: 9100431]
57. Yao S, Zou T, Zhu X, Abela JRZ, Auerbach RP, Tong X. Reliability and validity of the Chinese version of the multidimensional anxiety scale for children among Chinese secondary school students. Child Psychiatry Hum Dev 2006;38(1):1-16 [FREE Full text] [doi: 10.1007/s10578-006-0039-0]
58. Xu S, Qiu D, Hahne J, Zhao M, Hu M. Psychometric properties of the short-form UCLA loneliness Scale (ULS-8) among Chinese adolescents. Medicine (Baltimore) 2018;97(38):e12373 [FREE Full text] [doi: 10.1097/MD.0000000000012373] [Medline: 30235699]
59. Tsai FJ, Huang YH, Liu HC, Huang KY, Huang YH, Liu SI. Patient health questionnaire for school-based depression screening among Chinese adolescents. Pediatrics 2014;133(2):e402-e409 [FREE Full text] [doi: 10.1542/peds.2013-0204]
60. Leung DYP, Mak YW, Leung SF, Chiang VC, Loke AY. Measurement invariances of the PHQ-9 across gender and age groups in Chinese adolescents. Asia Pac Psychiatry 2020;12(3):e12381 [FREE Full text]
61. Manea L, Gilbody S, McMillan D. A diagnostic meta-analysis of the patient health questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. Gen Hosp Psychiatry 2015;37(1):67-75. [doi: 10.1016/j.genhosppsych.2014.09.009] [Medline: 25439733]
62. Zhao B, Kong F, Nam EW. Assessing knowledge, preventive practices, and depression among Chinese university students in Korea and China during the COVID-19 pandemic: an online cross-sectional study. In: Healthcare. Basel, Switzerland: Multidisciplinary Digital Publishing Institute; 2021.
63. Kline RB. Principles and Practice of Structural Equation Modeling. New York, NY: Guilford Publications; 2015.
64. Hooper D, Coughlan J, Mullen MR. Structural equation modelling: guidelines for determining model fit. Articles 2008;2(1):1-8 [FREE Full text]
65. Gong M, Zhang S, Li W, Wang W, Wu R, Guo L, et al. Association between childhood maltreatment and suicidal ideation and suicide attempts among Chinese adolescents: the moderating role of depressive symptoms. Int J Environ Res Public Health 2020;17(17):6025 [FREE Full text] [doi: 10.3390/ijerph17176025] [Medline: 32824995]

66. Nagamitsu S, Mimaki M, Koyanagi K, Tokita N, Kobayashi Y, Hattori R, et al. Prevalence and associated factors of suicidality in Japanese adolescents: results from a population-based questionnaire survey. BMC Pediatr 2020;20(1):467 [FREE Full text] [doi: 10.1186/s12887-020-02362-9] [Medline: 33023527]
67. Turner BJ, Robillard CL, Ames ME, Craig SG. Prevalence and correlates of suicidal ideation and deliberate self-harm in Canadian adolescents during the coronavirus disease 2019 pandemic. Can J Psychiatry 2022;67(5):403-406 [FREE Full text] [doi: 10.1177/07067437211036612] [Medline: 34378420]
68. Hofstra E, van Nieuwenhuizen C, Bakker M, Özgül D, Elfeddali I, de Jong SJ, et al. Effectiveness of suicide prevention interventions: a systematic review and meta-analysis. Gen Hosp Psychiatry 2020;63:127-140. [doi: 10.1016/j.genhosppsych.2019.04.011] [Medline: 31078311]
69. Torok M, Han J, Baker S, Werner-Seidler A, Wong I, Larsen ME, et al. Suicide prevention using self-guided digital interventions: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2020;2(1):e25-e36 [FREE Full text] [doi: 10.1016/S2589-7500(19)30199-2] [Medline: 33328037]
70. Yu Y, Mo PK, Zhang J, Li J, Lau JT. Maladaptive cognitions, loneliness, and social anxiety as potential moderators of the association between Internet gaming time and Internet gaming disorder among adolescent Internet gamers in China. Addictive Behaviors 2022;129:107239. [doi: 10.1016/j.addbeh.2022.107239] [Medline: 35092885]
71. Stevens MW, Dorstyn D, Delfabbro PH, King DL. Global prevalence of gaming disorder: a systematic review and meta-analysis. Aust N Z J Psychiatry 2021;55(6):553-568. [doi: 10.1177/0004867420962851] [Medline: 33028074]
72. Sussman CJ, Harper JM, Stahl JL, Weigle P. Internet and video game addictions: diagnosis, epidemiology, and neurobiology. Child Adolesc Psychiatr Clin N Am 2018;27(2):307-326. [doi: 10.1016/j.chc.2017.11.015] [Medline: 29502753]
73. Zajac K, Ginley MK, Chang R. Treatments of internet gaming disorder: a systematic review of the evidence. Expert Rev Neurother 2020;20(1):85-93 [FREE Full text] [doi: 10.1080/14737175.2020.1671824] [Medline: 31544539]
74. Fazeli S, Mohammadi Zeidi I, Lin CY, Namdar P, Griffiths MD, Ahorsu DK, et al. Depression, anxiety, and stress mediate the associations between internet gaming disorder, insomnia, and quality of life during the COVID-19 outbreak. Addict Behav Rep 2020;12:100307 [FREE Full text] [doi: 10.1016/j.abrep.2020.100307] [Medline: 33110934]
75. Teng Z, Pontes HM, Nie Q, Xiang G, Griffiths MD, Guo C. Internet gaming disorder and psychosocial well-being: a longitudinal study of older-aged adolescents and emerging adults. Addict Behav 2020;110:106530. [doi: 10.1016/j.addbeh.2020.106530] [Medline: 32683173]
76. Stein MB, Fuetsch M, Müller N, Höfler M, Lieb R, Wittchen HU. Social anxiety disorder and the risk of depression: a prospective community study of adolescents and young adults. Arch Gen Psychiatry 2001;58(3):251-256. [doi: 10.1001/archpsyc.58.3.251] [Medline: 11231832]
77. Tae H, Jeong BR, Chae JH. Sleep problems as a risk factor for suicide: are certain specific sleep domains associated with increased suicide risk? J Affect Disord 2019;252:182-189. [doi: 10.1016/j.jad.2019.04.053] [Medline: 30986733]
78. Eccles AM, Qualter P. Review: Alleviating loneliness in young people - a meta-analysis of interventions. Child Adolesc Ment Health 2021;26(1):17-33. [doi: 10.1111/camh.12389] [Medline: 32406165]
79. Joyce S, Shand F, Tighe J, Laurent SJ, Bryant RA, Harvey SB. Road to resilience: a systematic review and meta-analysis of resilience training programmes and interventions. BMJ Open 2018;8(6):e017858 [FREE Full text] [doi: 10.1136/bmjopen-2017-017858] [Medline: 29903782]
80. Robertson TW, Yan Z, Rapoza KA. Is resilience a protective factor of internet addiction? Comput Human Behav 2018;78:255-260. [doi: 10.1016/j.chb.2017.09.027]
81. Zhang S, Tian Y, Sui Y, Zhang D, Shi J, Wang P, et al. Relationships between social support, loneliness, and Internet addiction in Chinese postsecondary students: a longitudinal cross-lagged analysis. Front Psychol 2018;9:1707 [FREE Full text] [doi: 10.3389/fpsyg.2018.01707] [Medline: 30258387]
82. Marino C, Canale N, Vieno A, Caselli G, Scacchi L, Spada MM. Social anxiety and Internet gaming disorder: the role of motives and metacognitions. J Behav Addict 2020;9(3):617-628 [FREE Full text] [doi: 10.1556/2006.2020.00044] [Medline: 32750032]
83. Tras Z. Internet addiction and loneliness as predictors of internet gaming disorder in adolescents. Educ Res Rev 2019;14(13):465-473 [FREE Full text] [doi: 10.5897/err2019.3768]

Abbreviations

CFI: comparative fit index
DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)
ICD-11: International Classification of Diseases (11th Edition)
IGD: internet gaming disorder
RMSEA: root-mean-square error of approximation
SEM: structural equation modeling
UCLA: University of California, Los Angeles

Edited by N Cahill, T Leung; submitted 24.04.23; peer-reviewed by A AL-Asadi, V Sideropoulos; comments to author 11.10.23; revised version received 23.10.23; accepted 18.07.24; published 19.09.24.

Please cite as:
Yu Y, Wu AMS, Fong VWI, Zhang J, Li JB, Lau JTF
Association Between Internet Gaming Disorder and Suicidal Ideation Mediated by Psychosocial Resources and Psychosocial Problems Among Adolescent Internet Gamers in China: Cross-Sectional Study
JMIR Serious Games 2024;12:e48439
URL: https://games.jmir.org/2024/1/e48439
doi:10.2196/48439
PMID:

Correction: Effects of Virtual Reality Therapy for Patients With Breast Cancer During Chemotherapy: Randomized Controlled Trial

Mengdan Li[1], MPH; Zhifu Yu[1], MPH; Hui Li[1], MSN; Li Cao[1], BSN; Huihui Yu[1], MPH; Ning Deng[1], MD; Yunyong Liu[2], MM

1

2

Corresponding Author:
Yunyong Liu, MM

Related Article:

Correction of: https://games.jmir.org/2024/1/e53825/

(JMIR Serious Games 2024;12:e67718) doi:10.2196/67718

In “Effects of Virtual Reality Therapy for Patients With Breast Cancer During Chemotherapy: Randomized Controlled Trial” (JMIR Serious Games 2024;12: e53825) the authors noted two errors.

Affiliation 2 was previously listed as:

> *National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital & Shenzhen Hospital*

This has been changed to the following:

> *National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College*

As this affiliation was also included in the corresponding author’s address, the address has been modified in that location as well.

Additionally, the authorship’s degrees were previously listed as:

> *Mengdan Li, MD; Zhifu Yu, MD; Hui Li, MD; Li Cao, BM; Huihui Yu, MD; Ning Deng, PhD; Yunyong Liu, MD*

These have been replaced as follows:

> *Mengdan Li, MPH; Zhifu Yu, MPH; Hui Li, MSN; Li Cao, BSN; Huihui Yu, MPH; Ning Deng, MD; Yunyong Liu, MM*

The correction will appear in the online version of the paper on the JMIR Publications website on November 6, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 19.10.24; this is a non–peer-reviewed article;accepted 22.10.24; published 06.11.24.

Please cite as:
Li M, Yu Z, Li H, Cao L, Yu H, Deng N, Liu Y
Correction: Effects of Virtual Reality Therapy for Patients With Breast Cancer During Chemotherapy: Randomized Controlled Trial
JMIR Serious Games 2024;12:e67718
URL: https://games.jmir.org/2024/1/e67718
doi:10.2196/67718

Correction: Efficacy of a Virtual 3D Simulation–Based Digital Training Module for Building Dental Technology Students' Long-Term Competency in Removable Partial Denture Design: Prospective Cohort Study

KeXin Liu[1,2,3,*], BDS; YaQian Xu[1,2,3,*], BS; ChaoYi Ma[1,2,3], MMSc; Na Yu[1,2,3], MMSc; FaBing Tan[1,2,3], DDS; Yi Li[1,2,3], BS; YaXin Bai[1,2,3], BS; XiaoMing Fu[1,2,3], DDS; JiaWu Wan[4], BS; DongQi Fan[1,2,3], BDS; HuBin Yin[1,2,3], BS; MeiXi Chen[1,2,3], BS; HongJi Chen[1,2,3], BS; Lin Jiang[1,2,3], DDS; JinLin Song[1], DDS; Ping Ji[1,2,3], DDS; XiaoHan Zhao[5,*], MD; MengWei Pang[1,2,3,*], MDS

1

2

3

4

5

[*]these authors contributed equally

Corresponding Author:
MengWei Pang, MDS

Related Article:

Correction of: https://games.jmir.org/2024/1/e46789

Abstract

(JMIR Serious Games 2024;12:e63113) doi:10.2196/63113

In "Efficacy of a Virtual 3D Simulation–Based Digital Training Module for Building Dental Technology Students' Long-Term Competency in Removable Partial Denture Design: Prospective Cohort Study" (JMIR Serious Games 2024;12:e46789) the authors made one addition.

The "Acknowledgments" section has been amended within the manuscript to read as follows:

> *KXL and YQX are co-first authors of this work, and XHZ and MWP are co-corresponding authors (XHZ can be reached at haoxiaohan_buaa@163.com for correspondence).*
>
> *The engineers' group from Beijing Unidraw Virtual Reality Technology Research Institute Co. Ltd., led by XZ, helped develop the Objective Manipulative Skill Examination of Dental Technicians system, including the digital removable partial denture module. The Zhaozhi student team assisted in the early design and pre-experimental data collection for the digital removable partial denture module. This research was funded by the CQMU Program for Youth Innovation in Future Medicine (W0002 for LJ), and the Chongqing Municipal Teaching Reformation Fund (193070 for XF).*

The correction will appear in the online version of the paper on the JMIR Publications website on August 19, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 10.06.24; this is a non–peer-reviewed article;accepted 17.06.24; published 19.08.24.

Please cite as:
Liu K, Xu Y, Ma C, Yu N, Tan F, Li Y, Bai Y, Fu X, Wan J, Fan D, Yin H, Chen M, Chen H, Jiang L, Song J, Ji P, Zhao X, Pang M
Correction: Efficacy of a Virtual 3D Simulation–Based Digital Training Module for Building Dental Technology Students' Long-Term Competency in Removable Partial Denture Design: Prospective Cohort Study
JMIR Serious Games 2024;12:e63113
URL: https://games.jmir.org/2024/1/e63113
doi:10.2196/63113

Corrigenda and Addenda

Correction: Effects of Electronic Serious Games on Older Adults With Alzheimer's Disease and Mild Cognitive Impairment: Systematic Review With Meta-Analysis of Randomized Controlled Trials

Xinyi Zuo[1*], EdD; Yong Tang[1], PhD; Yifang Chen[2*], DPS; Zhimiao Zhou[3]

[1]Sociology Department, School of Government, Shenzhen University, Shenzhen, Guangdong, China
[2]Institution of Policy Studies, Lingnan University, Tuen Mun, Hong Kong SAR, China (Hong Kong)
[3]Shenzhen Senior High School, Shenzhen, China
[*]these authors contributed equally

Corresponding Author:
Yong Tang, PhD
Sociology Department
School of Government
Shenzhen University
1066 Xueyuan Avenue, Nanshan District
Shenzhen, Guangdong, 518055
China
Phone: 86 13670191924
Email: tangyong@szu.edu.cn

Related Article:

Correction of: https://games.jmir.org/2024/1/e55785

(JMIR Serious Games 2024;12:e65184) doi:10.2196/65184

In "Effects of Electronic Serious Games on Older Adults With Alzheimer's Disease and Mild Cognitive Impairment: Systematic Review With Meta-Analysis of Randomized Controlled Trials" (JMIR Serious Games 2024;12:e55785) the authors noted two errors.

In the "Search Strategy" section, the dates used for keyword search were changed from:

> *The keywords were used to search for papers published from January 1, 2017, to December 1, 2023.*

To the following:

> *The keywords were used to search for papers published from January 1, 2017, to December 31, 2023.*

Additionally, in the "Data Extraction" section, the following information appeared:

> *The first author (XZ) used the modified version of the data extraction table in the Systematic Review Manual of Cochrane Interventions to extract data.*

This has been changed to read as:

> *The co-first author (XZ) used the modified version of the data extraction table in the Systematic Review Manual of Cochrane Interventions to extract data.*

The correction will appear in the online version of the paper on the JMIR Publications website on August 14, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 07.08.24; this is a non–peer-reviewed article;accepted 07.08.24; published 14.08.24.

Please cite as:
Zuo X, Tang Y, Chen Y, Zhou Z
Correction: Effects of Electronic Serious Games on Older Adults With Alzheimer's Disease and Mild Cognitive Impairment: Systematic Review With Meta-Analysis of Randomized Controlled Trials
JMIR Serious Games 2024;12:e65184
URL: https://games.jmir.org/2024/1/e65184
doi:10.2196/65184
PMID:

Letter to the Editor

Therapeutic Uses of Gaming in Mental Health: An Untapped Potential

Jens Peter Eckardt[1], BSW, MSc

Bedre Psykiatri Research Unit (Videnscenter), Copenhagen, Denmark

Corresponding Author:
Jens Peter Eckardt, BSW, MSc
Bedre Psykiatri Research Unit (Videnscenter)
Gammeltorv 14. 2 sal
Copenhagen, 1457
Denmark
Phone: 45 28943288
Email: jp-mail@hotmail.com

Related Article:

Comment on: https://games.jmir.org/2024/1/e46697/

(JMIR Serious Games 2024;12:e57714) doi:10.2196/57714

KEYWORDS

digital mental health interventions; mental health; psychiatry; gaming; serious games; casual video games; commercial games; exergames; adolescent; anxiety; teenage; video game; youth

In their exploratory study, Pine et al [1] unveil promising results indicating the potential therapeutic benefits of using casual video games. While caution is necessary, particularly concerning the interpretation of student feedback, self-assessment effectiveness, recruitment, pandemic effects, and the absence of mental distress screening, the study demonstrates that video games integrating brief mental health messages provide a promising approach to merging digital intervention with the accessibility of commercial gaming. Although a randomized controlled trial is also required for precise clinical impact assessment, these preliminary findings bolster the notion that "gaming" (primarily commercial video games, serious games, and exergames) within mental health services is validated as a viable alternative or complement to conventional methods of prevention, assessment, and treatment.

Gaming, in a broad sense, is one of the most popular leisure activities globally, estimated to involve millions of gamers worldwide [2], making it ubiquitous and omnipresent, regardless of whether one has a mental illness or not. Considering the growing disparity between demand and supply for mental health assistance, combined with factors such as high disease burden, treatment costs, and long waiting lists, new alternative solutions must be explored. Coupled with accelerating technology-based game development and popularity, it may just be a matter of time before gaming truly disrupts several aspects of psychiatric work.

At present, gaming research has been conducted in the context of different psychiatric disorders such as anxiety, depression, eating disorders, attention-deficit/hyperactivity disorder (ADHD), stress symptoms, posttraumatic stress disorder (PTSD), autism, phobias, and schizophrenia, as well as in forensic psychiatry. The results vary, but reduced symptomatology, improved social, executive, and cognitive functions, as well as improved attention processes and problem-solving, have been reported. Gaming has also proven effective in offering temporary distraction from serious events, and it fosters social communities [3].

However, there are several challenges to research and practical application of gaming in mental health services. Moreover, there are critical concerns regarding the limited number of high-quality studies; weak research designs; methodological issues; and questions about generalizability, causality, mechanisms of action, control groups, effect sizes, definitions, terminology, comparability, theoretical strength, harmful effects, and transferability [2,3]. In addition, critics highlight concerns, such as gaming disorders as outlined in the *ICD-11 (International Classification of Diseases 11th Revision)*, prolonged sedentary screen time, exposure to violence, and instances of excessive or problematic gaming behavior [2,4]. Furthermore, critics argue that gaming encourages avoidance tactics, hindering physical interactions within communities. This challenge is compounded by distant communication, escapism, isolation, loneliness, emotional detachment, addiction, sleep disturbances, and physical inactivity, all posing risks of worsening the individual's condition. Critics argue gaming is not a treatment strategy but rather a tool for enhancing communication and presence among individuals.

Research and applications of gaming in psychiatry are expanding and proving beneficial for specific patient demographics, yet there is a pressing need for a more robust knowledge base to fully grasp both the potentials and challenges involved [5]. Capitalizing on these opportunities for clinical use will demand innovative thinking within multidisciplinary research environments [2]. In conclusion, it is evident that gaming, which is deeply embedded in our culture, possesses promising yet unexplored avenues to emerge as a vital component in forthcoming treatments for mental disorders.

Conflicts of Interest

None declared.

Editorial Notice

The corresponding author of "A Novel Casual Video Game With Simple Mental Health and Well-Being Concepts (Match Emoji): Mixed Methods Feasibility Study" declined to respond to this letter.

References

1. Pine R, Mbinta J, Te Morenga L, Fleming T. A novel casual video game with simple mental health and well-being concepts (Match Emoji): mixed methods feasibility study. JMIR Serious Games 2024 Feb 12;12:e46697. [doi: 10.2196/46697] [Medline: 38345848]
2. Colder Carras M, Van Rooij AJ, Spruijt-Metz D, Kvedar J, Griffiths MD, Carabas Y, et al. Commercial video games as therapy: a new research agenda to unlock the potential of a global pastime. Front Psychiatry 2017 Jan 22;8:300. [doi: 10.3389/fpsyt.2017.00300] [Medline: 29403398]
3. Zayeni D, Raynaud J, Revet A. Therapeutic and preventive use of video games in child and adolescent psychiatry: a systematic review. Front Psychiatry 2020;11:36. [doi: 10.3389/fpsyt.2020.00036] [Medline: 32116851]
4. Darvesh N, Radhakrishnan A, Lachance CC, Nincic V, Sharpe JP, Ghassemi M, et al. Exploring the prevalence of gaming disorder and internet gaming disorder: a rapid scoping review. Syst Rev 2020 Apr 02;9(1):68. [doi: 10.1186/s13643-020-01329-2] [Medline: 32241295]
5. Poppelaars M, Lichtwarck-Aschoff A, Otten R, Granic I. Can a commercial video game prevent depression? null results and whole sample action mechanisms in a randomized controlled trial. Front Psychol 2020 Jan 12;11:575962. [doi: 10.3389/fpsyg.2020.575962] [Medline: 33510666]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
***ICD-11*:** *International Classification of Diseases 11th Revision*
PTSD: posttraumatic stress disorder

Edited by T Leung; submitted 24.02.24; this is a non–peer-reviewed article;accepted 16.04.24; published 25.04.24.

Please cite as:
Eckardt JP
Therapeutic Uses of Gaming in Mental Health: An Untapped Potential
JMIR Serious Games 2024;12:e57714
URL: https://games.jmir.org/2024/1/e57714
doi:10.2196/57714
PMID:38662422

Review

The Role of AI in Serious Games and Gamification for Health: Scoping Review

Daniel Tolks[1,2*], Dr rer biol hum; Johannes Jeremy Schmidt[1*], MD; Sebastian Kuhn[1,3], Prof Dr Med

[1]Department of Digital Medicine, Medical Faculty OWL, Bielefeld University, Bielefeld, Germany

[2]Centre for Applied Health Science, Leuphana University Lueneburg, Lueneburg, Germany

[3]Institute for Digital Medicine, University Clinic of Gießen und Marburg, Philipps University Marburg, Marburg, Germany

[*]these authors contributed equally

Corresponding Author:
Sebastian Kuhn, Prof Dr Med
Institute for Digital Medicine
University Clinic of Gießen und Marburg
Philipps University Marburg
Baldingerstraße
Marburg, 35042
Germany
Phone: 49 15120053577
Email: sebastian.kuhn@uni-marburg.de

Abstract

Background: Artificial intelligence (AI) and game-based methods such as serious games or gamification are both emerging technologies and methodologies in health care. The merging of the two could provide greater advantages, particularly in the field of therapeutic interventions in medicine.

Objective: This scoping review sought to generate an overview of the currently existing literature on the connection of AI and game-based approaches in health care. The primary objectives were to cluster studies by disease and health topic addressed, level of care, and AI or games technology.

Methods: For this scoping review, the databases PubMed, Scopus, IEEE Xplore, Cochrane Library, and PubPsych were comprehensively searched on February 2, 2022. Two independent authors conducted the screening process using Rayyan software (Rayyan Systems Inc). Only original studies published in English since 1992 were eligible for inclusion. The studies had to involve aspects of therapy or education in medicine and the use of AI in combination with game-based approaches. Each publication was coded for basic characteristics, including the population, intervention, comparison, and outcomes (PICO) criteria; the level of evidence; the disease and health issue; the level of care; the game variant; the AI technology; and the function type. Inductive coding was used to identify the patterns, themes, and categories in the data. Individual codings were analyzed and summarized narratively.

Results: A total of 16 papers met all inclusion criteria. Most of the studies (10/16, 63%) were conducted in disease rehabilitation, tackling motion impairment (eg, after stroke or trauma). Another cluster of studies (3/16, 19%) was found in the detection and rehabilitation of cognitive impairment. Machine learning was the main AI technology applied and serious games the main game-based approach used. However, direct interaction between the technologies occurred only in 3 (19%) of the 16 studies. The included studies all show very limited quality evidence. From the patients' and healthy individuals' perspective, generally high usability, motivation, and satisfaction were found.

Conclusions: The review shows limited quality of evidence for the combination of AI and games in health care. Most of the included studies were nonrandomized pilot studies with few participants (14/16, 88%). This leads to a high risk for a range of biases and limits overall conclusions. However, the first results present a broad scope of possible applications, especially in motion and cognitive impairment, as well as positive perceptions by patients. In future, the development of adaptive game designs with direct interaction between AI and games seems promising and should be a topic for future reviews.

(JMIR Serious Games 2024;12:e48258) doi:10.2196/48258

KEYWORDS
artificial intelligence; AI; games; serious games; gamification; health care; review

Introduction

Background

Artificial intelligence (AI) and serious games are both relevant topics in the health sector, and the body of studies and literature is continuously growing. Interestingly, in terms of the research landscape, the 2 topics are not connected; rather, existing research views them independently.

The use of games for educational and serious purposes is nearly as old as the history of humankind and is an integral part of our culture [1]. In 1970, Abt [2] used the term "serious games" for the first time in his book with the same name. Sawyer and Smith [3] take a broad definition and consider serious games as "any computerized game whose chief mission is not entertainment and all entertainment games which can be reapplied to a different mission other than entertainment." What serious games have in common is that they pursue a concrete (pedagogical) intention and aim to provide information on a specific topic (eg, health) that is accessible in an entertaining and interactive way to deepen competencies or to achieve a change in behavior [4].

Serious games for health can be used in the fields of medical diagnostics, therapy, and prevention, as well as health promotion and medical or patient education [5]. From a didactic and learning psychology perspective, the effect of serious games is based on the integration of the motivating and multimedia aspects of computer and video games. Serious games can increase engagement, motivation, enthusiasm, and interest [6,7]. There are several existing use cases in health contexts [8-11]. One example is the game EndeavorRx. In 2020, the US Food and Drug Administration permitted its marketing as the first game-based digital therapeutic device to improve attention function in children with attention-deficit/hyperactivity disorder (ADHD) [12]. The game Re-Mission was developed for children with cancer and showed good results regarding compliance and the understanding of disease-related issues in the target group [13]. EMERGE is a simulation game that recreates an emergency department in real time to improve the clinical reasoning skills of physicians [14].

Next to serious games, gamification has emerged as a major trend in the health sector, which is reflected in a growing number of publications, including several meta-analyses [15-17]. The most used definition of this concept is "the use of game design elements in non-game contexts" [18]. The motivational effect of the game elements can be explained in different ways. Sailer et al [19] established the link between various gamification elements (eg, points, leaderboards, and badges) and the self-determination theory proposed by Ryan and Deci [20]. As a theory of motivation, this defines three universal psychological basic needs that determine human action: (1) competence, (2) autonomy, and (3) social inclusion. If ≥1 of these needs are addressed (eg, through gamification elements), this has positive effects on behavior and its determinants [19]. In the health sector, there are numerous studies that have demonstrated the effects of using gamification on motivation, performance, engagement, health, and well-being status [5,21,22].

According to Westera et al [23], computer games have been linked with AI since the first computer was programmed to play chess [24]. New AI methods have been used in computer games, for instance, to generate levels, scenarios, and storylines; to balance complexity; or to add intelligent behaviors to nonplayer characters (NPCs) [25]. However, over the years, various authors have pointed at the marginal penetration of academic game AI methods in industrial game production [26]. AI techniques will become indispensable to coordinate the ever-growing complexity and dynamics of games [23]. AI-driven adaptation and assessment systems are used to offer learner-centered environments [27]. As an example, NPCs controlled by AI can adapt to the behavior of the gamer and can enrich immersive and challenging experiences within the game play.

When transferring these principles to health care, the interaction between AI and games could provide a benefit, especially in the management of chronic diseases, which most game designs already target. The possibility to quickly adapt to new game-generated data or performance and provide live feedback could lead to more individual and thus more patient-centered game design in both illness detection and treatment. This could increase motivation and engagement for patients, leading to higher therapy adherence through more personal involvement. The vast body of evidence in the field of serious games and gamification, along with the growing body of evidence in the use of AI, may thus form a new field of research.

Scope

Therefore, this scoping review sought to generate an overview of the currently existing literature on the interaction of AI and game-based approaches in health care. At this point, to the best of our knowledge, this is the only review that targets this interaction.

The primary objective was to analyze the current body of evidence based on (1) the disease or health issue being evaluated, (2) the process of care in which these projects are located, and (3) the kind of AI and type of game-based approach used and the interaction of both techniques.

A secondary objective was to obtain an overview of publications on the interaction of AI and game-based approaches such as serious games, gamification, commercial games, and game periphery. Another secondary objective was to analyze the quality of the existing studies in this field regarding their grade of evidence and the conducted study types.

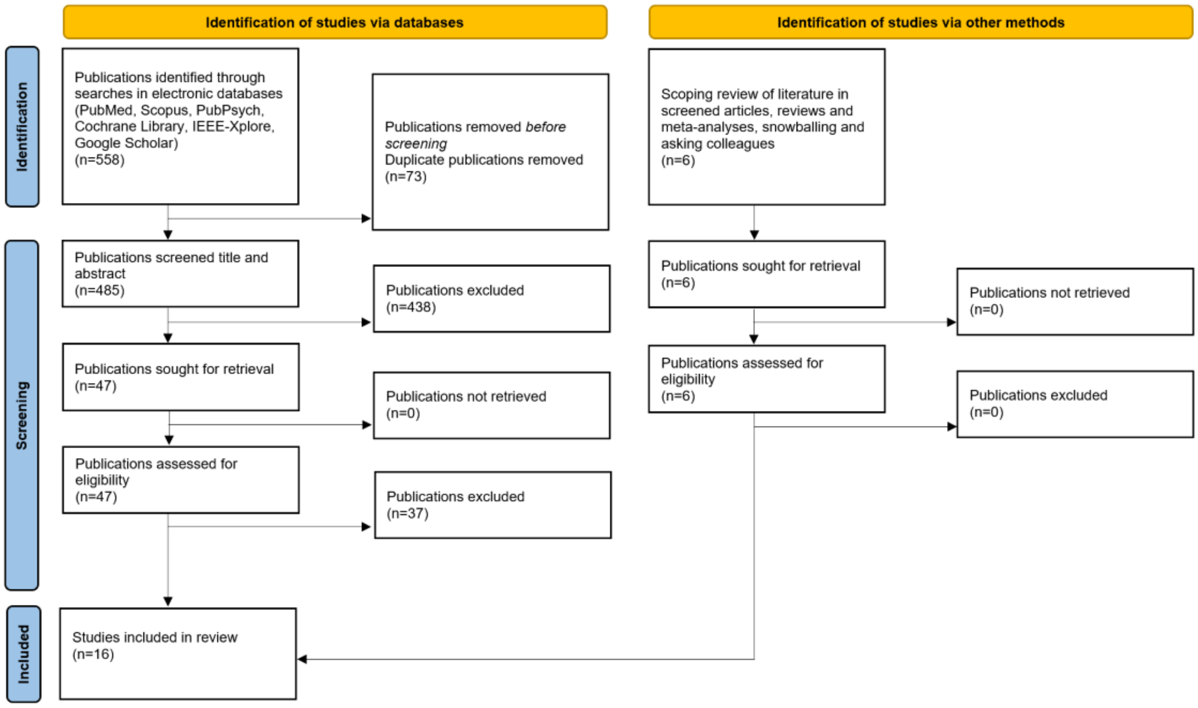
Methods

Overview

For this scoping review, we applied the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols; Figure 1) guidelines [28]. Furthermore, we used the recommendations of the Cochrane Consortium for conducting systematic reviews and the RefHunter website as guidance [29].

Before starting the review process, we defined the inclusion and exclusion criteria (Textbox 1).

Figure 1. PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) flow diagram.



Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Article type: original study, journal article, or conference paper
- Article scope: articles report the use of artificial intelligence (AI), machine learning, and deep learning in combination with game-based approaches (serious games, gamification, and game-based-learning)
- Health profession: medicine
- Area of application: articles that conducted research in the field of education, therapy, and health
- Language: English
- Publication period: last 20 years

Exclusion criteria

- Article type: opinion, commentary, or letter to the editor
- Article scope: not related to AI and game-based approaches
- Health profession: other than medicine
- Area of application: not related to health and medicine
- Language: not in English
- Publication period: published >20 years ago

For this review, we conducted the following steps:

1. Literature search
2. Title and abstract screening
3. Content screening
4. Further in-depth screening (snowballing method and asking colleagues)

Step 1: Literature Search

We applied the search terms primarily in the database PubMed on February 2, 2022. We tested and honed different search terms and Boolean operators (Multimedia Appendix 1) until sufficiently fitting results seemed to have been obtained (n=305). The final search term was defined as follows:

[“game” OR “gamification”] AND “artificial intelligence”

The search was extended to more open databases to assess studies that target AI and serious games in medicine-related research areas or interprofessional approaches that may include medical professions and in more technically oriented databases such as IEEE Xplore [30] to include papers from informatics and engineering with a focus on technical issues.

The same search term ([“game” OR “gamification”] AND “artificial intelligence”) AND “artificial intelligence”) was used for IEEE Xplore (n=98), Cochrane Library [31] (n=25), and PubPsych [32] (n=89). In Scopus, the search term used in the other databases showed fewer results and were modified to extend the range of hits (“serious” AND “game” AND “artificial intelligence”; n=41).

In addition, we conducted a search in Google Scholar [33]. However, the results from Google Scholar were not precise enough for inclusion in a review, which is consistent with the results of several studies [34-36].

After deduplication using Rayyan software (Rayyan Systems Inc), the combined search in these databases identified 545 (97.7%) publications out of the initial 558 identified. We then performed a manual deduplication, which resulted in 60 (11%) of the 545 publications being excluded; thus 485 (89%) publications remained (Multimedia Appendices 1 and 2; Multimedia Appendix 3 [37-51]).

Step 2: Title and Abstract Screening

For the second step of the scoping review (title and abstract screening), we used Rayyan software. The results from the literature search were transferred to the citation software Zotero (version 5.0.85; Rayyan) and to Rayyan software [52]. This software automatically identified duplicates. After iterative deduplication, the publications were subjected to manual screening. The first screening step was conducted using Rayyan and permitted publication inclusion based on their titles and abstracts. Given the volume of the publications to be screened, the title and abstract screening was distributed among 2 authors of this paper (JS and DT). To ensure the uniformity of the screening, the authors conducted several training sessions in Rayyan with the coreviewers. In addition, the authors randomly double-checked some of the excluded publications (25/485, 5.2%) to warrant the consistency of the screening by the other reviewer. This step was conducted independently by both researchers and was followed by a discussion of the results between the 2. As all analysis steps were conducted independently by the 2 researchers, a discussion of differently categorized literature (marked as “Conflict” in Rayyan) and subsequent adaptation were necessary in this step. Overall, only a few adjustments were necessary, and a good agreement between the 2 researchers could be reached. Possible conflicts and all included articles were discussed with a third team member (SK). The data generated by Rayyan can be found in Multimedia Appendix 2.

Step 3: Content Screening

After the primary screening, full-text publications were screened by the 2 lead authors. Toward this end, a table was prepared to compile relevant information (Textbox 2).

Textbox 2. Information compiled for full-text screening.

Relevant details obtained

- Authors, year, title, journal, and digital object identifier (DOI)
- Study type (according to Röhrig et al [53])
- Population, intervention, comparison, and outcomes (PICO) criteria
- Subject (topic of the study)
- Level of evidence (according to the Oxford Centre for Evidence-Based Medicine: Levels of Evidence [54])
- Disease or health issue
- Level of care (prevention, diagnostics, therapy, rehabilitation, nursing, organization or monitoring, and other [55])
- Game variants (serious games, gamification, games, and game controller or periphery)
- Artificial intelligence (AI) technology (machine learning, deep learning, and AI [not further specified])
- Function type (promoting health literacy, analysis and cognition, indirect intervention, direct intervention, documentation of health and medical history, organization and administration, and purchasing and supply) [56]

Some of the studies (4/16, 25%) showed an overlap among different categories (eg, in level of care). In these cases, double classifications were performed. All eligible studies were categorized and coded in detail (Multimedia Appendix 3 [37-51]).

Step 4: Further In-Depth Screening (Snowballing Method and Asking Colleagues)

After conducting the scoping review, we additionally used the “snowballing” approach described by Greenhalgh and Peacock [57], who have stated that in reviews of complex and heterogeneous evidence, formal protocol-driven search strategies may fail to identify important evidence. Informal approaches such as browsing and asking colleagues can substantially increase the efficiency of search efforts. Snowballing methods such as pursuing references of references and electronic citation tracking are very useful for identifying high-quality sources in obscure locations. Therefore, to validate the results of the review, the 2 reviewers searched the literature references used in meta-analyses, reviews, and papers that were closely related to the topic of the search. In addition to using the snowballing method, the method of asking colleagues, as recommended by Greenhalgh and Peacock [57], was applied as a last step.

Results

Overview

When applying the aforementioned search terms, we initially identified 335 studies on the topic of games and the use of AI in health care in the last 20 years. The subsequently performed step of title and abstract screening reduced the number of the initially identified studies from 335 to 47 (14%). In the next step, assessing the actual full-text literature, of the 47 papers, 3 (6%) were excluded because their full text was not in English, and the aforementioned inclusion and exclusion criteria were applied to the remaining 44 (94%). After the full-text screening, 10 (23%) of the 44 papers met all inclusion criteria. Using the snowballing method, 1 additional paper could be identified. Asking colleagues revealed 5 additional papers, which led to an overall total of 16 eligible papers (Figure 1). Not all criteria showed hits (eg, function type showed hits only in 2 categories, whereas level of care showed no hits in nursing).

Categories

Overview

The eligible papers showed a clear emphasis on certain categories (Table 1). Regarding the targeted diseases, the field of motion impairment was investigated the most (5/16, 31%). Cognitive impairment was targeted in 19% (3/16) of the studies, phantom limb pain or limb absence in 19% (3/16), rheumatoid arthritis in 13% (2/16), cancer in 6% (1/16), and ADHD in 6% (1/16). The primary focus on rehabilitation (10/16, 63%) was the most compelling. Of the 16 studies, 5 (31%) took place in the field of prevention, 4 (25%) in the field of diagnostics, and 1 (6%) in a nonrehabilitation therapeutic context (4 double assignments).

Most of the studies (12/16, 75%) applied machine learning as the AI technology, and 13% (2/16) used deep learning, whereas the remaining studies (2/16, 13%) did not specify the AI technology. Most of the studies (11/16, 69%) used a serious game, whereas 19% (3/16) used a commercial games approach. Despite the highly increased use of gamification in health and education, of the 16 studies, only 1 (6%) specifically used gamification to improve the motivation of patients, and 1 (6%) used game design–like interactions.

We further clustered and outlined the eligible papers according to the targeted disease (Table 2). A more detailed description of every included publication with a more specific outline of the use of AI and game variant can be found in Multimedia Appendix 4 [37-51].

Table 1. Categories.

Category	Studies, n (%)
Disease or health topic (n=16)	
Motion impairment	5 (31)
Phantom limb pain or limb absence	3 (19)
Cognitive impairment	3 (19)
Rheumatoid arthritis	2 (13)
Cancer	1 (6)
Attention-deficit/hyperactivity disorder	1 (6)
Other	1 (6)
Function type (n=16)	
Direct intervention	9 (56)
Analysis and cognition	7 (44)
Level of care (n=20[a])	
Prevention	5 (25)
Diagnostics	4 (20)
Therapy	1 (5)
Rehabilitation	10 (50)
AI[b] technology (n=16)	
Machine learning	12 (75)
Deep learning	2 (13)
AI (not further specified)	2 (13)
Game variant (n=16)	
Serious games	11 (69)
Gamification	1 (6)
Games	3 (19)
Game periphery	1 (6)

[a]A total of 4 studies showed an overlap between prevention and diagnostics and were double classified, resulting in an overall total of 20 studies.
[b]AI: artificial intelligence.

Table 2. Overview of included papers, structured by disease or health topic.

Authors; year	Target group (participants, n)	Subject	Study design	Level of evidence[a]	Level of care	Function type	AI[b] technology	Game variant
Disease or health topic: motion impairment								
Yeh et al [43]; 2014	Patients (48)	Noninvasive balance training	Case control study	3b	Therapy	Direct intervention	Machine learning	Games
Lyu et al [44]; 2019	Healthy individuals (8)	Electromyography-controlled knee exoskeleton	Quantitative, proof of concept	5	Rehabilitation	Analysis and cognition	Deep learning	Games
Nasri et al [45]; 2020	Patients (15)	Real-time hand gesture recognition	Case series	4	Rehabilitation	Direct intervention	Deep learning	Serious games
Burdea et al [42]; 2021	Healthy individuals (2)	Game controller–based telerehabilitation	Proof of concept	5	Rehabilitation	Direct intervention	AI	Game controller or periphery
Zhang et al [46]; 2021	Patients (5)	Gait analysis and waist motion capture	Case series	4	Rehabilitation	Direct intervention	Machine learning	Serious games
Disease or health topic: phantom limb pain or limb absence								
Ortiz-Catalan et al [47]; 2016	Patients (14)	Phantom motor execution	Quantitative clinical trial	4	Rehabilitation	Direct intervention	Machine learning	Games
Lendaro et al [48]; 2019	Patients (4)	Phantom motor execution	Quantitative clinical trial	4	Rehabilitation	Analysis and cognition	Machine learning	Serious games
Kristofferson et al [49]; 2021	Patients (4)	Prosthesis system	Explorative study	4	Rehabilitation	Direct intervention	Machine learning	Serious games
Disease or health topic: cognitive impairment								
Valladares-Rodriguez et al [50]; 2018	Healthy individuals or patients (16)	Early detection of mild cognitive impairment	Proof of concept	5	Prevention or diagnostics	Analysis and cognition	Machine learning	Serious games
Valladares-Rodriguez et al [51]; 2019	Patients (74)	Early detection of mild cognitive impairment	Case series	4	Prevention or diagnostics	Analysis and cognition	Machine learning	Serious games
Jung et al [37]; 2019	Patients (12)	Mini-Mental State Examination	Case series	4	Rehabilitation	Direct intervention	Machine learning	Serious games
Disease or health topic: cancer								
Good et al [39]; 2014	Registered players (1077)	Gene Selection for breast cancer survival prediction	Quantitative study	5	Prevention	Analysis and cognition	Machine learning	Serious games
Disease or health topic: attention-deficit/hyperactivity disorder								
Keshav et al [40]; 2019	Patients (7)	Digital attention-related augmented reality game	Case series	4	Prevention or diagnostics	Analysis and cognition	AI	Serious games
Disease or health topic: rheumatoid arthritis								
Varga et al [38]; 2021	Healthy individuals (7)	Virtual arthritis rehabilitation app	Proof of concept	5	Rehabilitation	Direct intervention	Machine learning	Serious games
Varga et al [58]; 2022	Patients (10)	Virtual arthritis rehabilitation app	Case series	4	Rehabilitation	Direct intervention	Machine learning	Serious games

Authors; year	Target group (participants, n)	Subject	Study design	Level of evidence[a]	Level of care	Function type	AI[b] technology	Game variant
Disease or health topic: other								
Pinto et al [41]; 2019	Older adults (11)	Active and assisted living for monitoring daily life activities	Case series	4	Prevention or diagnostics	Analysis and cognition	Machine learning	Gamification

[a]According to the Oxford Centre for Evidence-Based Medicine: Levels of Evidence [54].

[b]AI: artificial intelligence.

Motion Impairment

Almost one-third of the studies (5/16, 31%) targeted the objective of motion impairment. Studies included upper- and lower-limb rehabilitation with a broad range of possible medical indications, ranging from poststroke to vestibular dysfunction. Games were used to enhance motivation and provide a user-friendly at-home training experience. Some of the studies (5/16, 31%) achieved this through an integration of virtual reality and artificial reality. AI was integrated in different ways. Some of the studies (4/16, 25%) used games as a training tool and then analyzed and classified the collected data with AI. Other studies (3/16, 19%) first processed sensor data via AI to improve the quality of an associated game. Direct interaction between the AI and the games component was shown in 2 (40%) of the 5 studies, in which AI adapted the game design and difficulty to the ability level of the patient.

Only 1 (20%) of the 5 studies tested the design using a control group analyzing patient improvements in clinical parameters. All other studies demonstrated the functionality and usability of their technical approach in pilot studies.

Phantom Limb Pain or Limb Absence

Of the 16 studies, 3 (19%) targeted the topic of phantom limb pain or limb absence, where a game environment can support at-home therapy and provide enhanced visual feedback. Of the 3 studies, 2 (67%) by the same research group targeted phantom motor execution with similar approaches. Machine learning was used to improve the quality of electromyography sensor data and thus provide better data input for training. Different training methods in the spectrum of virtual reality and augmented reality and serious games were tested. Of the 3 studies, 1 (33%) focused on ethnographic user–type analysis, and 1 (33%) effected a decrease in phantom pain. The third study tested a machine learning–aided prothesis, comparing 2 different training approaches—1 conventional and 1 via a serious game—to collect electromyography data. Testing was only conducted on 4 patients; however, the results were insignificant.

Cognitive Impairment

In cognitive impairment, the included studies used a set of games covering different cognitive functions as diagnostic instruments. Data were then processed by machine learning techniques to further improve outcome quality. Of the 3 studies, 1 (33%) focused on evaluating patients with cognitive impairment after stroke. Scores acquired from a game set were analyzed by AI and compared with the clinically widely used Mini-Mental State Examination (MMSE) [37]. Of the 3 studies, 2 (67%) used a game set for predicting the future development of mild cognitive impairment, using AI to automatically distinguish between healthy individuals and individuals who were possibly affected. In both fields, pilot studies were conducted with patients, showing high motivation to participate and good usability of the game sets.

Rheumatoid Arthritis

In rheumatoid arthritis, a serious game for hand rehabilitation was developed. Neural networks for processing data and machine learning for testing the accuracy of hand movements for individually adapting difficulty were integrated. Two small pilot studies, 1 with healthy individuals and 1 with patients, showed high accuracy of the machine learning algorithm and good usability, whereas clinical benefits have not yet been measured [38].

Cancer

In cancer, a crowdsourcing campaign was set up via an open web-based game that captured inputs from players regarding their estimation of 5 specific genes, which can be used as predictors of breast cancer survival. Gene selections were processed by machine learning to identify the best prediction models. When only including inputs from people with a self-proclaimed Doctor of Medicine degree, a Doctor of Philosophy degree, or expertise in cancer, the resulting models performed similarly to clinically established gene sets [39].

ADHD Severity

A set of smartglasses was developed to assess ADHD severity through playing an attention-related augmented reality game designed as a social-emotional communication aid. AI was used to analyze video and audio as well as affective and behavioral data and provided users with in-game rewards based on their performance. The study showed significant correlation of the game score to validated clinical gold standard assessments for ADHD [40].

Other

To improve the prevention of cognitive and physical decline, an at-home innovative system consisting of remote monitoring and neurocognitive games was developed. Feedback to the user, including badges or benefits for real-life events, is provided via machine learning analysis. Older adult users indicated "great acceptability" of the system [41].

Discussion

Principal Findings

Currently, there are only a limited number of studies involving a combination of game-based methods and AI in health. Almost one-third of the included studies (5/16, 31%) were centered on addressing motion impairment. The primary emphasis of the research was on rehabilitation. In addition, most of the studies (9/16, 56%) focused on prevention and diagnostics. In terms of AI technology, machine learning was the most commonly used approach (12/16, 75%). Furthermore, serious games were used in most of the studies (11/16, 69%).

When analyzing the studies by disease category, most of the studies (5/16, 31%) used a rehabilitation approach for different aspects of motion impairment (eg, in poststroke conditions, phantom limb pain or limb absence, and rheumatoid arthritis). In this field, studies have a focus on providing individual, at-home, and complex training opportunities for improving motoric limb function, in which therapeutic concepts rely on long-term and self-guided exercising. Games take the role of a training tool, enhancing at-home training motivation and providing multidimensional and exercises compared with the current standard of care. In addition, in some of the studies (5/16, 31%), the integration of virtual reality and augmented reality provided an immersive experience. The role of AI in this context is diverse, sometimes to analyze and classify collected data to improve game setup and level, sometimes to analyze data resulting from game play itself.

In a second cluster, studies for neurological diseases, including those handling cognitive impairment in older adults as well as 1 study for ADHD in younger patients, there was a clear focus on diagnostic evaluations. Here, different sets of games were used to assess various cognitive subdomains, with AI processing these different data inputs and calculating scores and predictions. The advantages in this field are the wide range of possible game designs and the feasibility to play these games individually at home. This could reduce health professionals' time in assessing cognitive function during face-to-face visits or supplement them by enabling longitudinally acquired data sets and trajectories. The first results show promising results in comparison with standard clinical scores obtained using, for instance, the MMSE.

The direct interaction between the games approach and AI technology was only described in 3 (19%) of the 16 studies. Most of the time, the 2 entities follow each other, with the AI technology not analyzing live in-game playing data. However, direct interaction holds a promise of benefit through an AI-enabled assessment of the patient's ability during game play and individualized live adjustments of game design and difficulty. Examples using this approach showed good technical functioning and positive user feedback [42]. Even so, the limited number of published studies suggest that the potential of this integrated approach has so far not been fully used yet. This is rather surprising, given the fact that the direct link between AI and games is widely prevalent in the commercial games sector. The reasons for this are purely speculative. The transition of findings from 1 field to another is still pending, perhaps because studies in the commercial games field have a different scope than those in medicine and health. Another reason could be the resource-intensive nature of research. However, the future potential of this interaction seems promising, with the stimulation of user motivation by game design and gamification elements and with AI being used to process large and multimodal data sources and to perform individualized adaptations.

When analyzing further categories, our review shows that the studies so far have produced very limited quality evidence (all studies have an evidence level of 4 or 5, except for 1 study that has an evidence level of 3b), with most of the studies presenting either a rather technical proof of concept (15/16, 94%) or performing usability testing with a small sample size of healthy individuals and patients (14/16, 88%). Higher-quality studies with control groups and end points focusing on specific clinical outcomes are missing.

Of note, the research field is still young. All studies were conducted in the last 8 years, with 13 (81%) of the 16 studies being published in the last 3 years. All research settings however bear the potential of conducting higher-quality studies with bigger sample sizes and specific medical outcomes in the near future.

However, the studies in this review already show promising results, with overall well-functioning technical implementation of the game elements and high accuracy and usefulness of the AI integration. From the patients' and healthy individuals' perspective, generally high usability, motivation, and satisfaction were found, mostly assessed by established usability questionnaires and qualitative interviews. This is an encouraging perspective for the future because individualized patient-driven at-home diagnostic and therapeutic approaches are increasingly relevant in all fields of medicine.

All 16 studies identified in this review have a relatively low level of evidence (3b: n=1, 6%; 4: n=10, 63%; and 5: n=5, 31%). The risks of bias in these studies are multifaceted. Pilot studies, often conducted to assess the feasibility of a full-scale study, typically featured small sample sizes and often lacked rigorous methodology, randomization, and blinding procedures. As a result, they are susceptible to a range of biases, including selection bias, performance bias, and detection bias. Studies were characterized by weaker methodologies, which can lead to biases in data collection, analysis, and reporting. Nonrandomized studies were prone to selection bias, confounding, and other methodological flaws. The high heterogeneity of the identified studies encompassed a wide range of disease or health issues, populations, and interventions. This heterogeneity makes it challenging or impossible to integrate data and limits overall conclusions.

Limitations

First, as described earlier, the field of research is still very interdisciplinary, and the studies carried out are very diverse based on the vast variety of game-based approaches and therapeutic interventions.

This review only covered original studies in English, which were found in the PubMed, Scopus, IEEE Xplore, Cochrane Library, and PubPsych databases and published in the last 20 years. Although these are widely recognized and commonly used databases in the field of health care research, restricting the review to these 5 databases may have resulted in the exclusion of relevant studies published in other databases owing to this high interdisciplinarity. However, efforts have been made to minimize this limitation using comprehensive search strings, snowballing, and asking colleagues to identify additional relevant literature. In addition, this review also includes interdisciplinary databases such as the more technical-oriented IEEE Xplore and the more pedagogical-oriented PubPsych.

It especially remains unclear whether all projects conducted especially with a more technical focus have been published in scientific journals at all. For future reviews, a more holistic approach should be taken to assess more results from projects that may not have been included in a publication.

In addition, there might be a lack of awareness that research in the engineering, gaming, and fitness spectrum has a direct connection with health-related issues. Thus, it seems possible that certain publications were not fully covered by our already broad search strategy or that promising interventions have not been related to health care yet. This should be mitigated in future studies, considering the growing attention to this young research field.

Another limitation is that this review focused on therapeutic medical interventions rather than on health interventions. AI and game-based approaches in the field of prevention and health promotion have not been included, although this is an important aspect of population health. Game-based approaches especially are used a lot in this field to reach the target groups [8,9,21,22,59-61].

Future Directions

In the near future, the potential of games, which is already established in the commercial games sector, should be applied to the field of serious games and AI. Adaptive game design can be suitable in health care to improve the intervention outcome via AI-driven health care games that assess the skills level of the patient and adapt the difficulty in feedback loops, which could lead to a better harmonization with traditional therapy sessions. NPCs could be used as virtual patients or other health care personnel or relatives to simulate the interprofessional working environment and to improve the interaction and communication with virtual patients [26,27,34].

Finally, the integration of AI and games should carefully consider the ongoing discussions regarding ethical, moral, and data protection issues. In particular, studies describing ethical issues using game-based approaches are scarce [62,63].

Analyzing the currently limited evidence with promising future possibilities in study design and quality, as well as a dynamic research field, it seems, at this stage, that another review should be conducted in the next few years to assess this rapidly growing research field.

Acknowledgments

Open access funding was provided by the open access publishing funds of Philipps-University Marburg and Bielefeld University.

Authors' Contributions

All authors were responsible for conceptualization, methodology, validation, and visualization. DT and JJS were responsible for data curation and formal analysis. DT and SK were responsible for investigation. SK was responsible for funding acquisition, project administration, resources, software, and supervision. DT and JJS wrote the original draft; all authors reviewed and edited the manuscript (contributions of the authors are based on CRediT [contributor roles taxonomy] [64]).

Conflicts of Interest

SK is the founder and a shareholder of MED.digital. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Search term protocol.
[DOCX File , 31 KB - games_v12i1e48258_app1.docx]

Multimedia Appendix 2
Rayyan data of the review process of all studies.
[XLSX File (Microsoft Excel File), 320 KB - games_v12i1e48258_app2.xlsx]

Multimedia Appendix 3
Categorization of all eligible studies.
[XLSX File (Microsoft Excel File), 32 KB - games_v12i1e48258_app3.xlsx]

Multimedia Appendix 4
Detailed descriptions of all included studies.
[DOCX File , 25 KB - games_v12i1e48258_app4.docx]

Multimedia Appendix 5
PRISMA-P-Checklist.
[DOCX File , 32 KB - games_v12i1e48258_app5.docx]

References

1. Tolks D, Dadaczynski K, Horstmann D. Einführung in die Vergangenheit, Gegenwart und Zukunft von Serious Games (for Health). Prävention Gesundheitsförderung 2018 Sep 3;13:272-279. [doi: 10.1007/s11553-018-0667-9]
2. Abt CC. Serious Games. New York, NY: Viking Press; 1970.
3. Sawyer B, Smith P. Serious games taxonomy. The Digital Entertainment Alliance. 2008 Feb 19. URL: https://thedigitalentertainmentalliance.files.wordpress.com/2011/08/serious-games-taxonomy.pdf [accessed 2023-12-20]
4. Sostmann K, Tolks D, Buron S, Plener J, Gaedicke G, Gross M. Serious games healthcare: applying games to medical instruction. German Medical Science. 2010. URL: https://www.egms.de/static/en/meetings/cbt2010/10cbt06.shtml [accessed 2023-12-11]
5. Tolks D, Lampert C, Dadaczynski K, Maslon E, Paulus P, Sailer M. [Game-based approaches to prevention and health promotion: serious games and gamification]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Jun;63(6):698-707. [doi: 10.1007/s00103-020-03156-1] [Medline: 32451596]
6. Lujan HL, DiCarlo SE. Too much teaching, not enough learning: what is the solution? Adv Physiol Educ 2006 Mar;30(1):17-22 [FREE Full text] [doi: 10.1152/advan.00061.2005] [Medline: 16481604]
7. Pitt MB, Borman-Shoap EC, Eppich WJ. Twelve tips for maximizing the effectiveness of game-based learning. Med Teach 2015;37(11):1013-1017. [doi: 10.3109/0142159X.2015.1020289] [Medline: 25776226]
8. Lu AS, Kharrazi H. A state-of-the-art systematic content analysis of games for health. Games Health J 2018 Feb;7(1):1-15 [FREE Full text] [doi: 10.1089/g4h.2017.0095] [Medline: 29293368]
9. Maheu-Cadotte MA, Cossette S, Dubé V, Fontaine G, Lavallée A, Lavoie P, et al. Efficacy of serious games in healthcare professions education: a systematic review and meta-analysis. Simul Healthc 2021 Jun 01;16(3):199-212. [doi: 10.1097/SIH.0000000000000512] [Medline: 33196609]
10. Wang Y, Wang Z, Liu G, Wang Z, Wang Q, Yan Y, et al. Application of serious games in health care: scoping review and bibliometric analysis. Front Public Health 2022 Jun 10;10:896974 [FREE Full text] [doi: 10.3389/fpubh.2022.896974] [Medline: 35757619]
11. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
12. Kollins SH, DeLoss DJ, Cañadas E, Lutz J, Findling RL, Keefe RS, et al. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial. Lancet Digit Health 2020 Apr;2(4):e168-e178 [FREE Full text] [doi: 10.1016/S2589-7500(20)30017-0] [Medline: 33334505]
13. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. Pediatrics 2008 Aug;122(2):e305-e317. [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
14. Chon SH, Timmermann F, Dratsch T, Schuelper N, Plum P, Berlth F, et al. Serious games in surgical medical education: a virtual emergency department as a tool for teaching clinical reasoning to medical students. JMIR Serious Games 2019 Mar 05;7(1):e13028 [FREE Full text] [doi: 10.2196/13028] [Medline: 30835239]
15. Bai S, Hew KF, Huang B. Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. Educ Res Rev 2020 Jun;30:100322. [doi: 10.1016/j.edurev.2020.100322]
16. Zainuddin Z, Chu SK, Shujahat M, Perera CJ. The impact of gamification on learning and instruction: a systematic review of empirical evidence. Educ Res Rev 2020 Jun;30:100326. [doi: 10.1016/j.edurev.2020.100326]
17. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev 2019 Aug 15;32:77-112. [doi: 10.1007/s10648-019-09498-w]
18. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11: Academic MindTrek 2011; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
19. Sailer M, Hense JU, Mayr SK, Mandl H. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput Hum Behav 2017 Apr;69:371-380. [doi: 10.1016/j.chb.2016.12.033]
20. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
21. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov 2;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]

22. Schmidt-Kraepelin M, Thiebes S, Baumsteiger D, Sunyaev A. State of play: a citation network analysis of healthcare gamification studies. In: Proceedings of the 26th European Conference on Information Systems. 2018 Presented at: 26th European Conference on Information Systems; June 23-28, 2018; Portsmouth, UK.
23. Westera W, Prada R, Mascarenhas S, Santos PA, Dias J, Guimarães M, et al. Artificial intelligence moving serious gaming: presenting reusable game AI components. Educ Inf Technol 2019 Jul 30;25:351-380 [FREE Full text] [doi: 10.1007/s10639-019-09968-2]
24. Shannon CE. XXII. Programming a computer for playing chess. Lond Edinb Dublin Philos Mag J Sci 1950;41(314):256-275. [doi: 10.1080/14786445008521796]
25. Yannakakis GN, Togelius J. A panorama of artificial and computational intelligence in games. IEEE Trans Comput Intell AI Games 2015 Dec;7(4):317-335. [doi: 10.1109/tciaig.2014.2339221]
26. Yannakakis GN, Togelius J. Artificial Intelligence and Games. Cham, Switzerland: Springer; 2018.
27. Nyamsuren E, van der Vegt W, Westera W. Automated adaptation and assessment in serious games: a portable tool for supporting learning. In: Proceedings of the 15th International Conferences, ACG 2017. 2017 Presented at: 15th International Conferences, ACG 2017; July 3–5, 2017; Leiden, The Netherlands. [doi: 10.1007/978-3-319-71649-7_17]
28. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4(1):1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
29. Nordhausen T, Hirt J. Manual zur literaturrecherche in fachdatenbanken. Medizinische Fakultät Institut für Gesundheits- und Pflegewissenschaft. 2018 Sep 10. URL: https://www.hwg-lu.de/fileadmin/user_upload/service/bibliothek/RefHunter.pdf [accessed 2023-12-20]
30. IEEE Xplore. URL: https://ieeexplore.ieee.org/Xplore/home.jsp [accessed 2023-10-10]
31. Advanced search. Cochrane Library. URL: https://www.cochranelibrary.com/advanced-search [accessed 2023-10-10]
32. PubPsych. URL: https://www.pubpsych.eu/ [accessed 2023-10-10]
33. Google Scholar. URL: https://scholar.google.com/ [accessed 2023-12-20]
34. Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Res Synth Methods 2020 Mar;11(2):181-217 [FREE Full text] [doi: 10.1002/jrsm.1378] [Medline: 31614060]
35. Giustini D, Boulos MN. Google Scholar is not enough to be used alone for systematic reviews. Online J Public Health Inform 2013 Jul 01;5(2):214 [FREE Full text] [doi: 10.5210/ojphi.v5i2.4623] [Medline: 23923099]
36. Boeker M, Vach W, Motschall E. Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. BMC Med Res Methodol 2013 Oct 26;13:131 [FREE Full text] [doi: 10.1186/1471-2288-13-131] [Medline: 24160679]
37. Jung HT, Daneault JF, Lee H, Kim K, Kim B, Park S, et al. Remote assessment of cognitive impairment level based on serious mobile game performance: an initial proof of concept. IEEE J Biomed Health Inform 2019 May;23(3):1269-1277. [doi: 10.1109/JBHI.2019.2893897] [Medline: 30668485]
38. Varga G, Stoicu-Tivadar L, Nicola S. Serious gaming and AI supporting treatment in rheumatoid arthritis. Stud Health Technol Inform 2021 May 27;281:699-703. [doi: 10.3233/SHTI210262] [Medline: 34042666]
39. Good BM, Loguercio S, Griffith OL, Nanis M, Wu C, Su AI. The cure: design and evaluation of a crowdsourcing game for gene selection for breast cancer survival prediction. JMIR Serious Games 2014 Jul 29;2(2):e7 [FREE Full text] [doi: 10.2196/games.3350] [Medline: 25654473]
40. Keshav NU, Vogt-Lowell K, Vahabzadeh A, Sahin NT. Digital attention-related augmented-reality game: significant correlation between student game performance and validated clinical measures of attention-deficit/hyperactivity disorder (ADHD). Children (Basel) 2019 May 28;6(6):72 [FREE Full text] [doi: 10.3390/children6060072] [Medline: 31142022]
41. Pinto M, Pereira M, Raposo D, Simões M, Castelo-Branco M. GameAAL - an AAL solution based on gamification and machine learning techniques. In: Proceedings of the 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). 2019 Presented at: 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); July 9-11, 2019; Siena, Italy. [doi: 10.1109/cibcb.2019.8791454]
42. Burdea G, Kim N, Polistico K, Kadaru A, Grampurohit N, Roll D, et al. Assistive game controller for artificial intelligence-enhanced telerehabilitation post-stroke. Assist Technol 2021 May 04;33(3):117-128 [FREE Full text] [doi: 10.1080/10400435.2019.1593260] [Medline: 31180276]
43. Yeh SC, Huang MC, Wang PC, Fang TY, Su MC, Tsai PY, et al. Machine learning-based assessment tool for imbalance and vestibular dysfunction with virtual reality rehabilitation system. Comput Methods Programs Biomed 2014 Oct;116(3):311-318. [doi: 10.1016/j.cmpb.2014.04.014] [Medline: 24894180]
44. Lyu M, Chen WH, Ding X, Wang J. Knee exoskeleton enhanced with artificial intelligence to provide assistance-as-needed. Rev Sci Instrum 2019 Sep;90(9):094101. [doi: 10.1063/1.5091660] [Medline: 31575258]
45. Nasri N, Orts-Escolano S, Cazorla M. An sEMG-controlled 3D game for rehabilitation therapies: real-time time hand gesture recognition using deep learning techniques. Sensors (Basel) 2020 Nov 12;20(22):6451 [FREE Full text] [doi: 10.3390/s20226451] [Medline: 33198083]

46. Zhang Q, Jin T, Cai J, Xu L, He T, Wang T, et al. Wearable triboelectric sensors enabled gait analysis and waist motion capture for IoT-based smart healthcare applications. Adv Sci (Weinh) 2022 Feb;9(4):e2103694 [FREE Full text] [doi: 10.1002/advs.202103694] [Medline: 34796695]
47. Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, Zepeda-Echavarria A, Caine-Winterberger K, Kulbacka-Ortiz K, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain. Lancet 2016 Dec 10;388(10062):2885-2894. [doi: 10.1016/S0140-6736(16)31598-7] [Medline: 27916234]
48. Lendaro E, Hermansson L, Burger H, Van der Sluis CK, McGuire BE, Pilch M, et al. Phantom motor execution as a treatment for phantom limb pain: protocol of an international, double-blind, randomised controlled clinical trial. BMJ Open 2018 Jul 16;8(7):e021039 [FREE Full text] [doi: 10.1136/bmjopen-2017-021039] [Medline: 30012784]
49. Kristoffersen MB, Franzke AW, Bongers RM, Wand M, Murgia A, van der Sluis CK. User training for machine learning controlled upper limb prostheses: a serious game approach. J Neuroeng Rehabil 2021 Feb 12;18(1):32 [FREE Full text] [doi: 10.1186/s12984-021-00831-5] [Medline: 33579326]
50. Valladares-Rodriguez S, Pérez-Rodriguez R, Fernandez-Iglesias JM, Anido-Rifón LE, Facal D, Rivas-Costa C. Learning to detect cognitive impairment through digital games and machine learning techniques. Methods Inf Med 2018 Sep;57(4):197-207. [doi: 10.3414/ME17-02-0011] [Medline: 30248709]
51. Valladares-Rodriguez S, Fernández-Iglesias MJ, Anido-Rifón L, Facal D, Rivas-Costa C, Pérez-Rodríguez R. Touchscreen games to detect cognitive impairment in senior adults. A user-interaction pilot study. Int J Med Inform 2019 Jul;127:52-62. [doi: 10.1016/j.ijmedinf.2019.04.012] [Medline: 31128832]
52. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
53. Röhrig B, du Prel JB, Wachtlin D, Blettner M. Types of study in medical research: part 3 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2009 Apr;106(15):262-268 [FREE Full text] [doi: 10.3238/arztebl.2009.0262] [Medline: 19547627]
54. OCEBM levels of evidence. Centre for Evidence-Based Medicine (CEBM), University of Oxford. URL: https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence [accessed 2023-12-20]
55. Kuhn S, Thranberend T. Medizinische begutachtung digitaler gesundheitsanwendungen im projekt "trusted-health-apps". Bertelsmann Stiftung. URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/medizinische-begutachtung-digitaler-gesundheitsanwendungen-im-projekt-trusted-health-apps-all [accessed 2023-12-07]
56. AppQ: gütekriterien-kernset für mehr qualitätstransparenz bei digitalen gesundheitsanwendungen. Bertelsmann Stiftung. 2019 Oct. URL: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Studienbericht_AppQ_191028.pdf [accessed 2023-12-07]
57. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 2005 Nov 05;331(7524):1064-1065 [FREE Full text] [doi: 10.1136/bmj.38636.593461.68] [Medline: 16230312]
58. Varga G, Stoicu-Tivadar L, Nicola S. Serious gaming and artificial intelligence in rehabilitation of rheumatoid arthritis. Stud Health Technol Inform 2022 Jun 29;295:562-565. [doi: 10.3233/SHTI220790] [Medline: 35773936]
59. Koivisto J, Hamari J. Gamification of physical activity: a systematic literature review of comparison studies. In: Proceedings of the GamiFIN Conference 2019. 2019 Presented at: GamiFIN Conference 2019; April 8-10, 2019; Levi, Finland.
60. van Gaalen AE, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma AD, Georgiadis JR. Gamification of health professions education: a systematic review. Adv Health Sci Educ Theory Pract 2021 May;26(2):683-711 [FREE Full text] [doi: 10.1007/s10459-020-10000-3] [Medline: 33128662]
61. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
62. Barnes S, Chandler T, Salazar MG. Help over harm: practical and ethical considerations for the evaluation and deployment of therapeutic games. In: Proceedings of ICACTCE'23 — The International Conference on Advances in Communication Technology and Computer Engineering. 2023 Presented at: ICACTCE'23; February 24-25, 2023; Bolton, UK. [doi: 10.1007/978-3-031-37164-6_54]
63. Kim TW, Werbach K. More than just a game: ethical issues in gamification. Ethics Inf Technol 2016 May 12;18:157-173. [doi: 10.1007/s10676-016-9401-5]
64. CRediT - contributor roles taxonomy. CRediT. URL: https://credit.niso.org/ [accessed 2023-10-11]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
AI: artificial intelligence
MMSE: Mini-Mental State Examination
NPC: nonplayer character

PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols

Edited by A Mavragani; submitted 19.04.23; peer-reviewed by N Mungoli, R Gorantla, M Gasmi ; comments to author 21.09.23; revised version received 12.10.23; accepted 18.10.23; published 15.01.24.

Please cite as:
Tolks D, Schmidt JJ, Kuhn S
The Role of AI in Serious Games and Gamification for Health: Scoping Review
JMIR Serious Games 2024;12:e48258
URL: https://games.jmir.org/2024/1/e48258
doi:10.2196/48258
PMID:38224472

Review

The Effects of Serious Games on Cardiopulmonary Resuscitation Training and Education: Systematic Review With Meta-Analysis of Randomized Controlled Trials

Pengfei Cheng[1*], MSc; Yangxi Huang[2*], MSc; Pengyu Yang[3*], MSc; Haizhen Wang[1], MSc; Baichao Xu[4], PhD; Chaoran Qu[5], MSc; Hua Zhang[6,7], PhD

[1]Department of Nursing, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China

[2]School of Nursing, The University of Hong Kong, Hong Kong, China

[3]Department of Nursing, West China Hospital, Sichuan University, Chengdu, China

[4]Department of Physical Education, Hainan Medical University, Haikou, China

[5]Department of the Operating Room, Shenzhen People's Hospital (The Second Clinical Medical College, Jinan University; The First Affiliated Hospital, Southern University of Science and Technology), Shenzhen, China

[6]International Nursing School, Hainan Medical University, Haikou, China

[7]Key Laboratory of Emergency and Trauma, Ministry of Education, Haikou, China

[*]these authors contributed equally

Corresponding Author:
Hua Zhang, PhD
International Nursing School
Hainan Medical University
Number 3, Xueyuan Road.
Haikou, 571199
China
Phone: 86 13698954876
Email: zhanghuashelley@hotmail.com

Abstract

Background: Serious games have emerged as an innovative educational strategy with the potential to significantly enhance the quality and effectiveness of cardiopulmonary resuscitation (CPR) training. Despite their promise, there remains a degree of controversy when comparing the advantages of serious games with traditional CPR training methods. This study seeks to provide a comprehensive assessment of the impact of serious games on CPR training and education by systematically analyzing the results of previous research.

Objective: This study aimed to assess the effect of serious games on CPR training and education by summarizing and pooling the results of previous studies.

Methods: We conducted a thorough and systematic search across 9 prominent web-based databases, encompassing the period from the inception of these databases until April 1, 2023. The databases included in our search were PubMed, Cochrane Library, Wiley Online Library, EBSCO (PsycInfo), SpringerLink, Chinese Biology Medicine Disc, Vip Journal Integration Platform, Wanfang Database, and Chinese National Knowledge Infrastructure. The studies selected adhered to the following criteria: (1) being a randomized controlled trial comparing serious games and traditional methods for CPR training; (2) having participants aged 12 years or older in CPR; (3) having an experimental group using serious games and a control group using nongame methods for CPR instruction; and (4) having outcomes including theoretical and skill assessments, compression depth, and rate. The Cochrane risk of bias assessment tool was used to evaluate the risk of bias. Data analysis was performed using RevMan (version 5.3; Cochrane Training), and mean differences (MDs) and standardized mean differences (SMDs) with 95% CIs were used to calculate continuous variables.

Results: A total of 9 articles were included, involving 791 study participants, of whom 395 in the experimental group taught CPR training using serious games and 396 in the control group taught CPR training using traditional methods. The results of our meta-analysis indicate that the use of serious games in CPR training yields outcomes that are comparable in effectiveness to traditional training methods across several key areas. Specifically, serious games demonstrated equivalence to traditional formats in theory assessment (SMD –0.22, 95% CI – 0.96 to 0.51; *P*=.55), skill assessment (SMD –0.49, 95% CI –1.52 to 0.55; *P*=.36), compression depth (MD –3.17, 95% CI –0.18 to 6.53; *P*=.06), and compression rate (MD –0.20, 95% CI –7.29 to 6.89; *P*=.96).

Conclusions: In summary, serious games offer a viable and effective CPR education approach, yielding results comparable to traditional formats. This modality is a valuable addition to CPR training methodologies. However, caution is warranted in interpreting these findings due to limited controlled trials, small sample sizes, and low-quality meta-analyzed evidence.

(JMIR Serious Games 2024;12:e52990) doi:10.2196/52990

KEYWORDS

CPR; education; meta-analysis; serious game; training

Introduction

Background

Out-of-hospital cardiac arrest (OHCA) is a critical medical emergency characterized by the sudden cessation of heart function, resulting in an abrupt loss of blood flow. OHCA incidents frequently occur in community settings, schools, homes, and public places [1]. Despite sustained efforts, OHCA survival rates remain disheartening, largely due to modifiable factors such as bystander cardiopulmonary resuscitation (CPR), automated external defibrillator (AED) use, and the timing of emergency medical services (EMS) intervention [2,3]. In the United States, OHCA affects over 88.8 adults per 100,000 adults annually, with a mere 9.0% discharge survival rate, as reported by the American Heart Association [4]. Similarly, in Europe, the annual incidence of OHCA among adults ranges from 67 to 170 per 100,000, with discharge survival rates varying from 0% to 18% [5]. In China, more than 540,000 individuals experience OHCA each year, but the survival rate remains at approximately 2% [6]. These statistics underscore that OHCA, despite regional disparities, has emerged as a substantial public health challenge, imperiling the well-being of citizens [7]. OHCA is typified by sudden respiratory distress, pulse cessation, and loss of consciousness, necessitating immediate and effective first-aid measures within the critical 4-minute window [8]. However, current prehospital EMS services often struggle to reach the scene promptly to address emergencies in public spaces [9]. Consequently, first responders (FRs), nonmedical professionals in public areas, shoulder the responsibility of on-site rescue efforts [10]. Swift and efficient basic life support interventions administered by FRs not only create a vital time buffer for EMS teams to arrive but also substantially elevate the chances of patients with OHCA surviving [11].

CPR, encompassing artificial respiration and chest compressions, stands as one of the simplest and most universally applicable techniques for basic life support during OHCA emergencies [12]. The quality of chest compressions holds immense significance in preserving organ perfusion. Consequently, the timely and effective administration of CPR plays a pivotal role in determining both the survival rate and neurological outcomes for patients with OHCA [3]. To enhance the widespread adoption of CPR and ensure that more individuals are proficient in this vital first-aid technique, the World Health Organization and the International Liaison Committee on Resuscitation endorsed the “Kids Save Lives” statement, which calls for CPR training for students, adolescents, and adults aged 12 years or older who already have the physical fitness and learning ability to understand and remember CPR skills to empower young people, including children aged 12 years, with CPR skills. Develop a generation of proactive and empowered community members who are expected to make a difference in emergency situations, especially in the context of OHCA, with the goal of increasing survival and improving long-term outcomes for patients with OHCA [13,14].

Serious games are increasingly used in medical education, encompassing medical theory instruction, clinical skills training, cognitive rehabilitation exercises, and patient health education. The integration of serious games into medical simulation programs is seen as a means to enhance the efficiency and effectiveness of training programs [15,16]. Otero-Agra et al [17] used serious games to instruct middle school students in CPR, revealing that 61.7% of participants acquired correct CPR techniques, with 93.4% achieving an average chest compression depth exceeding 50 mm. These results endorse serious games as effective tools for knowledge acquisition and the mastery of high-quality CPR skills. To optimize their use as an educational strategy, serious games must possess robust content and cater to the target audience. Integrating learning theory with game requirements enhances student engagement and ensures the efficacy of learning [18]. High fidelity is crucial, especially for medical students, as the knowledge and skills acquired in serious games will be applied in future clinical practice involving real patients. High-fidelity serious games bridge the gap between virtual gaming scenarios and clinical reality, boosting rescue confidence and self-efficacy [19]. Creutzfeldt et al [20] used serious games based on massively multiplayer virtual worlds technology to train 36 high school students in CPR. After 90-120 minutes of game-based sessions, participants reported a significant increase in self-efficacy, endorsing the effectiveness of serious games for CPR instruction. Moreover, serious games can incorporate adaptive learning features, adjusting difficulty and content based on the learner’s proficiency, ensuring tailored learning for individuals with varying CPR skill levels [21].

The incorporation of serious games into CPR training aims to enhance the learning process by rendering it more engaging, interactive, and effective. Compared to conventional methods relying on lectures, videos, and hands-on practice, serious games make the learning experience more enjoyable, interactive, and motivation-driven, integrating features such as scores, levels, and rewards [21,22]. Notably, serious games for CPR training are user-facing, offering immediate training opportunities, flexible learning schedules, and detailed real-time feedback on CPR performance [23]. In contrast, traditional teaching models often limit training opportunities, providing delayed feedback, particularly in large-scale group activities where individual feedback is frequently overlooked [24]. A systematic review by Lim et al [25] underlines that the absence of regular retraining and effective feedback in traditional CPR education can impact skill retention. Serious games address these shortcomings by providing continuous opportunities for practice and feedback. Moreover, serious games support collaborative learning, enabling learners to respond jointly to virtual CPR scenarios and develop teamwork and communication skills. They also offer diverse immersive first aid scenarios with varying causes of cardiac arrest, an aspect unattainable in traditional teaching formats [16,26]. This multifaceted approach not only compensates for the deficiencies in traditional methods but also promotes a dynamic and engaging learning environment in CPR training. Considering the advantages mentioned above, the 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care recommended the incorporation of serious games into CPR training and education to enhance teaching methods and improve instructional quality, taking into account advancements in training equipment and teaching formats [27]. However, Dankbaar [28] concluded that serious games have limitations in terms of time and their ability to provide learners with sufficient knowledge acquisition and complex skill improvement. In summary, there exists a degree of controversy regarding the impact of serious games on CPR training and education. Therefore, we aimed to conduct a meta-analysis to determine the effectiveness of serious games in CPR training and education.

Research Gap and Aim

While numerous researchers have explored and experimented with serious games for CPR training, published randomized controlled trial (RCT) studies have explored and experimented with the effect of serious games applied to CPR training, and their effectiveness has been proven and supported [16,29]. However, due to the limitations of research, the generalization of research conclusions is affected. Specifically, (1) these RCTs were single-center studies with small sample sizes; (2) specific serious games limit the reliability of the findings in different settings of serious games or target populations; (3) outcomes were mostly assessed by questionnaires, and there were a lack of reliable, automated, and repeatable methods to measure their efficacy; and (4) there is a lack of methodological specifications and standard protocols for the use of serious games. Furthermore, there is a lack of systematic evaluation or meta-analysis of the effectiveness of serious games-based CPR training. Consequently, it is necessary to quantitatively analyze the objective effect of serious games–based training through meta-analysis. In view of this, we conducted a meta-analysis to comprehensively evaluate the effect of serious games on CPR training and teaching.

Methods

Overview and Registration

This systematic review adheres to the guidelines set forth by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [30] and was registered in advance in the PROSPERO (International Prospective Register of Systematic Reviews) database (registration number CRD42023423089).

Search Strategy

Our search was conducted in several databases, including PubMed, Cochrane Library, Wiley Online Library, EBSCO (PsycInfo), and SpringerLink. Besides, Chinese databases, including the China National Knowledge Infrastructure (CNKI), China Biomedical Literature Database, VIP Journal Integration Platform, and Wanfang Database, were searched. The search was conducted from the inception of the databases until April 1, 2023. We limited the publication language to English and Chinese. English search terms included "serious game," "gam*," "cardiopulmonary resuscitation," "CPR," "basic life support," "BLS," "first aid training," "resuscitation education," "emergency skill," etc. The search involved a combination of subject terms and free words, with a manual retrospective search of references and associated literature to ensure a comprehensive search of relevant studies. Multimedia Appendix 1 provides detailed information on the search strategies, including search terms, and the process used.

Eligibility Criteria for This Review

The eligibility of studies was assessed based on the following criteria: (1) the study type should be an RCT comparing the effectiveness of serious games with other traditional training methods for teaching CPR; (2) the study population should include participants aged 12 years or older who participated in CPR training or first aid training that covered basic life support for CPR; (3) interventions in the experimental group should involve the use of serious games for CPR training instruction, while the control group should receive other methods of CPR theory and skills training instruction excluding serious games, with no limitations on the types of games or software used; and (4) outcome measures should include one or more of the indicators of theoretical assessment, CPR skill assessment, compression depth, and compression rate. Additionally, duplicate or multiple manuscripts, literature in languages other than Chinese or English, literature with inaccessible full text, incomplete or missing data, improper data collection, or errors in statistical methods were excluded.

Screening Process

Two authors (PC and PY), who were trained in evidence-based methods, independently conducted the screening of literature and extraction of data. All references were managed using EndNote X9 (Clarivate), a reference management software. After removing duplicates, the remaining references were first screened based on titles and abstracts. Subsequently, full-text screening was performed independently by the authors in duplicate to determine the inclusion of literature. Disagreements were resolved through discussion or adjudication by a third author (HZ).

Quality Assessment

The Cochrane handbook's criteria for assessing the risk of bias in RCTs were used to evaluate the methodological quality of the trials [31]. The assessment covered various aspects, including selection bias, concealment of the allocation scheme, implementation bias, measurement bias, missed visit bias, reporting bias, and other biases. Each item was categorized as "low risk of bias," "unclear," or "high risk of bias." In cases where differing opinions arose, a third author (HZ) was involved to reach a consensus.

Data Extraction

For data extraction, we used Excel (2010; Microsoft Corporation) to create a standardized form. The form included the following information: (1) basic details such as the first author, publication year, and country of the study; (2) population characteristics, sample size, and information about the serious games used in training and teaching; (3) specific interventions for the test and control groups; and (4) outcome measures and the tools used for measurement.

Statistical Analysis

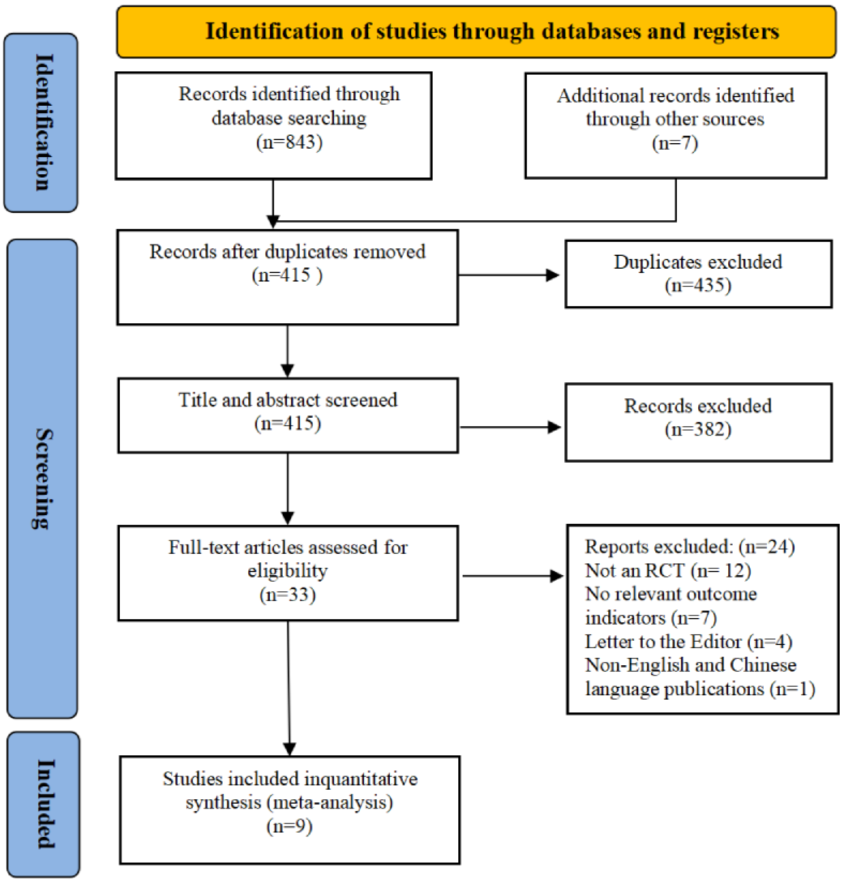
Data analysis was carried out using RevMan (version 5.3; Cochrane Training). To assess heterogeneity, the Q test and the I^2 test were used. If the resulting *P* value was greater than or equal to .1 and I^2 was less than or equal to 50%, it indicated low heterogeneity among the findings, leading to the selection of the fixed-effects model for meta-analysis. Otherwise, the random-effects model was used. When comparing groups, continuous variables were analyzed using mean difference (MD) if the same measurement instrument was used or standardized mean difference (SMD) if different instruments were used. Both effect measures were reported with 95% CIs. For continuous data that did not follow a normal distribution in the included studies and were expressed as medians, extreme values, or quartiles, a specific web-based formula calculator developed by Professor Luo et al [32] from Hong Kong Baptist University was used. This calculator, designed for meta-analysis data conversion, enabled statistical estimation of the data. Leave-one-out analysis was used to conduct sensitivity analysis, that is, omitting one study at a time from the meta-analysis and examining the impact on the overall effect size, then judging the robustness and reliability of the results and exploring the sources of heterogeneity [33]. Statistical significance was determined at a $P<.05$. The level of evidence was evaluated using the GRADEpro GDT web-based tool.

Results

Study Selection

After conducting a comprehensive search across various databases, a total of 843 RCTs were found. Additionally, 7 more studies were obtained by snowballing. Following the removal of duplicates, 415 articles were screened based on their titles and abstracts. Out of these, 382 articles were excluded, and the remaining 33 articles were examined in their entirety. Ultimately, a total of 9 full-text articles were considered for quantitative synthesis. This included 5 papers in English and 4 papers in Chinese. More specific information can be found in the study's PRISMA flowchart (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the selection process. RCT: randomized controlled trial.



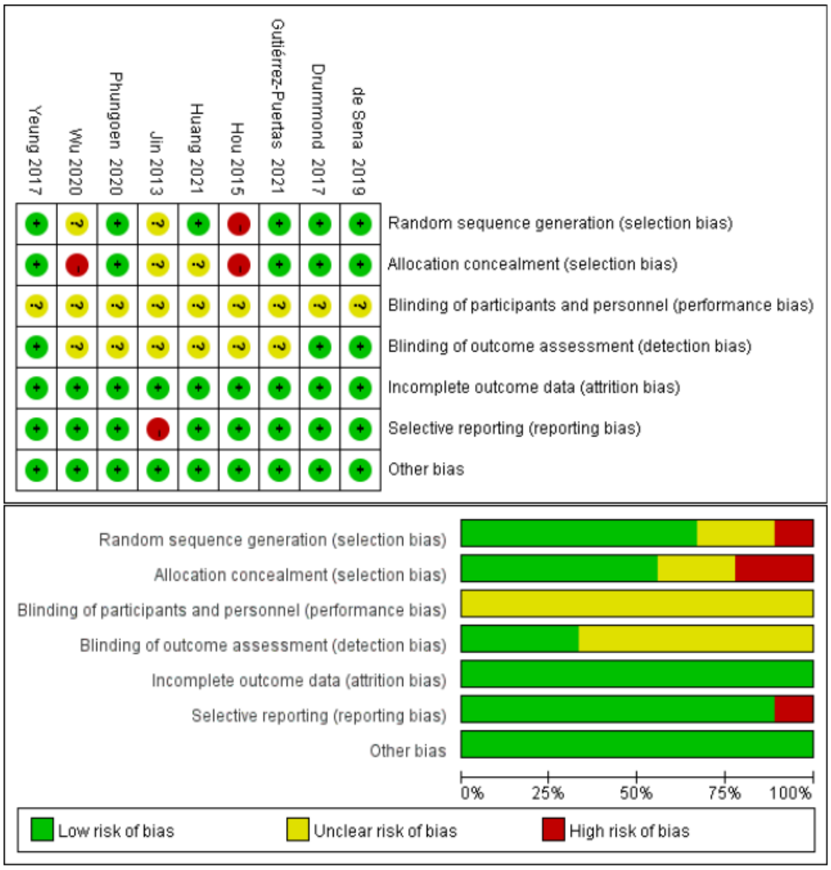
Characteristics of Included Studies

We included 9 RCTs [34-42] from 6 countries. There were no statistical differences in general information between the trial and control groups in each study. A total of 791 study participants were included, with 395 in the experimental group taught CPR training using serious games and 396 in the control group taught CPR training using traditional methods. Additional information can be found in Multimedia Appendix 2 [34-42].

Quality Assessment

The risk of bias evaluation of the included literature is presented in Figure 2 [34-42] (the colors green, yellow, and red in the figure mean “low risk of bias,” “unclear risk of bias,” and “high risk of bias,” respectively). The quality of the included studies was found to be acceptable. In 6 RCTs [36-38,40-42], they described the generation of random sequences, of which 5 RCTs [36-38,40,41] described methods of allocation concealment. Due to CPR training and teaching, it was not possible to blind participants. In 3 RCTs [36-38], they applied the blinding method for researchers. Additionally, in 1 RCT [34], they had a high risk of reporting bias, and all 9 RCTs had complete data and did not have any other bias.

Figure 2. Methodological quality assessment of risk of bias for the included trials.

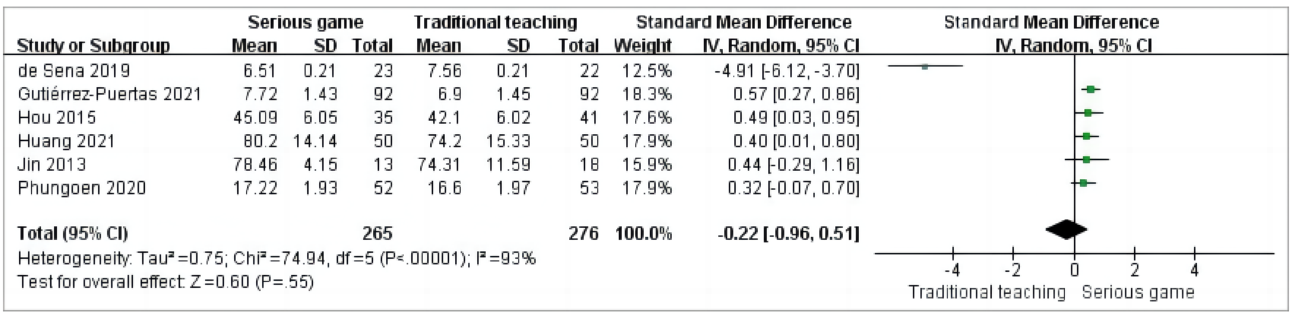


Meta-Analysis Results

The Effect of Serious Games Teaching on CPR Theory Performance

In the analysis, 6 out of the 9 studies [34,35,38,40-42] used posttraining CPR theory assessment as an outcome measure in RCTs. The pooled results revealed significant heterogeneity among the studies (P<.001; I^2=93%), necessitating the use of a random-effects model for the meta-analysis. Figure 3 [34,35,38,40-42] demonstrates that there was no significant disparity in the theory assessment between the 2 groups under investigation (SMD –0.22, 95% CI –0.96 to 0.51; P=.55).

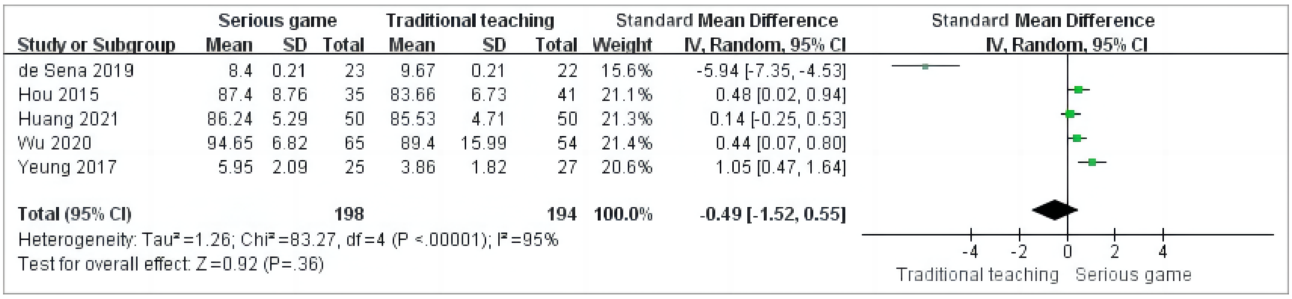
Figure 3. Meta-analysis of the effect of serious games on theory assessment.



The Effect of Serious Games Teaching on the Performance of CPR Skills Operations

Posttraining CPR skill manipulation performance was assessed as an outcome indicator in 5 RCTs [35,36,38,39,42] out of the 9 studies included. Meta-analysis was conducted using a random-effects model due to heterogeneity among the studies (P<.001; I^2=95%). The results indicated that there was no significant difference in skills assessment between the 2 study groups (SMD –0.49, 95% CI –1.52 to 0.55; P=.36). This suggests that the use of serious games for CPR training did not lead to a significantly different skill level compared to other traditional training methods (Figure 4 [35,36,38,39,42]).

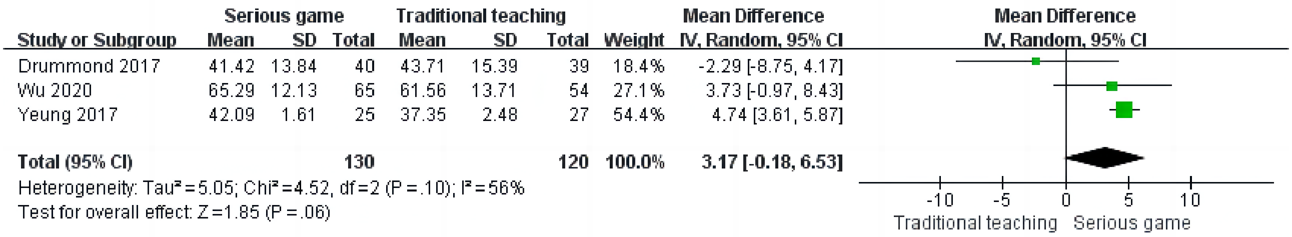
Figure 4. Meta-analysis of the effect of serious games on skill assessment.



The Effect of Serious Games Teaching on the Depth of CPR Compression

A total of 3 studies [36,37,39] presented findings on the impact of serious games on CPR compression depth. The assessment of heterogeneity demonstrated variability among the included studies (P=.10; I^2=56%), necessitating the application of a random effects model. The analysis depicted in Figure 5 [36,37,39] revealed that the disparity between the 2 groups did not reach statistical significance (MD 3.17, 95% CI –0.18 to 6.53; P=.06).

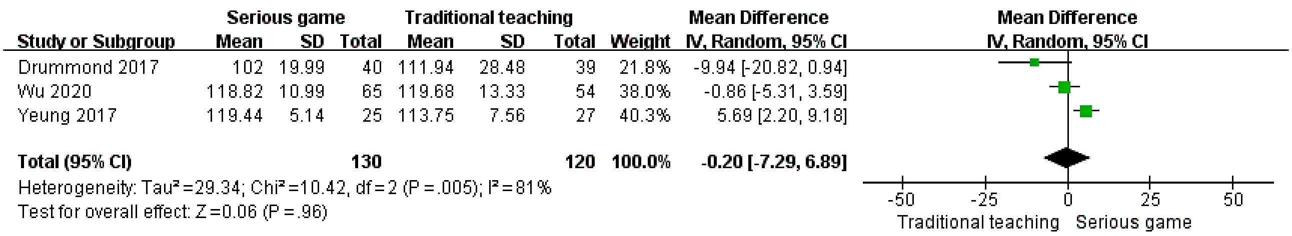
Figure 5. Meta-analysis of the effect of serious games on cardiopulmonary resuscitation (CPR) compression depth.



The Effect of Serious Games Teaching on the Frequency of CPR Compression

A meta-analysis was performed on 3 studies [36,37,39] that investigated the impact of serious games training on the frequency of CPR compression. Due to the variation among these studies (P=.005; I^2=81%), a random effects model was used. The results, as illustrated in Figure 6 [36,37,39], indicated that there was no significant difference in the theory of CPR compression rate between the 2 study groups (MD –0.20, 95% CI –7.29 to 6.89; P=.96).

Figure 6. Meta-analysis of the effect of serious games on cardiopulmonary resuscitation (CPR) compression frequency.



Sensitivity Analysis

We conducted separate analyses using both fixed effects and random effects models to examine the SMD, MD, and 95% CI of each model. By systematically excluding studies one by one, when the study by de Sena et al [38] was excluded, we observed a decrease in heterogeneity from 93% to 0% for theoretical assessment (Figure 7 [34,35,40-42]) and from 95% to 54% for skill assessment (Figure 8 [35,36,39,42]), respectively. This indicates that the study conducted by de Sena et al [38] may have contributed to the observed heterogeneity. In the meta-analysis of CPR compression depth, heterogeneity decreased from 56% to 0% after the exclusion of the study by Drummond et al [37], indicating that this study was the source of heterogeneity (Figure 9 [36,39]). After the exclusion of the study by Yeung et al [36], the heterogeneity of the meta-analysis on compression frequency of CPR decreased from 81% to 56%, indicating that this study was one of the sources of heterogeneity (Figure 10 [37,39]).

Figure 7. Sensitivity analysis of meta-analysis of cardiopulmonary resuscitation (CPR) theory performance.

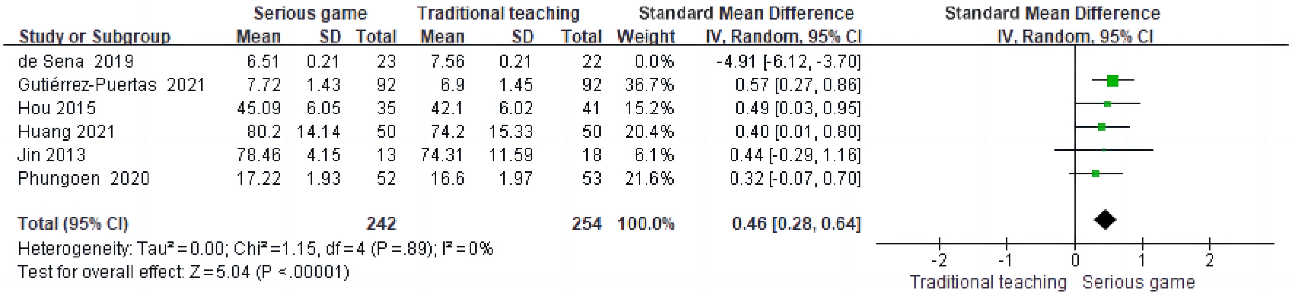


Figure 8. Sensitivity analysis of meta-analysis of cardiopulmonary resuscitation (CPR) skills operations.

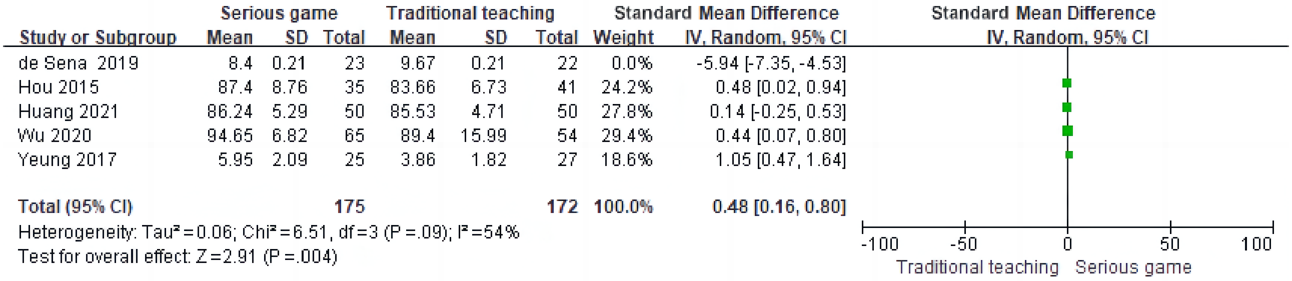


Figure 9. Sensitivity analysis of meta-analysis of cardiopulmonary resuscitation (CPR) compression depth.

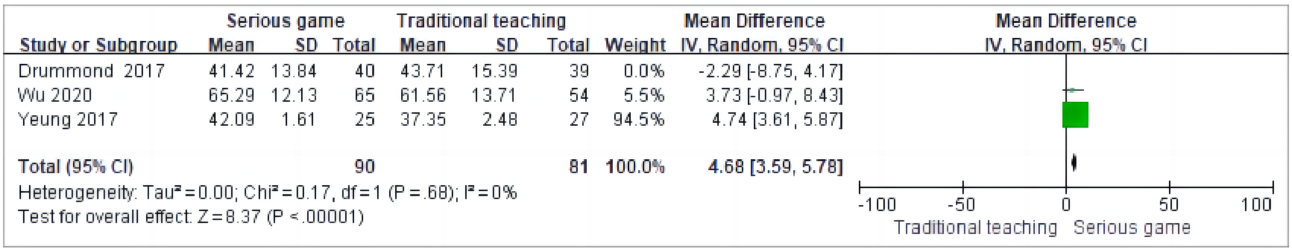
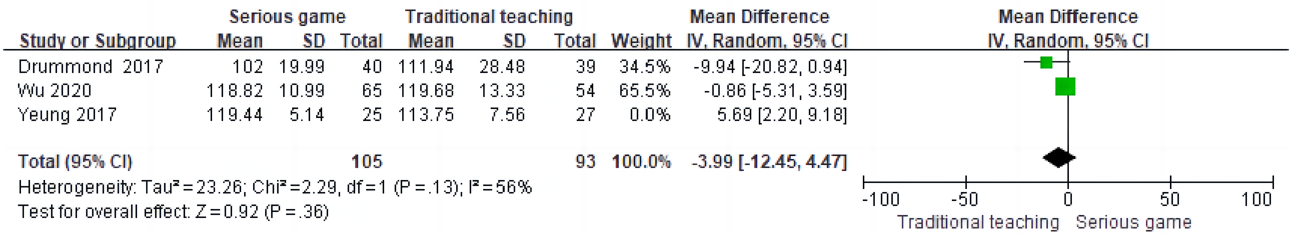


Figure 10. Sensitivity analysis of meta-analysis of cardiopulmonary resuscitation (CPR) compression frequency.



GRADE Evidence Quality Levels

Table S1 in Multimedia Appendix 3 presents the GRADE (Grading of Recommendations, Assessment, Development, and Evaluations) system evidence level for each outcome indicator in the meta-analysis of this study. The 4 outcome indicators considered were theory assessment, skill assessment, compression depth, and compression frequency.

Discussion

Principal Findings

This study systematically evaluated the efficacy of serious games-based training in CPR education, drawing upon data from 9 studies with a total of 791 participants. Our findings reveal no significant differences in theoretical exam scores, skill assessment scores, compression depth, or compression frequency between serious games-based and traditional CPR training methods. This suggests that serious games offer a highly effective alternative for CPR education. In alignment with the 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care [27], which recommend the incorporation of serious games into CPR education, our results underscore the positive impact of virtualized, gamified learning models on knowledge acquisition and CPR skill mastery. Theory and skills assessments are pivotal components of CPR training, serving as key indicators of training effectiveness and student proficiency. Our meta-analysis demonstrates that serious games-based CPR training is on par with traditional methods in enhancing both knowledge acquisition and skill levels. Consequently, serious games represent a valuable addition to the spectrum of CPR teaching and training methods, fostering innovation and aligning with the American Heart Association's guidelines for modernizing teaching tools and approaches.

Comparison With Previous Work

This study aligns with a previous meta-analysis [43], indicating that both lay and medical school students exhibit enhanced knowledge following web-based digital resuscitation training. Moreover, they demonstrate comparable cognitive outcomes to those undergoing traditional training sessions. The inclination of younger individuals toward serious games for acquiring new skills stems from their immersive and interactive nature, offering a secure trial-and-error environment [44]. This, coupled with engaging and positive learning experiences, reduces reliance on educational resources and fosters active, independent learning—especially when conventional training methods are inaccessible. This approach helps in sustaining knowledge levels, preventing decay over time, and attaining learning outcomes equivalent to traditional education forms [45]. Despite these benefits, serious games' applications for CPR training face challenges, presenting a mixed landscape concerning usability and enjoyment quality. Issues range from outdated guidelines and unupdated advice to overly detailed, professional information hindering learning efficacy. Such drawbacks may discourage public engagement with CPR learning [46]. For nonmedical learners, serious games must ensure acceptable usability, simplifying the comprehension and retention of CPR theoretical knowledge. Regular updates aligning with the latest guidelines can transform serious games into dynamic electronic textbooks [46,47]. To maximize the potential of serious games over traditional training, it is crucial to identify and evaluate functions that motivate learners to increase frequency and actively embrace knowledge updates. This strategic approach could position serious games as a superior alternative for enhancing the theoretical understanding of CPR, offering distinct advantages over traditional training methods [46,48].

Numerous guidelines [27,49-51] underscore the primary goal of CPR training: imparting participants with the skills necessary for high-quality CPR. This involves maintaining the correct compression rate and depth, ensuring thoracic recoil, and minimizing interruptions and hyperventilation. However, traditional training methods have presented challenges, particularly for nonmedical personnel [52,53], in mastering these vital competencies. Previous studies [52,54] have noted that simulated scenarios and repetitive practice often fall short of achieving adequate compression depth and frequency. Aksoy [55] and Siqueira et al [56] propose that a CPR teaching mode based on serious games could enhance learners' motivation and attitude, consequently improving compression quality. This study echoes Lau et al's [57] systematic review, indicating equivalence between serious games and traditional training methods in enhancing compression depth and frequency. However, electronic CPR training, including serious games, may not independently enhance skills without some influence from instructors, particularly for beginners. In other words, teacher involvement remains crucial to refining CPR skill performance through serious games training. Lim et al [25] discovered that content learned in serious games may not seamlessly transfer to skill operations during assessments, particularly for students with autonomous learning based on serious games. Scores in the pressing position, crucial for CPR quality, were notably worse than those in traditional training. Factors such as incorrect anatomical positions directly impact compression quality, making it challenging to achieve better performance in practical measures such as compression depth and frequency. While there was no significant difference in CPR compression skill or rate between the 2 training models, serious games-based CPR training revealed imperfections. To address this, integrating and emphasizing the impactful elements and advantageous attributes of traditional training into serious games may compensate for their shortcomings in skill practice. This approach has the potential to amplify the comparative advantages of serious games in CPR training.

In summary, the results of this study are similar to those of similar previous systematic reviews or studies. Nevertheless, given the limited number of studies included in this meta-analysis and the low GRADE evidence level, these results warrant cautious interpretation. Therefore, we recommend future CPR training efforts prioritize conducting high-quality, large-sample studies. This will enable a more comprehensive analysis of the effectiveness of serious games-based training, providing substantial evidence for the refinement of guidelines and the development of related teaching methodologies.

Strengths and Limitations

Strengths

This review compensates for the shortcomings of the previous literature in English by focusing on all types of serious games and conducting a comprehensive search of massive Chinese databases. Certainly, this study was conducted in strict accordance with highly recommended guidelines (ie, PRISMA), with early registration of the protocol for the systematic review and final grading of the evidence based on the GRADEpro GDT web-based tool, so it can be considered a robust, high-quality review. In addition, the meta-analysis conducted in this study involved 9 RCTs [34-42]. These RCTs provided detailed information on the study population, training protocol, serious games used, and measurement tools for outcome indicators. As for blinding implementation, it was challenging to blind interventionists due to the nature of CPR teaching training, which resulted in some degree of implementation bias. On the other hand, blinding the measurer effectively prevented measurement bias, particularly when assessing CPR theoretical knowledge and skills. Objective outcome indicators such as CPR compression depth and frequency, as recorded by the simulator, were less susceptible to measurement bias. The literature also addressed missed visits, had a low risk of selective reporting bias, and demonstrated baseline comparability between groups. Therefore, the included literature was of high quality, and the findings can be considered credible.

Limitations

This study acknowledges several limitations that merit consideration. First, our research only encompassed studies available in Chinese and English, which may introduce a linguistic bias. Second, heterogeneity in our meta-analysis results emerged due to variations in study populations, the use of different serious games, and diverse tools used to measure outcome indicators. Despite our efforts to explore the sources of heterogeneity through sensitivity analysis, a complete explanation remained elusive. In particular, it is worth noting that the use of different instruments by the included studies to evaluate training outcomes may have influenced the judgment of the results. Third, the relatively small number of included studies prevented us from conducting tests for publication bias. Additionally, some data underwent statistical transformations during the meta-analysis, potentially influencing the accuracy of the results. Lastly, this study focused primarily on CPR theory assessment, skill evaluation, compression depth, and compression rate as outcome indicators, without delving into knowledge and skill retention post-training, trainees' self-efficacy, or other facets of compression quality.

Implications for Future Research and Practice

Serious games, as an innovative model for CPR teaching and training, offer a promising avenue for first aid education, catering to diverse populations. However, this approach is still in its developmental and exploratory phases, and its cost-effectiveness warrants discussion. Future research should consider incorporating outcome indicators from the field of health economics to address economic barriers and promote the adoption of serious games in professional medical education and broader first aid training. Additionally, many studies lack standardized training specifications for serious games, including training duration, frequency, trainer intervention levels, and evaluation methods and tools for assessing training effectiveness. While serious games are recommended for CPR education, the specific details of this training mode require further standardization. Moreover, the quality of serious games, which serve as the platform for CPR training, significantly impacts training effectiveness. Developing serious games that align with international guidelines and cater to the diverse characteristics of trainees is undoubtedly challenging but essential. In conclusion, future research should prioritize conducting multicenter, large-sample RCTs to advance our understanding of the potential of serious games in CPR education.

Conclusion

This study conducted a meta-analysis of RCTs to assess the efficacy of serious games in CPR training. The findings indicate that serious games are equally effective as traditional training methods in enhancing CPR theory assessment and skill evaluation. Meanwhile, no significant differences emerged between serious games and traditional training methods regarding CPR compression depth and frequency. Notably, the current body of high-quality studies on serious games in CPR training is limited, often characterized by small sample sizes. Therefore, future research should prioritize conducting additional high-quality RCTs to provide further evidence and offer a more comprehensive understanding of the impact of serious games in CPR training and education.

Acknowledgments

This study was supported by the Hainan Provincial Natural Science Foundation of China (822MS071), the National Natural Science Foundation of China (81960420), and the Zhejiang Provincial Medical and Health Technology Project (2024KY1065).

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Authors' Contributions

PC conceived the idea for the review. PC, YH, and PY conducted the data curation, methodology, validation, and formal analysis and wrote the first draft of the manuscript. PC, YH, PY, HW, and BX were involved in the study selection, quality assessment, and data extraction. PC, BX, and CQ conducted the statistical analysis. HZ is responsible for the writing, methodology, conceptualization, supervision, and editing of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[DOCX File , 19 KB - games_v12i1e52990_app1.docx]

Multimedia Appendix 2

Detailed characteristics of the included trials.

[DOCX File , 22 KB - games_v12i1e52990_app2.docx]

Multimedia Appendix 3

Supplementary table S1.

[DOCX File , 15 KB - games_v12i1e52990_app3.docx]

Multimedia Appendix 4
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 101 KB - games_v12i1e52990_app4.pdf]

References

1. Morley PT, Atkins DL, Finn JC, Maconochie I, Nolan JP, Rabi Y, et al. Evidence evaluation process and management of potential conflicts of interest: 2020 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. Resuscitation 2020;156:A23-A34. [doi: 10.1016/j.resuscitation.2020.09.011] [Medline: 33099418]
2. Liang LD, Leung KHB, Chan TCY, Deakin J, Heidet M, Meckler G, et al. Pediatric and adult out-of-hospital cardiac arrest incidence within and near public schools in British Columbia: missed opportunities for systematic AED deployment strategies. Resuscitation 2022;181:20-25. [doi: 10.1016/j.resuscitation.2022.09.020] [Medline: 36208861]
3. Amacher SA, Blatter R, Briel M, Appenzeller-Herzog C, Bohren C, Becker C, et al. Predicting neurological outcome in adult patients with cardiac arrest: systematic review and meta-analysis of prediction model performance. Crit Care 2022;26(1):382 [FREE Full text] [doi: 10.1186/s13054-022-04263-y] [Medline: 36503620]
4. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. Circulation 2022;145(8):e153-e639 [FREE Full text] [doi: 10.1161/CIR.0000000000001052] [Medline: 35078371]
5. Perkins GD, Graesner JT, Semeraro F, Olasveengen T, Soar J, Lott C, et al. European Resuscitation Council Guidelines 2021: executive summary. Resuscitation 2021;161:1-60 [FREE Full text] [doi: 10.1016/j.resuscitation.2021.02.003] [Medline: 33773824]
6. Zhou G, Wang Y, Sun Z, Yuan M, Ma Y, Wu Q, et al. Survival outcome among patients with out-of-hospital cardiac arrest who received cardiopulmonary resuscitation in China: a systematic review and meta-analysis. Eur J Med Res 2023;28(1):8 [FREE Full text] [doi: 10.1186/s40001-022-00955-x] [Medline: 36600249]
7. Yoo KH, Cho Y, Oh J, Lee J, Ko BS, Kang H, et al. Association of socioeconomic status with long-term outcome in survivors after out-of-hospital cardiac arrest: nationwide population-based longitudinal study. JMIR Public Health Surveill 2023;9:e47156 [FREE Full text] [doi: 10.2196/47156] [Medline: 37432716]
8. Hasselqvist-Ax I, Riva G, Herlitz J, Rosenqvist M, Hollenberg J, Nordberg P, et al. Early cardiopulmonary resuscitation in out-of-hospital cardiac arrest. N Engl J Med 2015;372(24):2307-2315 [FREE Full text] [doi: 10.1056/NEJMoa1405796] [Medline: 26061835]
9. Baldi E, D'Alto A, Benvenuti C, Caputo ML, Cresta R, Cianella R, et al. Perceived threats and challenges experienced by first responders during their mission for an out-of-hospital cardiac arrest. Resusc Plus 2023;14:100403 [FREE Full text] [doi: 10.1016/j.resplu.2023.100403] [Medline: 37287957]
10. Boet S, Waldolf R, Bould C, Lam S, Burns JK, Moffett S, et al. Early or late booster for basic life support skill for laypeople: a simulation-based randomized controlled trial. CJEM 2022;24(4):408-418. [doi: 10.1007/s43678-022-00291-3] [Medline: 35438450]
11. Nguyen DT, Lauridsen KG, Krogh K, Løfgren B. Bystander performance using the 2010 vs 2015 ERC guidelines: a post-hoc analysis of two randomised simulation trials. Resusc Plus 2021;6:100123 [FREE Full text] [doi: 10.1016/j.resplu.2021.100123] [Medline: 34223381]
12. Zenani NE, Bello B, Molekodi M, Useh U. Effectiveness of school-based CPR training among adolescents to enhance knowledge and skills in CPR: a systematic review. Curationis 2022;45(1):e1-e9 [FREE Full text] [doi: 10.4102/curationis.v45i1.2325] [Medline: 36453813]
13. Böttiger BW, Van Aken H. Kids save lives—training school children in cardiopulmonary resuscitation worldwide is now endorsed by the World Health Organization (WHO). Resuscitation 2015;94:A5-A7. [doi: 10.1016/j.resuscitation.2015.07.005] [Medline: 26209417]
14. Kiguchi T, Okubo M, Nishiyama C, Maconochie I, Ong MEH, Kern KB, et al. Out-of-hospital cardiac arrest across the world: first report from the International Liaison Committee on Resuscitation (ILCOR). Resuscitation 2020;152:39-49. [doi: 10.1016/j.resuscitation.2020.02.044] [Medline: 32272235]
15. Oo AM, Vallabhajosyula R. Perceived effectiveness of an innovative mobile-based serious game on the improvement of soft skills in minimally invasive surgical training. Asian J Endosc Surg 2023;16(1):41-49. [doi: 10.1111/ases.13115] [Medline: 36594159]
16. Ghoman SK, Patel SD, Cutumisu M, von Hauff P, Jeffery T, Brown MRG, et al. Serious games, a game changer in teaching neonatal resuscitation? A review. Arch Dis Child Fetal Neonatal Ed 2020;105(1):98-107 [FREE Full text] [doi: 10.1136/archdischild-2019-317011] [Medline: 31256010]
17. Otero-Agra M, Barcala-Furelos R, Besada-Saavedra I, Peixoto-Pino L, Martínez-Isasi S, Rodríguez-Núñez A. Let the kids play: gamification as a CPR training methodology in secondary school students. A quasi-experimental manikin simulation study. Emerg Med J 2019;36(11):653-659. [doi: 10.1136/emermed-2018-208108] [Medline: 31409636]

18. Warsinsky S, Schmidt-Kraepelin M, Rank S, Thiebes S, Sunyaev A. Conceptual ambiguity surrounding gamification and serious games in health care: literature review and development of Game-Based Intervention Reporting Guidelines (GAMING). J Med Internet Res 2021;23(9):e30390 [FREE Full text] [doi: 10.2196/30390] [Medline: 34505840]
19. Perron JE, Coffey MJ, Lovell-Simons A, Dominguez L, King ME, Ooi CY. Resuscitating cardiopulmonary resuscitation training in a virtual reality: prospective interventional study. J Med Internet Res 2021;23(7):e22920 [FREE Full text] [doi: 10.2196/22920] [Medline: 34326040]
20. Creutzfeldt J, Hedman L, Heinrichs L, Youngblood P, Felländer-Tsai L. Cardiopulmonary resuscitation training in high school using avatars in virtual worlds: an international feasibility study. J Med Internet Res 2013;15(1):e9 [FREE Full text] [doi: 10.2196/jmir.1715] [Medline: 23318253]
21. Tan JW, Zary N. Diagnostic markers of user experience, play, and learning for digital serious games: a conceptual framework study. JMIR Serious Games 2019;7(3):e14620 [FREE Full text] [doi: 10.2196/14620] [Medline: 31313660]
22. Taramarcaz V, Herren T, Golay E, Regard S, Martin-Achard S, Mach F, et al. A short intervention and an interactive e-learning module to motivate medical and dental students to enlist as first responders: implementation study. J Med Internet Res 2022;24(5):e38508 [FREE Full text] [doi: 10.2196/38508] [Medline: 35583927]
23. Fijačko N, Masterson Creber R, Gosak L, Štiglic G, Egan D, Chaka B, et al. Evaluating quality, usability, evidence-based content, and gamification features in mobile learning apps designed to teach children basic life support: systematic search in app stores and content analysis. JMIR Mhealth Uhealth 2021;9(7):e25437 [FREE Full text] [doi: 10.2196/25437] [Medline: 34283034]
24. Chong KM, Yang HW, He HC, Lien WC, Yang MF, Chi CY, et al. The effectiveness of online-only blended cardiopulmonary resuscitation training: static-group comparison study. J Med Internet Res 2023;25:e42325 [FREE Full text] [doi: 10.2196/42325] [Medline: 37018023]
25. Lim XMA, Liao WA, Wang W, Seah B. The effectiveness of technology-based cardiopulmonary resuscitation training on the skills and knowledge of adolescents: systematic review and meta-analysis. J Med Internet Res 2022;24(12):e36423 [FREE Full text] [doi: 10.2196/36423] [Medline: 36520524]
26. Naylor K, Torres K. Transitioning to web-based learning in basic life support training during the COVID-19 pandemic to battle the fear of out-of-hospital cardiac arrest: presentation of novel methods. J Med Internet Res 2021;23(5):e27108 [FREE Full text] [doi: 10.2196/27108] [Medline: 33886488]
27. Cheng A, Magid DJ, Auerbach M, Bhanji F, Bigham BL, Blewer AL, et al. Part 6: resuscitation education science: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020;142(16_suppl_2):S551-S579 [FREE Full text] [doi: 10.1161/CIR.0000000000000903] [Medline: 33081527]
28. Dankbaar M. Serious games and blended learning; effects on performance and motivation in medical education. Perspect Med Educ 2017;6(1):58-60 [FREE Full text] [doi: 10.1007/s40037-016-0320-2] [Medline: 27975195]
29. Cheng A, Nadkarni VM, Mancini MB, Hunt EA, Sinz EH, Merchant RM, et al. Resuscitation education science: educational strategies to improve outcomes from cardiac arrest: a scientific statement from the American Heart Association. Circulation 2018;138(6):e82-e122 [FREE Full text] [doi: 10.1161/CIR.0000000000000583] [Medline: 29930020]
30. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
31. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
32. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res 2018;27(6):1785-1805. [doi: 10.1177/0962280216669183] [Medline: 27683581]
33. Higgins JPT, Green S, The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. Chichester: John Wiley & Sons, Inc; 2011.
34. Jin XQ, Zhao Y, Chen ZQ, Jiang S, Lei H, Fan JY. Exploration on autonomous game learning method in cardiopulmonary cerebral resuscitation. Chin J Med Edu 2013;33(6):906-908. [doi: 10.3760/cma.j.issn.1673-677X.2013.06.036]
35. Hou WW, Lu YL, Deng RL, Sun LL, Wang X, Guo LZ, et al. Application of the educational game software in the drowning first aid training for college students. Chin J Nurs Educ 2015;12(8):590-593. [doi: 10.3761/j.issn.1672-9234.2015.08.009]
36. Yeung J, Kovic I, Vidacic M, Skilton E, Higgins D, Melody T, et al. The school Lifesavers study-a randomised controlled trial comparing the impact of Lifesaver only, face-to-face training only, and Lifesaver with face-to-face training on CPR knowledge, skills and attitudes in UK school children. Resuscitation 2017;120:138-145. [doi: 10.1016/j.resuscitation.2017.08.010] [Medline: 28974344]
37. Drummond D, Delval P, Abdenouri S, Truchot J, Ceccaldi PF, Plaisance P, et al. Serious game versus online course for pretraining medical students before a simulation-based mastery learning course on cardiopulmonary resuscitation: a randomised controlled study. Eur J Anaesthesiol 2017;34(12):836-844 [FREE Full text] [doi: 10.1097/EJA.0000000000000675] [Medline: 28731928]
38. de Sena DP, Fabrício DD, da Silva VD, Bodanese LC, Franco AR. Comparative evaluation of video-based on-line course versus serious game for training medical students in cardiopulmonary resuscitation: a randomised trial. PLoS One 2019;14(4):e0214722 [FREE Full text] [doi: 10.1371/journal.pone.0214722] [Medline: 30958836]

39. Wu RY, Qin CZ, He CY, Ma QX, Zhou Q, Zhang L. The effectiveness of a game-based CPR training in high school students. China Journal of Emergency Resuscitation and Disaster Medicine 2020;15(10):1163-1167. [doi: 10.3969/j.issn.1673-6966.2020.10.007]
40. Phungoen P, Promto S, Chanthawatthanarak S, Maneepong S, Apiratwarakul K, Kotruchin P, et al. Precourse preparation using a serious smartphone game on advanced life support knowledge and skills: randomized controlled trial. J Med Internet Res 2020;22(3):e16987 [FREE Full text] [doi: 10.2196/16987] [Medline: 32149711]
41. Gutiérrez-Puertas L, García-Viola A, Márquez-Hernández VV, Garrido-Molina JM, Granados-Gámez G, Aguilera-Manrique G. Guess it (SVUAL): an app designed to help nursing students acquire and retain knowledge about basic and advanced life support techniques. Nurse Educ Pract 2021;50:102961. [doi: 10.1016/j.nepr.2020.102961] [Medline: 33421681]
42. Huang D, Chen YR, Jiang ZX, Zhou J, Wang Y, Dai YY. Effects of CPR educational games on CPR knowledge acquisition and retention in nursing undergraduate students. J Nurs Sci 2021;36(16):59-62. [doi: 10.3870/j.issn.1001-4152.2021.16.059]
43. Hsieh MJ, Bhanji F, Chiang WC, Yang CW, Chien KL, Ma MHM. Comparing the effect of self-instruction with that of traditional instruction in basic life support courses-a systematic review. Resuscitation 2016;108:8-19. [doi: 10.1016/j.resuscitation.2016.08.021] [Medline: 27581252]
44. Aksoy ME, Özkan AE, Kitapcioglu D, Usseli T. Comparing the outcomes of virtual reality-based serious gaming and lecture-based training for advanced life support training: randomized controlled trial. JMIR Serious Games 2023;11:e46964 [FREE Full text] [doi: 10.2196/46964] [Medline: 37768719]
45. Akbari F, Nasiri M, Rashidi N, Zonoori S, Amirmohseni L, Eslami J, et al. Comparison of the effects of virtual training by serious game and lecture on operating room novices' knowledge and performance about surgical instruments setup: a multi-center, two-arm study. BMC Med Educ 2022;22(1):268 [FREE Full text] [doi: 10.1186/s12909-022-03351-5] [Medline: 35410279]
46. Kalz M, Lenssen N, Felzen M, Rossaint R, Tabuenca B, Specht M, et al. Smartphone apps for cardiopulmonary resuscitation training and real incident support: a mixed-methods evaluation study. J Med Internet Res 2014;16(3):e89 [FREE Full text] [doi: 10.2196/jmir.2951] [Medline: 24647361]
47. Metelmann B, Metelmann C, Schuffert L, Hahnenkamp K, Brinkrolf P. Medical correctness and user friendliness of available apps for cardiopulmonary resuscitation: systematic search combined with guideline adherence and usability evaluation. JMIR Mhealth Uhealth 2018;6(11):e190 [FREE Full text] [doi: 10.2196/mhealth.9651] [Medline: 30401673]
48. Ahn C, Cho Y, Oh J, Song Y, Lim TH, Kang H, et al. Evaluation of smartphone applications for cardiopulmonary resuscitation training in South Korea. Biomed Res Int 2016;2016:6418710 [FREE Full text] [doi: 10.1155/2016/6418710] [Medline: 27668257]
49. Merchant RM, Topjian AA, Panchal AR, Cheng A, Aziz K, Berg KM, et al. Part 1: executive summary: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020;142(16_suppl_2):S337-S357 [FREE Full text] [doi: 10.1161/CIR.0000000000000918] [Medline: 33081530]
50. Semeraro F, Frisoli A, Loconsole C, Mastronicola N, Stroppa F, Ristagno G, et al. Kids (learn how to) save lives in the school with the serious game Relive. Resuscitation 2017;116:27-32. [doi: 10.1016/j.resuscitation.2017.04.038] [Medline: 28476478]
51. Böttiger BW, Lockey A, Georgiou M, Greif R, Monsieurs K, Mpotos N, et al. KIDS SAVE LIVES: ERC position statement on schoolteachers' education and qualification in resuscitation. Resuscitation 2020;151:87-90 [FREE Full text] [doi: 10.1016/j.resuscitation.2020.04.021] [Medline: 32339597]
52. Chamdawala H, Meltzer JA, Shankar V, Elachi D, Jarzynka SM, Nixon AF. Cardiopulmonary resuscitation skill training and retention in teens (CPR START): a randomized control trial in high school students. Resusc Plus 2021;5:100079 [FREE Full text] [doi: 10.1016/j.resplu.2021.100079] [Medline: 34223345]
53. Minna S, Leena H, Tommi K. How to evaluate first aid skills after training: a systematic review. Scand J Trauma Resusc Emerg Med 2022;30(1):56 [FREE Full text] [doi: 10.1186/s13049-022-01043-z] [Medline: 36348427]
54. Lin Y, Cheng A, Grant VJ, Currie GR, Hecker KG. Improving CPR quality with distributed practice and real-time feedback in pediatric healthcare providers—a randomized controlled trial. Resuscitation 2018;130:6-12. [doi: 10.1016/j.resuscitation.2018.06.025] [Medline: 29944894]
55. Aksoy ME. Comparing basic life support serious gaming scores with hands-on training platform performance scores: pilot simulation study for basic life support training. JMIR Serious Games 2020;8(4):e24166 [FREE Full text] [doi: 10.2196/24166] [Medline: 33237035]
56. Siqueira TV, da Silva Garcia Nascimento J, de Oliveira JLG, da Silva Garcia Regino D, Dalri MCB. The use of serious games as an innovative educational strategy for learning cardiopulmonary resuscitation: an integrative review. Rev Gaucha Enferm 2020;41:e20190293 [FREE Full text] [doi: 10.1590/1983-1447.2020.20190293] [Medline: 33111759]
57. Lau Y, Nyoe RSS, Wong SN, Ab Hamid ZB, Leong BSH, Lau ST. Effectiveness of digital resuscitation training in improving knowledge and skills: a systematic review and meta-analysis of randomised controlled trials. Resuscitation 2018;131:14-23. [doi: 10.1016/j.resuscitation.2018.07.033] [Medline: 30071263]

Abbreviations

AED: automated external defibrillator
CPR: cardiopulmonary resuscitation
EMS: emergency medical services
FR: first responder
MD: mean difference
OHCA: out-of-hospital cardiac arrest
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
SMD: standardized mean difference

Edited by T Leung; submitted 24.09.23; peer-reviewed by Z Zhang, L Suppan; comments to author 11.11.23; revised version received 07.12.23; accepted 30.12.23; published 06.02.24.

Please cite as:
Cheng P, Huang Y, Yang P, Wang H, Xu B, Qu C, Zhang H
The Effects of Serious Games on Cardiopulmonary Resuscitation Training and Education: Systematic Review With Meta-Analysis of Randomized Controlled Trials
JMIR Serious Games 2024;12:e52990
URL: https://games.jmir.org/2024/1/e52990
doi:10.2196/52990
PMID:38319697

Review

Electronic Interactive Games for Glycemic Control in Individuals With Diabetes: Systematic Review and Meta-Analysis

WenQi Yao[1*], BM; YiBing Han[1*], MM; Li Yang[1,2*], MD; Ying Chen[1], MM; ShengZhe Yan[1], BM; YanZhen Cheng[1], MD

[1]Department of Endocrinology, Zhujiang Hospital of Southern Medical University, Guangzhou, China
[2]Department of Nutrition, Zhujiang Hospital of Southern Medical University, Guangzhou, China
[*]these authors contributed equally

Corresponding Author:
YanZhen Cheng, MD
Department of Endocrinology
Zhujiang Hospital of Southern Medical University
No.253 Gongye Avenue Central
Guangzhou, 510280
China
Phone: 86 0 206 278 2330
Email: chengyzx@163.com

Abstract

Background: Several electronic interventions have been used to improve glycemic control in patients with diabetes. Electronic interactive games specific to physical activity are available, but it is unclear if these are effective at improving glycemic control in patients with diabetes.

Objective: This study aimed to determine the effects of electronic game–based interventions on glycemic control in patients with diabetes.

Methods: Relevant studies that were published before April 1, 2023, were searched from 5 databases: PubMed, Embase, Web of Science, Scopus, and Cochrane Library. Eligibility criteria included prospective studies examining the relationship between electronic games with physical activities or diet education and glycemic control as the outcome. The risk of bias was assessed using the Cochrane risk-of-bias tool. All analyses were conducted using RevMan5.4.1. Depending on the heterogeneity across studies, the pooled effects were calculated using fixed-effects or random-effects models.

Results: Participants from 9 studies were included and assessed. Glycated hemoglobin (HbA_{1c}) and fasting blood glucose improved in the intervention group, although the analysis revealed no significant reduction in HbA_{1c} (−0.09%, 95% CI −0.29% to 0.10%) or fasting blood glucose (−0.94 mg/dL, 95% CI −9.34 to 7.46 mg/dL). However, the physical activity of individuals in the intervention group was significantly higher than that of those in the control group (standardized mean difference=0.84, 95% CI 0.30 to 1.38; P=.002). Other outcomes, such as weight and blood lipids, exhibited no significant improvement (all P>.05).

Conclusions: Electronic games had a good impact on participants' physical activity and offered an advantage in glycemic control without reaching statistical significance. Electronic games are convenient for reminders and education. Low-intensity exercise games may not be considered a better adjuvant intervention to improve diabetes self-management care.

(JMIR Serious Games 2024;12:e43574) doi:10.2196/43574

KEYWORDS
electronic game; physical activity; diet; diabetes mellitus; glycemic control

Introduction

Diabetes mellitus is one of the 4 major noncommunicable diseases and is also among the top 10 global causes of death. Throughout the world, the number of patients with diabetes mellitus is increasing, probably due to changes in lifestyle. According to the International Diabetes Federation, in 2021, approximately 536.6 million adults (aged 20-79 years) were living with diabetes; this is expected to rise to 12.2% in 2045 [1]. Because of the rise in type 1 and type 2 diabetes, the burden of health care expenditures and its complications continues to increase, whereas the complications are the main causes of morbidity and mortality [2]. To address the health challenge resulting from diabetes, effective and efficient management is needed [3-5].

Lifestyle management, an efficacious method for diabetes prevention [6], is a fundamental aspect of diabetes care. It includes diabetes self-management education and support, medical nutrition therapy, physical activity, smoking cessation counseling, and psychosocial care [7]. Food intake and physical activity are associated with significantly improved control of diabetes [8]. With advances in technology, lifestyle management incorporating novel technologies and formats meets the needs of various populations for diabetes treatment [9]. New methods, such as electronic games and wearable devices, aim to contribute to better patient compliance [10].

It has been reported that electronic games can help players learn more about healthy diets and encourage exercise [11,12]. Although they play a role as facilitators in motivating and accelerating physical activity, they offer little benefit to patients with chronic disease [13]. Previous systematic reviews have evaluated the impact of app-based or electronic health interventions to support changes in blood glucose management, physical activity, or diet [9,14,15]. However, previous papers analyzed relatively few articles or articles that were not solely on using games. They also used educational or regulation applications, robots, or virtual worlds that do not contain game elements. Electronic games specific to physical activity and dietary education are available; however, we currently lack an understanding of how effective electronic games can be for glycemic control.

In this study, we performed a comprehensive literature search to select studies on the effects of electronic game–based interventions on glycemic control in patients with diabetes for meta-analysis. Electronic gaming interventions are defined as containing an element of gaming that involves virtual reality, serious gaming, or exergaming [15].

Methods

Data Sources and Search Strategy

This review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement and its associated checklist (Multimedia Appendix 1). Relevant studies that were published before April 1, 2023, were searched from 5 databases: PubMed, Embase, Web of Science, Scopus, and Cochrane Library. The references of the included studies were hand-searched to identify any additional articles. The following terms were used during the search: (“Diabetes” OR “diabetic” OR “diabetes mellitus” OR “glycemic control” OR “glucose control” OR “glucose”) AND (“game” OR “gamification” OR “exergaming” OR “avatar” OR “wii” OR “virtual” OR “konami” OR “wii-fit” OR “kinect” OR “tierone” OR “video-game” OR “serious-games” OR “serious video-games” OR “Augmented reality” OR “mixed reality” OR “second life” OR “TierOne” OR “Konami Dance Dance Revolution” OR “Sony Eyetoy” OR “Microsoft Kinec”). Detailed search strategies for each database are given in Multimedia Appendix 2. The reference lists of the searched articles and the relevant reviews were then screened to identify any pertinent studies.

Study Selection

Studies included in this meta-analysis met the following criteria: (1) participants were diagnosed as having type 1 diabetes or type 2 diabetes; (2) the articles were published in English or Chinese; (3) the articles presented the electronic management intervention with a gaming element, such as a virtual reality game, serious game, or exergame; and (4) the outcome indicators were blood glucose and glycated hemoglobin (HbA_{1c}).

Studies that met the following criteria were excluded: (1) participants had gestational diabetes mellitus, had other special types of diabetes mellitus, underwent surgery, had an operation, or were in the emergency department; (2) participants had a previous history of mental illnesses, eating disorders, or cancer; (3) the management intervention was only based on an online, mobile, or virtual application but did not use a gaming element; and (4) articles that were protocols, conference abstracts, case reports, reviews, or meta-analyses.

Articles were screened in a 2-step process. First, all titles and abstracts were examined by 2 investigators. Any citations that clearly did not meet the inclusion criteria were excluded. Second, all abstracts and full-text articles were examined independently by 2 investigators. Any disagreements in the selection process were resolved through discussion with a third investigator.

Risk of Bias

The included trials were independently assessed by 2 investigators for the risk of bias using the Cochrane risk-of-bias tool [16]. An assessment was performed across 5 domains of bias (sequence generation, allocation concealment, blinding, incomplete outcome data, and selective reporting). The risk of bias was assessed as either low (proper methods taken to reduce bias), high (improper methods creating bias), or unclear (insufficient information provided to determine the bias level). All discrepancies and disagreements were resolved through consensus or, where necessary, by a third author.

Data Extraction

A Microsoft Excel table was used to extract data on the year of publication, country, sample size, participant characteristics, study setting and design, intervention and control arms, duration, and outcome data. The main outcomes included HbA_{1c} or fasting blood glucose (FBG). The secondary outcomes included daily steps (regarded as a physical activity outcome), blood pressure, and weight, among others. The data were obtained from the original text and attachments supplied. Data from different studies were converted to common units. Data extraction was carried out by 2 reviewers independently. All discrepancies and disagreements were resolved through consensus.

Missing Data

Study authors were contacted by email where there were missing or unclear data (for instance, relating to the primary outcome). Studies for which insufficient primary data were available (eg, missing data cannot be obtained) were excluded from analysis but not from the review.

Data Synthesis and Quality Assessment

All analyses were conducted using RevMan5.4.1 (Cochrane). Data were expressed as the mean difference (MD) and 95% CI and pooled using fixed-effect or random-effects models according to the heterogeneity. A random-effects model assumes that the study estimates are estimating different, yet related, intervention effects and thus incorporates the heterogeneity among studies. This is a more appropriate method to pool studies that may differ slightly in the distribution of risk factors, population, size, and outcomes.

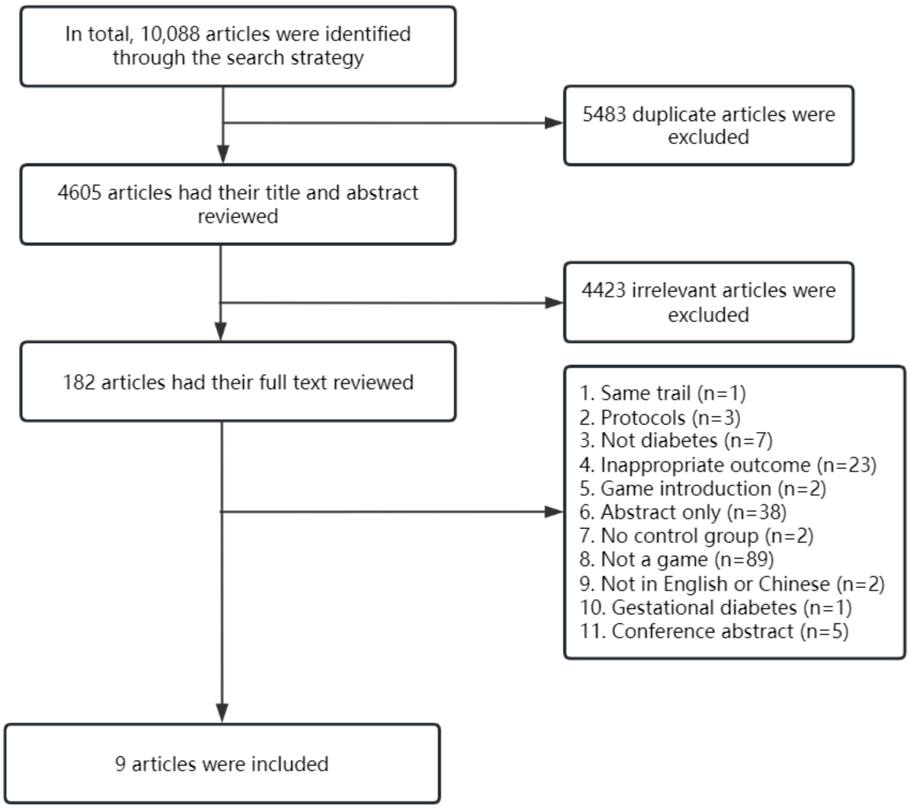
Heterogeneity was assessed using a χ^2 test and quantified using the I^2 statistic. Significance for the heterogeneity was set at $P<.05$, with an $I^2>50\%$ considered to be evidence of high heterogeneity, which prompted us to use the random-effects model to pool the data.

Results

Overview

Our search identified 10,088 articles, of which 4605 were screened after removing duplicate records. Of these, 182 were identified for further evaluation. Of these, 173 were excluded, resulting in 9 included studies (Figure 1). Of the excluded articles, 18 were excluded because they only had abstracts and we could not access the original text and data.

Figure 1. Flowchart of the study selection process.



The results of the remaining 9 studies, comprising 913 participants and 747 cases of type 2 diabetes, were included in the meta-analysis [17-25]. The characteristics of all 9 studies are shown in Multimedia Appendix 3 [17-25]. The duration of trials ranged from 1 month to 1 year. Of the 9 studies, 4 were undertaken in the United States [18,21,23,25], 2 in Europe [17,20], and 3 in Asia [19,22,24]. Of the 9 studies, 3 assessed FBG and 8 assessed HbA_{1c}.

The studies included 2 non–randomized controlled trials (RCTs) and 7 RCTs, the quality of which was assessed using the Cochrane risk-of-bias tool. We determined that 3 studies were of high quality, whereas 4 were of moderate quality and 2 were of low quality (Figures 2 and 3 [17-25]). The 2 non-RCTs were not random and the allocations were unclear. Blinding was difficult in game interventions; 1 study was unblinded [25] and 4 were unclear, but the studies made an effort to blind either patients or personnel. One study was not blinded to the outcome assessment, but it was still analyzed as low risk, considering its main outcome was the objective index. The 9 studies had no elective outcome reporting.

Figure 2. Risk-of-bias graph showing the authors' judgments about each risk-of-bias item presented as percentages across all the included studies. A total of 9 trials were assessed for risk of bias.

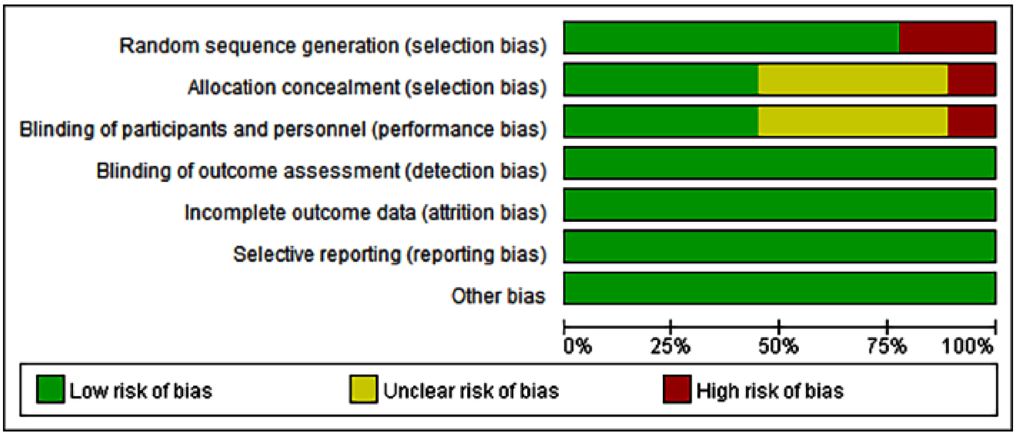
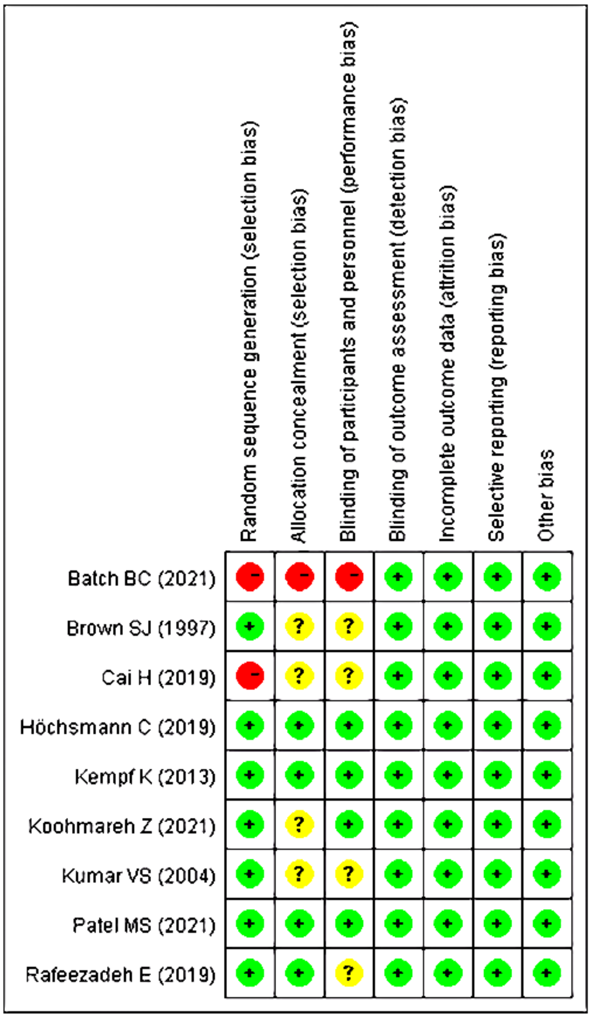


Figure 3. Risk-of-bias summary showing the authors' judgments about each risk-of-bias item for each included study. Green "+": low risk of bias; red "–": high risk of bias; yellow "?": unknown risk of bias.



Publication bias was not assessed for any outcome as <10 trials were available.

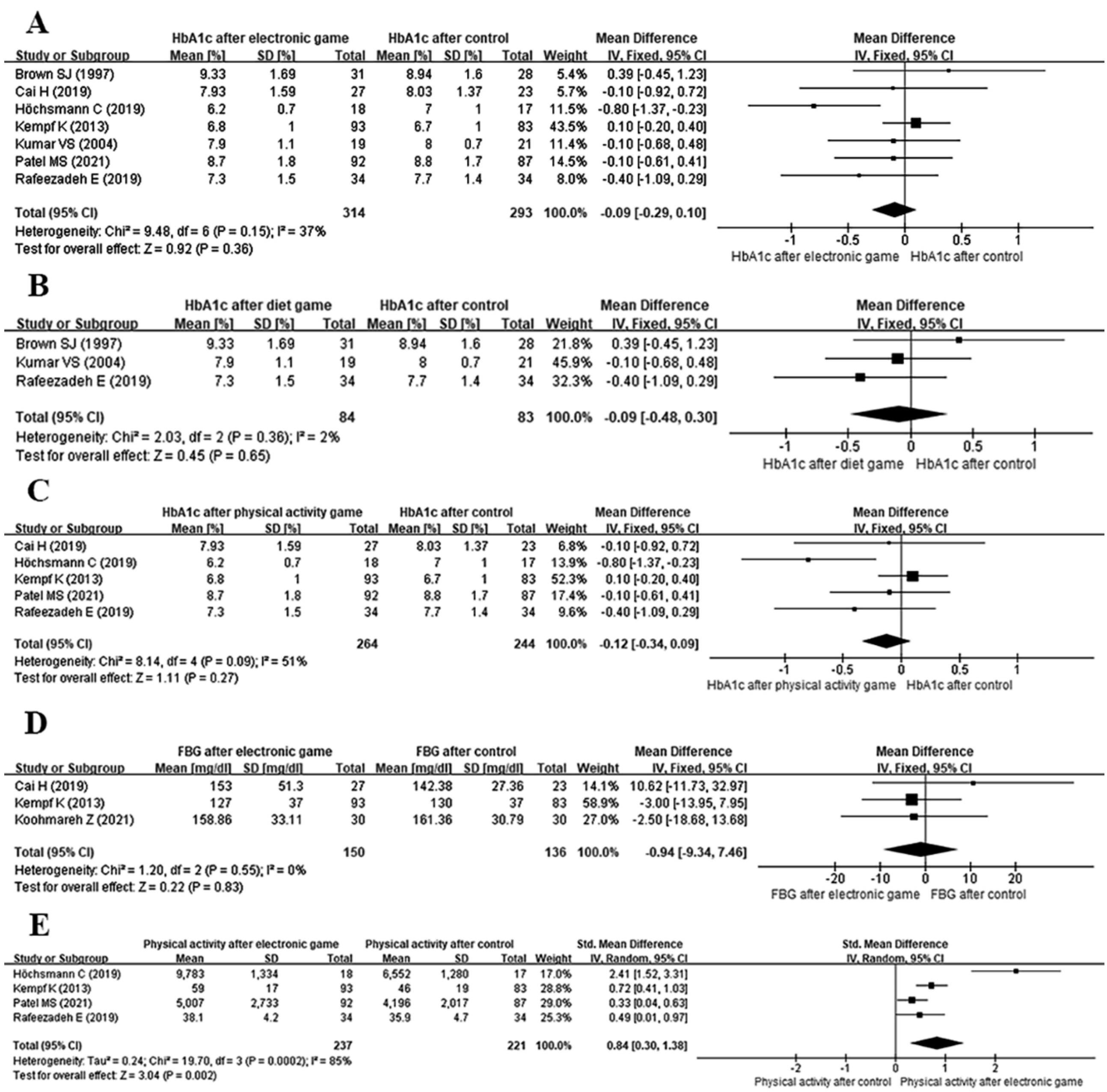
Meta-Analysis

HbA_{1c} Level

A total of 8 articles had HbA_{1c} testing but 1 did not provide postintervention data [25]. We sent an email to the author with a request to provide the raw data but received no reply.

As shown in Figure 4A [17,21,23-25], this analysis showed a clinically important improvement in HbA_{1c}, but there was no significant reduction after the intervention among individuals with diabetes mellitus (7 studies; n=607; MD=−0.09%, 95% CI −0.29% to 0.10%; I^2=37%; P=.36). Figure 4B shows the change in HbA_{1c} after a diet-based game intervention (3 studies; n=167; MD=−0.09%, 95% CI −0.48% to 0.30%; I^2=2%; P=.65). Figure 4C shows the change in HbA_{1c} after a physical activity–based game intervention (5 studies; n=508, MD=−0.12%, 95% CI −0.34% to 0.09%; I^2=51%; P=.27).

Figure 4. Meta-analysis of the effect of electronic games on HbA_{1c}, FBG, and physical activity. (A) HbA_{1c} after a diet intervention or physical activity intervention; (B) HbA_{1c} after a diet intervention; (C) HbA_{1c} after a physical activity game intervention; (D) FBG after an electronic game intervention; and (E) physical activity after an electronic game intervention. FBG: fasting blood glucose; HbA_{1c}: glycated hemoglobin; IV: inverse variance; Std.: standardized.



Fasting Blood Glucose Level

The meta-analysis showed that the FBG level of the intervention groups was not statistically different from that of the control groups (3 studies; n=286; MD=−0.94 mg/dL, 95% CI −9.34 to 7.46 mg/dL; I^2=0%, P=.83; Figure 4D).

Physical Activity

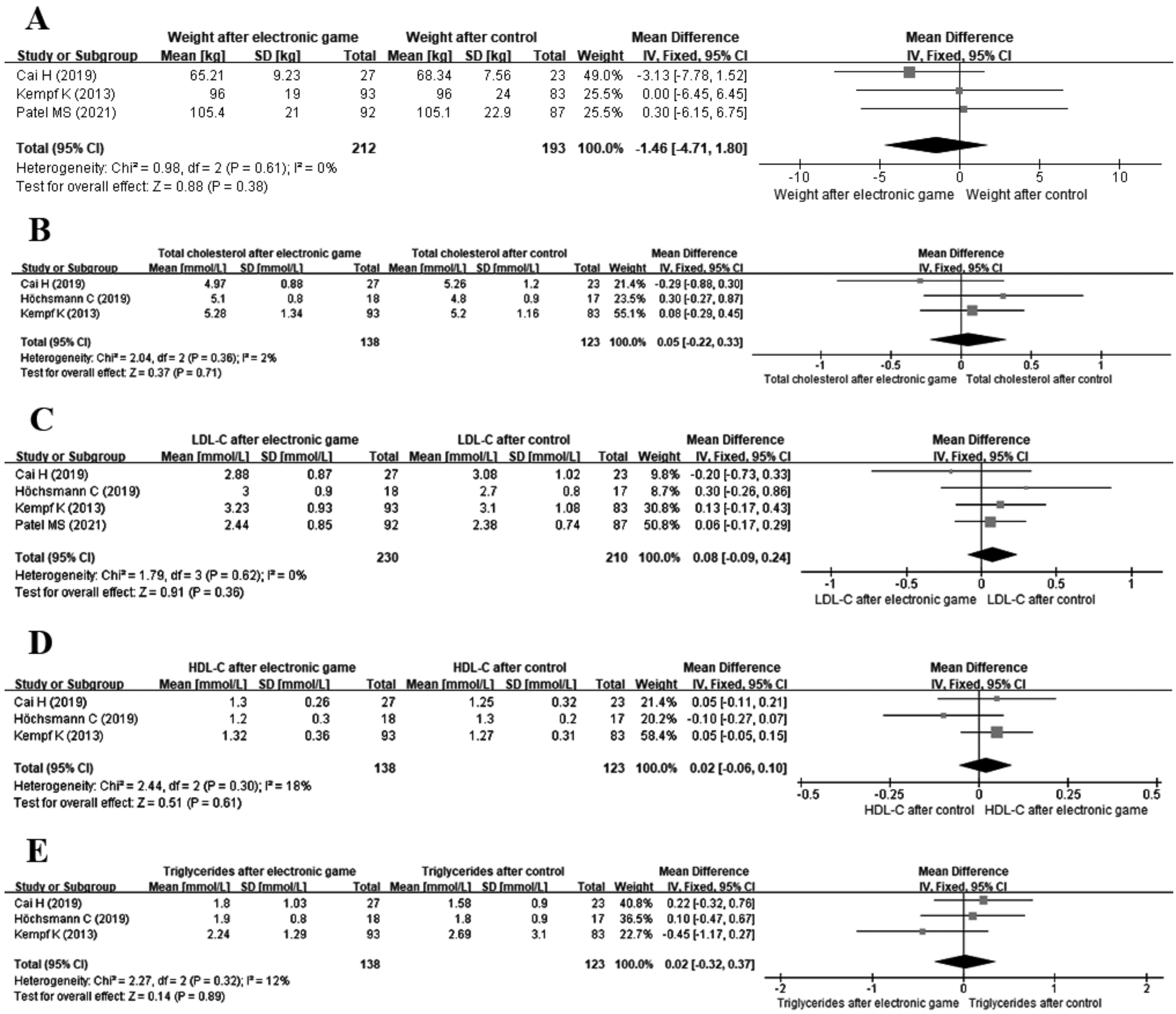
Of the 7 RCTs, 2 assessed self-reported physical activity and 2 counted participants' daily steps during the intervention to assess the patients' physical activity. Because of the differences in measurement instruments, we calculated standardized mean differences (SMDs). These results were statistically heterogeneous with respect to the effect (χ^2_3=19.70; P<.001; I^2=85%); we found a significant increase in physical activity above baseline in the intervention groups. Moreover, participants assigned to the intervention groups increased their physical activity significantly more than participants in the control groups (SMD=0.84; 95% CI 0.30 to 1.38; P=.002; Figure 4E).

Weight

Weight also trended toward decreases in the intervention groups, with an MD of −1.46 kg (95% CI −4.71 to 1.80 kg; Figure 5A

[17-19]). However, the decreases did not reach statistical significance (P=.38).

Figure 5. Meta-analysis of the effect of electronic games on (A) weight, (B) total cholesterol, (C) LDL-C, (D) HDL-C, and (E) triglycerides. HDL-C: high-density lipoprotein cholesterol; IV: inverse variance; LDL-C: low-density lipoprotein cholesterol; Std.: standardized.



Blood Lipids

There was no significant reduction in total cholesterol (3 studies; n=261; MD=0.05 mmol/L, 95% CI −0.22 to 0.33 mmol/L; I^2=2%; P=.71; Figure 5B), low-density lipoprotein cholesterol (4 studies; n=440; MD=0.08 mmol/L, 95% CI −0.09 to 0.24 mmol/L; I^2=0%; P=.36; Figure 5C), high-density lipoprotein cholesterol (3 studies; n=261; MD=0.02 mmol/L, 95% CI −0.06 to 0.10 mmol/L; I^2=18%; P=.61; Figure 5D), or triglycerides (3 studies; n=261; MD=0.02 mmol/L, 95% CI −0.32 to 0.37 mmol/L; I^2=12%; P=.89; Figure 5E) after the intervention among patients with diabetes mellitus.

Discussion

Principal Findings

This study demonstrated that electronic interactive games were associated with a good impact on participants' physical activity. However, we found that electronic interactive games did not present a significant benefit for HbA_{1c} levels, FBG levels, weight, or blood lipids compared to the control group. The game interventions were intended for education to manage diabetes through games.

Effects of Diet Education Games on Blood Glucose

Plant-based diets and exercise are major diabetes-protective factors [26]. The Da Qing Diabetes Prevention Study showed an overall 51% reduction in diabetes incidence in participants after a 6-year intervention with diet, exercise, or both; its 30-year follow-up showed that lifestyle interventions reduced the incidence of serious diabetes complications and diabetes-related mortality [27]. However, Hemmingsen et al [28] did not find firm evidence that diet alone or physical activity alone influences the risk of type 2 diabetes mellitus or its associated complications in people at increased risk of developing type 2 diabetes mellitus compared to standard treatment [28]. The trials included in this study had little data on the impact of games on diet, and only 3 articles evaluated participants' postintervention diet. From the results, education through games was effective, although the improvements in glycemic control were not statistically significant. The most important reason was that the 3 trials studied patients with type 1 diabetes mellitus aged 8 to 18 years. The games provided diabetes-related diet education to the patients, but family-based diet intervention may also not impact glycemic control [29].

Effects of Games Related to Physical Activity on Blood Glucose

Physical activity with different intensities impacts glycemic control in individuals with diabetes. Of the included studies, 4 trials [17,18,20,24] assessed physical activity by daily steps or self-reported activity, and this analysis found a significant increase. These results are consistent with findings from other meta-analyses showing increased physical activity among patients with chronic disease [30-32]. Some studies find positive effects with low-intensity physical activity, although these are not reflected by a decrease in HbA_{1c} or FBG in patients with type 2 diabetes [33-35]. A meta-analysis showed that high-intensity interval exercise significantly reduced HbA_{1c} levels compared to no or low-intensity exercise [36]. Low exercise intensity in the 9 studies we included may be the reason why there was no significant difference in HbA_{1c} and FBG in patients with diabetes between the groups. However, the games in the virtual reality group were relatively novel, which was very helpful for improving cognition, physical skills that are directly involved in functional abilities, and enthusiasm for sports [19].

The study by Höchsmann et al [20] contributed a substantial amount of heterogeneity; without this study, I^2 was 11%. The high heterogeneity may have been caused by the baseline of the participants in this trial being better than those in the other trials. In their trial, Höchsmann et al [20] used a dilapidated garden to symbolize the patient's physical condition, and exercise and daily physical activity execution were tracked by mobile phones, allowing for feedback. After 24 weeks of intervention, there was no significant change in HbA_{1c} levels in the intervention group, while HbA_{1c} levels in the control group receiving 1-time lifestyle counseling increased. In the trial, the intervention group had a higher increase in daily steps than the control group, providing evidence that physical activity can be encouraged by electronic games.

Effects of Games on Blood Lipids, Blood Pressure, and BMI

In our study, game-based intervention resulted in no significant decrease in blood lipids in patients with diabetes. Only 2 trials reported the outcomes of blood pressure [17,20] and BMI [17,19], and the 2 indexes were both reduced. Systolic blood pressure was below 140 mm Hg but above 130 mm Hg, which is still high for patients with diabetes. Treatment with medication may be indispensable.

Effects of Games on Weight

Lifestyle intervention can be effective for achieving clinically important reductions in body weight [37,38]. It has been demonstrated that electronic game activities are engaging, which encourages their use on a regular basis, improving the long-term outcome of a treatment for obesity [39,40]. However, an intervention using a different avatar did not improve physical activity practice or self-efficacy expectations [41]. Gomez et al [42] showed that high exercise intensity from active electronic games elicited significant increases in energy expenditure. In this study, electronic games did not result in significant weight reduction, and BMI was reduced slightly in 2 trials. Possible reasons include insufficient physical activity and that participants did not strictly control their diet. Whether electronic games are beneficial for weight control by encouraging appropriate intensity exercise in patients with diabetes requires more clinical evidence in the future.

The reasons for the lack of significant results in this meta-analysis may be as follows. First, participants in the control groups were also familiar with what the game taught. Second, patients with type 2 diabetes mellitus in the intervention groups, who were all older than 40 years, could not make full use of electronic devices and adapt to the games. Third, for exercise-based interventions, not all studies involved regular exercise monitoring for participants and established appropriate feedback or interaction mechanisms.

Limitations and Future Directions

This study had several limitations. First, not every included study reported the HbA_{1c} and FBG levels. Some excluded studies had relevant interventions but did not observe blood glucose changes or failed to give detailed trial data results. Second, the studies that were included in this meta-analysis were not homogeneous. Different games or game mechanisms were used in different patient populations. The number of participants was not large in several of the included studies, and each study used different games. Therefore, it is difficult to conduct detailed hierarchical verification of the effects of different games on blood glucose. We strived to ensure that the included studies were high-quality RCTs with strict inclusion and exclusion criteria, excluding nongaming electronic interventions. Existing studies have evaluated the effectiveness of electronic games as an alternative for traditional diabetes education. As diabetes continues, it is necessary to promote this management model. However, future studies should not only design the game in terms of increased knowledge and improved self-management but should encourage enhanced physical activity intensity.

Conclusion

As an alternative treatment tool in diabetes management, the studies on electronic games explored in this study showed a clinical improvement in glycemic control and weight control, although this improvement was not superior to that observed in the control participants. Thus, such interventions may complement existing treatment courses for diet, self-management education, and high-intensity physical activity to potentially increase the compliance of patients with diabetes. More new technologies can be used for diabetes control, and electronic games can be designed for different groups of patients with diabetes. For example, immersive virtual reality is an emerging strategy to enhance exercise performance for young patients with diabetes, and the metaverse may be a new community enabling older patients to form new social connections and share their experiences of living with diabetes. Interactive exercise games can be used in children to increase interest in education and family companionship time, and thus improve exercise compliance.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (grants 82170796 and 82003370), the Natural Science Foundation of Guangdong province (grant 2019A1515012225), the Science and Technology Plan of Guangdong province (grant 2022A0505030005), and Science and Technology Program of Guangzhou (grants 202102020165 and 2024A04J9990).

Authors' Contributions

Y Cheng conceived the idea of the overview and wrote the protocol and full overview. WY and YH developed the concept and details of the overview (ie, participants, intervention, comparison, outcomes). WY and Y Chen carried out searches and selected reviews for inclusion (YH acted as an arbitrator). WY and SY carried out the assessment of methodological quality (YH acted as an arbitrator). WY and SY extracted the data and interpreted the initial findings. YH and Y Chen directed data analyses. WY, YH, LY, and Y Cheng formulated the focus of the discussion and made suggestions for future studies. All authors were involved in the interpretation of the results and in approving the final review.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.
[DOCX File , 28 KB - games_v12i1e43574_app1.docx]

Multimedia Appendix 2
Search strategy.
[PDF File (Adobe PDF File), 116 KB - games_v12i1e43574_app2.pdf]

Multimedia Appendix 3
Basic characteristics of the included studies.
[DOCX File , 21 KB - games_v12i1e43574_app3.docx]

References

1. Diabetes around the world in 2021. IDF Diabetes Atlas. URL: https://www.diabetesatlas.org [accessed 2022-05-13]
2. Diabetes. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/diabetes# [accessed 2022-05-13]
3. Gregg EW, Sattar N, Ali MK. The changing face of diabetes complications. Lancet Diabetes Endocrinol 2016 Jun;4(6):537-547. [doi: 10.1016/S2213-8587(16)30010-9] [Medline: 27156051]
4. Molania T, Alimohammadi M, Akha O, Mousavi J, Razvini R, Salehi M. The effect of xerostomia and hyposalivation on the quality of life of patients with type II diabetes mellitus. Electron Physician 2017 Nov;9(11):5814-5819. [doi: 10.19082/5814] [Medline: 29403624]
5. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol 2018 Feb;14(2):88-98. [doi: 10.1038/nrendo.2017.151] [Medline: 29219149]
6. Pan X, Li G, Hu Y, Wang J, Yang W, An Z, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. the Da Qing IGT and Diabetes Study. Diabetes Care 1997 Apr;20(4):537-544. [doi: 10.2337/diacare.20.4.537] [Medline: 9096977]
7. American Diabetes Association. 5. Lifestyle management: standards of medical care in diabetes-2019. Diabetes Care 2019 Jan;42(Suppl 1):S46-S60. [doi: 10.2337/dc19-S005] [Medline: 30559231]
8. Liu M, Liu C, Zhang Z, Zhou C, Li Q, He P, et al. Quantity and variety of food groups consumption and the risk of diabetes in adults: a prospective cohort study. Clin Nutr 2021 Dec;40(12):5710-5717. [doi: 10.1016/j.clnu.2021.10.003] [Medline: 34743048]
9. Weber MB, Hassan S, Quarells R, Shah M. Prevention of type 2 diabetes. Endocrinol Metab Clin North Am 2021 Sep;50(3):387-400 [FREE Full text] [doi: 10.1016/j.ecl.2021.05.003] [Medline: 34399952]
10. Cahn A, Akirov A, Raz I. Digital health technology and diabetes management. J Diabetes 2018 Jan;10(1):10-17. [doi: 10.1111/1753-0407.12606] [Medline: 28872765]

11. Kerfoot BP, Gagnon DR, McMahon GT, Orlander JD, Kurgansky KE, Conlin PR. A team-based online game improves blood glucose control in veterans with type 2 diabetes: a randomized controlled trial. Diabetes Care 2017 Sep;40(9):1218-1225. [doi: 10.2337/dc17-0310] [Medline: 28790131]
12. Ruppert B. New directions in virtual environments and gaming to address obesity and diabetes: industry perspective. J Diabetes Sci Technol 2011 Mar 01;5(2):277-282 [FREE Full text] [doi: 10.1177/193229681100500212] [Medline: 21527094]
13. Jo A, Coronel BD, Coakes CE, Mainous AG. Is there a benefit to patients using wearable devices such as Fitbit or health apps on mobiles? a systematic review. Am J Med 2019 Dec;132(12):1394-1400.e1. [doi: 10.1016/j.amjmed.2019.06.018] [Medline: 31302077]
14. Martos-Cabrera MB, Membrive-Jiménez MJ, Suleiman-Martos N, Mota-Romero E, Cañadas-de la Fuente GA, Gómez-Urquiza JL, et al. Games and health education for diabetes control: a systematic review with meta-analysis. Healthcare (Basel) 2020 Oct 14;8(4):399. [doi: 10.3390/healthcare8040399] [Medline: 33066372]
15. Christensen J, Valentiner LS, Petersen RJ, Langberg H. The effect of game-based interventions in rehabilitation of diabetics: a systematic review and meta-analysis. Telemed e-Health 2016 Oct;22(10):789-797. [doi: 10.1089/tmj.2015.0165]
16. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
17. Kempf K, Martin S. Autonomous exercise game use improves metabolic control and quality of life in type 2 diabetes patients - a randomized controlled trial. BMC Endocr Disord 2013 Dec 10;13:57 [FREE Full text] [doi: 10.1186/1472-6823-13-57] [Medline: 24321337]
18. Patel MS, Small DS, Harrison JD, Hilbert V, Fortunato MP, Oon AL, et al. Effect of behaviorally designed gamification with social incentives on lifestyle modification among adults with uncontrolled diabetes: a randomized clinical trial. JAMA Netw Open 2021 May 03;4(5):e2110255 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.10255] [Medline: 34028550]
19. Cai H, Li G, Jiang S, Yin H, Liu P, Chen L. Effect of low-intensity, Kinect-based Kaimai-style qigong exercise in older adults with type 2 diabetes. J Gerontol Nurs 2019 Feb 01;45(2):42-52. [doi: 10.3928/00989134-20190111-05] [Medline: 30690653]
20. Höchsmann C, Müller O, Ambühl M, Klenk C, Königstein K, Infanger D, et al. Novel smartphone game improves physical activity behavior in type 2 diabetes. Am J Prev Med 2019 Jul;57(1):41-50 [FREE Full text] [doi: 10.1016/j.amepre.2019.02.017] [Medline: 31128953]
21. Brown SJ, Lieberman DA, Germeny BA, Fan YC, Wilson DM, Pasta DJ. Educational video game for juvenile diabetes: results of a controlled trial. Med Inform (Lond) 1997;22(1):77-89. [doi: 10.3109/14639239709089835] [Medline: 9183781]
22. Koohmareh Z, Karandish M, Hadianfard AM. Effect of implementing a mobile game on improving dietary information in diabetic patients. Med J Islam Repub Iran 2021;35:68 [FREE Full text] [doi: 10.47176/mjiri.35.68] [Medline: 34277505]
23. Kumar VS, Wentzell KJ, Mikkelsen T, Pentland A, Laffel LM. The DAILY (Daily Automated Intensive Log for Youth) trial: a wireless, portable system to improve adherence and glycemic control in youth with diabetes. Diabetes Technol Ther 2004 Aug;6(4):445-453. [doi: 10.1089/1520915041705893] [Medline: 15320998]
24. Rafeezadeh E, Ghaemi N, Miri HH, Rezaeian A. Effect of an educational video game for diabetes self-management on adherence to a self-care regimen in children with type 1 diabetes. Evid Based Care J 2019;9:74-83. [doi: 10.22038/ebcj.2019.40293.2058]
25. Batch BC, Spratt SE, Blalock DV, Benditz C, Weiss A, Dolor RJ, et al. General behavioral engagement and changes in clinical and cognitive outcomes of patients with type 2 diabetes using the Time2Focus mobile app for diabetes education: pilot evaluation. J Med Internet Res 2021 Jan 20;23(1):e17537 [FREE Full text] [doi: 10.2196/17537] [Medline: 33470947]
26. Kolb H, Martin S. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. BMC Med 2017 Jul 19;15(1):131 [FREE Full text] [doi: 10.1186/s12916-017-0901-x] [Medline: 28720102]
27. Gong Q, Zhang P, Wang J, Ma J, An Y, Chen Y, et al. Morbidity and mortality after lifestyle intervention for people with impaired glucose tolerance: 30-year results of the Da Qing Diabetes Prevention Outcome Study. The Lancet Diabetes & Endocrinology 2019;7:452-461. [doi: 10.1016/s2213-8587(19)30093-2]
28. Hemmingsen B, Gimenez-Perez G, Mauricio D, Roqué IFM, Metzendorf MI, Richter B. Diet, physical activity or both for prevention or delay of type 2 diabetes mellitus and its associated complications in people at increased risk of developing type 2 diabetes mellitus. Cochrane Database Syst Rev 2017;12:CD003054. [doi: 10.1002/14651858.cd003054.pub4]
29. Nansel TR, Laffel LMB, Haynie DL, Mehta SN, Lipsky LM, Volkening LK, et al. Improving dietary quality in youth with type 1 diabetes: randomized clinical trial of a family-based behavioral intervention. Int J Behav Nutr Phys Act 2015 May 08;12(1):58 [FREE Full text] [doi: 10.1186/s12966-015-0214-4] [Medline: 25952160]
30. Hodkinson A, Kontopantelis E, Adeniji C, van Marwijk H, McMillian B, Bower P, et al. Interventions using wearable physical activity trackers among adults with cardiometabolic conditions: a systematic review and meta-analysis. JAMA Netw Open 2021 Jul 01;4(7):e2116382 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.16382] [Medline: 34283229]
31. Qiu S, Cai X, Chen X, Yang B, Sun Z. Step counter use in type 2 diabetes: a meta-analysis of randomized controlled trials. BMC Med 2014 Feb 27;12(1):36 [FREE Full text] [doi: 10.1186/1741-7015-12-36] [Medline: 24571580]
32. Martos-Cabrera MB, Velando-Soriano A, Pradas-Hernández L, Suleiman-Martos N, Cañadas-de la Fuente GA, Albendín-García L, et al. Smartphones and apps to control glycosylated hemoglobin (HbA1c) level in diabetes: a systematic

review and meta-analysis. J Clin Med 2020 Mar 04;9(3):693 [FREE Full text] [doi: 10.3390/jcm9030693] [Medline: 32143452]
33. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013 Sep 08;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
34. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. JAMA 2018 Nov 20;320(19):2020-2028 [FREE Full text] [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
35. Lee I, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of step volume and intensity with all-cause mortality in older women. JAMA Intern Med 2019 Aug 01;179(8):1105-1112 [FREE Full text] [doi: 10.1001/jamainternmed.2019.0899] [Medline: 31141585]
36. Ruffino JS, Songsorn P, Haggett M, Edmonds D, Robinson AM, Thompson D, et al. A comparison of the health benefits of reduced-exertion high-intensity interval training (REHIT) and moderate-intensity walking in type 2 diabetes patients. Appl Physiol Nutr Metab 2017 Feb;42(2):202-208 [FREE Full text] [doi: 10.1139/apnm-2016-0497] [Medline: 28121184]
37. Franz MJ, Boucher JL, Rutten-Ramos S, VanWormer JJ. Lifestyle weight-loss intervention outcomes in overweight and obese adults with type 2 diabetes: a systematic review and meta-analysis of randomized clinical trials. J Acad Nutr Diet 2015 Sep;115(9):1447-1463. [doi: 10.1016/j.jand.2015.02.031] [Medline: 25935570]
38. Salinardi TC, Batra P, Roberts SB, Urban LE, Robinson LM, Pittas AG, et al. Lifestyle intervention reduces body weight and improves cardiometabolic risk factors in worksites. Am J Clin Nutr 2013 Apr;97(4):667-676 [FREE Full text] [doi: 10.3945/ajcn.112.046995] [Medline: 23426035]
39. Skip Rizzo A, Lange B, Suma EA, Bolas M. Virtual reality and interactive digital game technology: new tools to address obesity and diabetes. J Diabetes Sci Technol 2011 Mar 01;5(2):256-264 [FREE Full text] [doi: 10.1177/193229681100500209] [Medline: 21527091]
40. Cesa GL, Manzoni GM, Bacchetta M, Castelnuovo G, Conti S, Gaggioli A, et al. Virtual reality for enhancing the cognitive behavioral treatment of obesity with binge eating disorder: randomized controlled study with one-year follow-up. J Med Internet Res 2013 Jun 12;15(6):e113 [FREE Full text] [doi: 10.2196/jmir.2441] [Medline: 23759286]
41. Navarro J, Cebolla A, Llorens R, Borrego A, Baños RM. Manipulating self-avatar body dimensions in virtual worlds to complement an internet-delivered intervention to increase physical activity in overweight women. Int J Environ Res Public Health 2020 Jun 05;17(11):4045 [FREE Full text] [doi: 10.3390/ijerph17114045] [Medline: 32517117]
42. Gomez DH, Bagley JR, Bolter N, Kern M, Lee CM. Metabolic cost and exercise intensity during active virtual reality gaming. Games Health J 2018 Oct;7(5):310-316. [doi: 10.1089/g4h.2018.0012] [Medline: 30325233]

Abbreviations

FBG: fasting blood glucose
HbA_{1c}: glycated hemoglobin
MD: mean difference
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMD: standardized mean difference

Edited by T Leung; submitted 16.10.22; peer-reviewed by A Teles, J Chen, J Egger, T Baranowski; comments to author 09.03.23; revised version received 05.05.23; accepted 24.10.23; published 12.02.24.

Please cite as:
Yao W, Han Y, Yang L, Chen Y, Yan S, Cheng Y
Electronic Interactive Games for Glycemic Control in Individuals With Diabetes: Systematic Review and Meta-Analysis
JMIR Serious Games 2024;12:e43574
URL: https://games.jmir.org/2024/1/e43574
doi:10.2196/43574
PMID:38345856

Review

Virtual Reality Therapy for the Management of Chronic Spinal Pain: Systematic Review and Meta-Analysis

Tongtong Zhang[1,2*], BS; Xin Li[1,2*], MS; Xuan Zhou[1*], MM; Lixia Zhan[3*], MM; Fan Wu[1*], BM; Zefan Huang[1], BM; Yuxun Sun[4]; Yufei Feng[2]; Qing Du[1,5], PhD

[1]Department of Rehabilitation, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
[2]School of Exercise and Health, Shanghai University of Sport, Shanghai, China
[3]The Second People's Hospital of Beihai, Beihai, China
[4]College of Rehabilitation Sciences, Shanghai University of Medicine & Health Sciences, Shanghai, China
[5]Chongming Hospital, Shanghai University of Medicine & Health Sciences, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Qing Du, PhD
Department of Rehabilitation
Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine
1665 Kongjiang Road, Yangpu District, Shanghai, China
Shanghai, 200092
China
Phone: 86 021 25078600
Email: duqing@xinhuamed.com.cn

Abstract

Background: The effectiveness of virtual reality (VR) therapy in adults with chronic spinal pain (CSP) is unclear.

Objective: This study was conducted to compare the effectiveness of VR therapy and other therapies in adults with CSP, especially patients with inflammation-related pain.

Methods: PubMed, Web of Science, Cochrane Library, Embase, and CINAHL databases were searched up to November 11, 2023. Randomized controlled trials (RCTs) comparing adults with CSP receiving VR therapy with those receiving other therapies were included. The trial registration platform as well as the reference lists of included studies and previous systematic reviews and meta-analyses were manually searched. Two independent reviewers performed study selection, data extraction, risk-of-bias assessment, and evaluation of the quality of the evidence. The weighted mean difference (WMD) was used as the effect size used to synthesize the outcome measure.

Results: In total, 16 RCTs involving 800 participants were included in this meta-analysis. The pooled data from 15 (94%) RCTs including 776 (97%) participants showed that VR therapy was superior in improving pain intensity (WMD=–1.63, 95% CI –2.11 to –1.16, P<.001, I^2=90%) and reducing inflammatory markers, including C-reactive protein (WMD=–0.89, 95% CI –1.07 to –0.70, P<.001, I^2=0%), tumor necrosis factor-alpha (WMD=–6.60, 95% CI –8.56 to –4.64, P<.001, I^2=98%), and interleukin-6 (WMD=–2.76, 95% CI –2.98 to –2.53, P<.001, I^2=0%). However, no significant differences were found in terms of the spinal range of motion (ROM), disability level, or fear of movement. In addition, 10 (63%) of the included RCTs had a high risk of bias.

Conclusions: VR therapy may be an effective and safe intervention for reducing symptoms in patients with CSP, as it is shown to exert significant analgesic effects and beneficial improvements in inflammatory factor levels. However, this approach may not have significant effects on the spinal ROM, disability level, or fear of movement. Notably, the quality of the evidence from the RCTs included in this study ranged from moderate to low. Therefore, we recommend that readers interpret the results of this study with caution.

Trial Registration: PROSPERO CRD42022382331; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=382331

(JMIR Serious Games 2024;12:e50089) doi:10.2196/50089

KEYWORDS
virtual reality; chronic spinal pain; inflammation-related pain; systematic review; meta-analysis

Introduction

Chronic spinal pain (CSP), which most commonly includes chronic low back pain (CLBP) and chronic neck pain (CNP), is the leading cause of years with disability worldwide [1,2] and constitutes the most frequent reason for patients to seek medical care in any given year. The lifetime prevalence of low back pain (LBP) is 84%; more specifically, the lifetime prevalence of CLBP is 23%, and LBP accounts for approximately 11%-12% of cases of disability [3]. CSP is recognized as a biopsychosocial syndrome [4]. Prolonged pain can lead to anxiety, depression, and other negative emotions and is particularly significant in patients with CSP, as it is associated with decreased quality of sleep and reduced physical activity, thus placing tremendous strain on health care systems and world economies [5].

Previous studies have reported that an intervertebral disc undergoes aging or pathological changes in the adjacent region in patients with CSP, exposing cells within the nucleus pulposus to macrophages, resulting in an inflammatory response that might trigger pain [6,7]. The guidelines recommend that nonsteroidal anti-inflammatory drugs (NSAIDs) be the primary choice for patients with chronic pain [8]. However, compared with a placebo, NSAIDs can reduce CSP by controlling the level of inflammation but do not achieve clinically important efficacy [9]. Additionally, long-term use may be associated with adverse effects (eg, gastrointestinal reactions, hepatic and renal damage, and cardiovascular risk) [10]. Several studies have shown that conventional nonpharmacological therapies, such as spinal manipulation, acupuncture, exercise therapy, yoga, and cognitive-behavioral therapy, are beneficial for reducing CSP and improving psychological symptoms but have limited effects (small to moderate) [11-14]. Effective cognitive-behavioral therapies are not widely accessible due to the reliance on therapist experience, and the long-term effectiveness of these therapies remains unclear [15]. Notably, the majority of patients with CSP have goals of pain management (using ongoing care) rather than "curing" (care with a specific end) for their therapeutic care because of the complexity of the causes of chronic pain [16]. Thus, pain management is as important as the control of inflammation levels for patients with CSP. There is an urgent need for an alternative analgesic nonpharmacological and anti-inflammatory strategy for patients with CSP.

Virtual reality (VR) is typically characterized by low cost, easy availability, reusability, and personalized customization; VR therapy has been used as an alternative approach for pain management in various populations, such as individuals with spinal cord injuries, burns, and phantom limb pain [17-19]. VR can be categorized into 2 types: nonimmersive virtual reality (NIVR) and immersive virtual reality (IVR). NIVR is managed using a computer or console gaming system and a 2D interface device (mouse, keyboard, or gamepad, joystick), and patients do not need to be fully immersed in a virtual environment for experience [20].With the use of professional equipment, hardware, and configuration of the corresponding software, IVR can mimic reality by enabling the user to interact with the virtual environment [21]. A recent study demonstrated that regular exercise with the use of VR might be related to a decrease in inflammation in participants undergoing chronic hemodialysis [22], and inflammatory arthritis–targeting innovative teaching approaches based on VR technology are considered feasible [23]. There is limited evidence regarding the beneficial effects of VR therapy on pain in patients with CNP [24] and CLBP [25,26]; furthermore, there is insufficient focus on inflammatory factors. Therefore, this study aimed to investigate the potential efficacy of VR in reducing pain intensity and the levels of inflammatory factors in patients with CSP, thereby providing an updated summary of the existing evidence.

Methods

Study Protocol and Registration

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. The PRISMA checklist is given in Multimedia Appendix 1. The study protocol was registered in the PROSPERO database (CRD42022382331). The *Cochrane Handbook for Systematic Reviews of Interventions* (version 5.1.0) was followed [27].

Search Strategy

Search Sources

PubMed, Web of Science, Cochrane Library, Embase, and CINAHL electronic databases were searched from inception to November 11, 2023, to identify relevant studies. The reference lists of the included studies, as well as systematic reviews and meta-analyses that examined the efficacy of VR in patients with CSP, were manually searched for additional eligible studies. The trial registration platform ClinicalTrials was also searched for ongoing studies that reported sufficient data on the efficacy of VR for CSP.

Search Terms

The studies on VR for CSP were identified by formulating appropriate search terms. These terms were selected based on the target population (spinal pain, neck pain, thoracic pain, back pain, LBP, sacral pain, and intervertebral disc pain), target intervention (eg, VR), and target study design (eg, randomized controlled trial [RCT]). The detailed search strategy is shown in Multimedia Appendix 2.

Study Eligibility Criteria

The inclusion criteria were as follows:

- Participants: adults older than 18 years with chronic pain (more than 12 weeks) in the spinal region were included, except those who were receiving analgesic medication and who had cancer-related pain or neuropathic pain (eg, neuropathic pain after spinal cord injury, herniated disc with compression, sciatica, or lumbosacral radiculitis).

- Intervention: VR therapy.
- Comparisons: sham stimulation, usual care, and conventional treatment.
- Outcomes: pain intensity, inflammatory markers (eg, C-reactive protein [CRP], tumor necrosis factor-alpha [TNF-α], and interleukin [IL]-2, IL-4, and IL-6), fear of movement, spinal range of motion (ROM), and disability level.
- Study design: RCT.

No restrictions were imposed on language or publication date.

Study Selection

The retrieved studies were imported into Endnote X9 software (Clarivate), which was used to eliminate duplicate studies. Two independent reviewers (authors TTZ and FW) performed the initial screening of the literature by reading the titles and abstracts of all retrieved studies, and studies that did not meet the inclusion criteria were excluded. Next, the full texts of the remaining studies were screened. Any disagreements were resolved by negotiation and discussion with a third reviewer (author XZ).

Data Extraction

Two independent reviewers (authors FW and XL) extracted detailed information, including the name of the first author, the year and country of publication, the language of publication, study design, the number of included subjects (% female), diagnosis, and outcome indicators. Information about the characteristics of the interventions, including dose, frequency, and duration, was also collected for both the VR group and the control group. The sample size and mean (SD) of the outcome indicators in each group were collected. When the same group of participants was reported in different studies, the group with the largest sample size was selected for inclusion in this review to avoid duplicate publications [28]. For information that could not be confirmed, the authors were contacted by email. The 2 reviewers cross-checked the data at the end of the extraction, and any disagreements were resolved by negotiation.

Risk-of-Bias Assessment

The methodological quality of the included studies was independently assessed by 2 reviewers (authors XL and ZFH) using the Cochrane Risk of Bias tool, and the studies were classified as having a low, unclear, or high risk of bias [29]. Disagreements were resolved by consulting a third reviewer (author QD). The Egger test and funnel plots generated with Stata 14.0 software (StataCorp) were used to evaluate potential publication bias. The trim-and-fill method was used to adjust for funnel plot asymmetry due to publication bias [30]. Sensitivity analyses were performed by removing each study separately to assess the robustness of the results [29]. The overall strength of the evidence was assessed using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) criteria [31].

Meta-Analysis and Subgroup Analysis

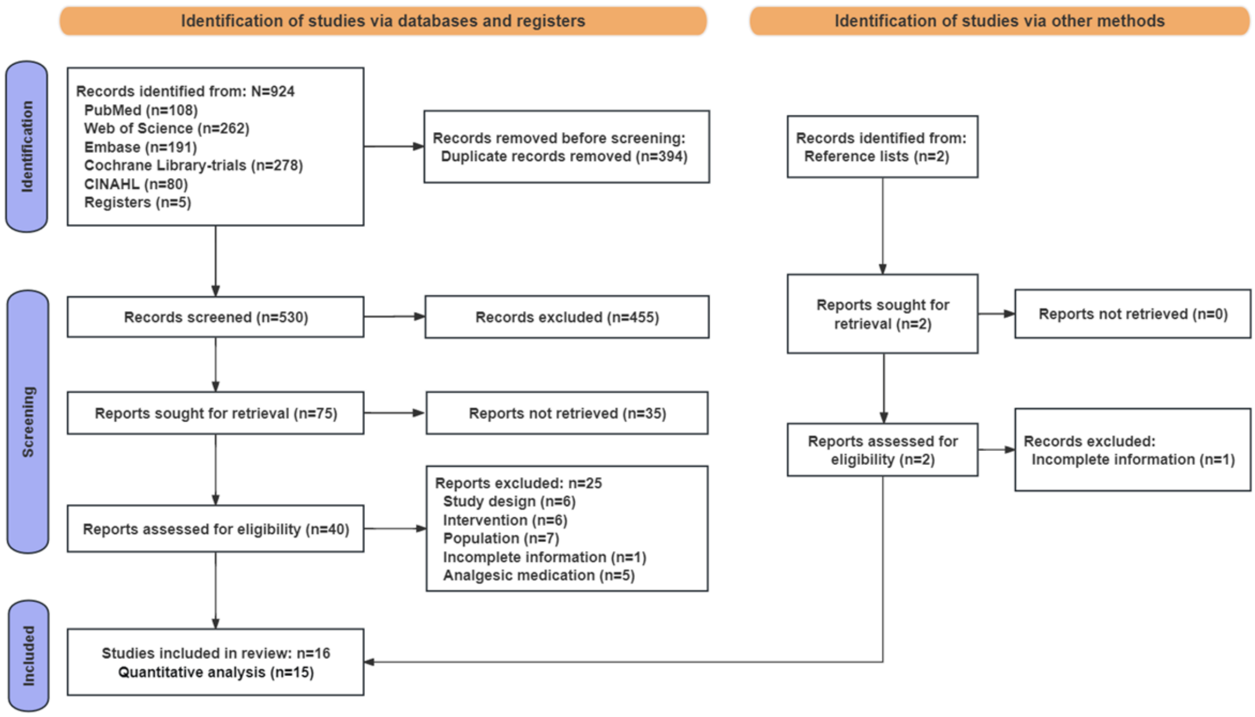
This systematic review and meta-analysis were performed using Review Manager software 5.4 (Informer Technologies) and Stata 14.0 software. Heterogeneity was tested using the I^2 statistic. A fixed effects model was selected for the outcome indicators if I^2<50%, while a random effects model was used when there was significant statistical heterogeneity (I^2>50%, P<.05). The effect size used to synthesize the outcome measure was the weighted mean difference (WMD). Three subgroup analyses were performed to explore the possible causes of heterogeneity among the studies: the region of CSP (CNP vs CLBP), VR types (IVR vs NIVR), and treatment duration (<4 weeks vs ≥4 weeks).

Results

Search Results

A total of 924 records were obtained from the 5 databases and the trial registration platform. A total of 394 (42.6%) duplicates were identified and removed using Endnote X9 software. After screening the titles and abstracts, 40 (7.5%) of the remaining 530 RCTs were retained, and 490 (92.5%) were excluded for the following reasons: (1) the study population included patients without CSP, (2) the intervention did not use VR therapy, (3) the type of study was a non-RCT, (4) the information was incomplete, and (5) the patients also received analgesic medication. Of the 40 studies, 15 (38%) were retained after reading the full text and 25 (62%) were excluded for the following reasons: (1) the study population included patients without CSP, (2) the intervention did not use VR therapy, (3) the type of study was a non-RCT, (4) the information was incomplete, and (5) the patients also received analgesic medication. Two additional RCTs were retrieved from the reference lists of the included studies. One RCT was retained after the full text was read, and the other was excluded due to incomplete information. A total of 16 studies were included in this review, 15 (94%) of which reported sufficient data (eg, mean [SD], sample size) on the analgesic effect of VR for CSP. Therefore, 15 studies were included in the meta-analysis. The PRISMA flowchart of selecting the included studies is shown in Figure 1.

Figure 1. PRISMA flowchart: database and clinical trial register search and other sources. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis.



The CSP reported in the included studies included CLBP [32-43] and CNP [44-47]. All patients had chronic pain that persisted for more than 3 months. The sample size varied from 8 to 90 participants, and the mean age ranged from 18 to 85 years. The characteristics of all the studies are summarized in Table 1.

Table 1. Characteristics of the included studies [32-47].

First author	Patient characteristics			Outcome measures	Time points	Dropout rate (%)	Country, language
	Participants, n (% female)	Age (years), mean (SD)	Diagnosis				
Garcia et al [32]	T[a]: 179 I[b]: 89 (75) C[c]: 90 (78)	I: 51.5 (13.5) C: 51.4 (12.9)	CLBP[d]	DVPRS[e], Pain Catastrophizing Scale (PCS), 8-item Chronic Pain Acceptance Questionnaire (CPAQ-8)	Baseline, –7, 0, 4, 7, 11, 14, 18, 21, 25, 28, 32, 35, 39, 42, 46, 49, 53, 56 days	I: 0 C: 0	United States, English
Nambi et al [33]	T: 60 I (VR[f]): 20 I (core stabilization [CS]): 20 C: 20	I (VR): 21.45 (1.50) I (CS): 21.39 (1.40) C: 20.97 (1.50)	CLBP	NPRS[g], quality of life (physical fitness index)	Baseline, 4 weeks, 8 weeks, 6 months	I (VR): 0.05 I (CS): 0.05 C: 0	Saudi Arabia, English
Nambi et al [34]	T: 45 I (VR): 15 I (isokinetic training [IKT]): 15 C: 15	I (VR): 20.23 (1.60) I (IKT): 21.25 (1.20) C: 20.78 (1.60)	CLBP	NPRS	Baseline, 4 weeks	I (VR): 0 I (IKT): 0 C: 0	Saudi Arabia, English
Yalfani et al [35]	T: 25 I: 13 C: 12	I: 68.00 (2.94) C: 67.08 (2.90)	CLBP	VAS[h], 36-item Short Form Health Survey (SF-36)	Baseline, 8 weeks	I: 0 C: 0	Iran, English
Park et al [36]	T: 24 I (NWE[i]): 8 I (lumbar stabilization exercise [LSE]): 8 C: 8	I (NWE): 44.12 (5.48) I (LSE): 43.37 (5.42) C: 45.50 (5.34)	CLBP	VAS	Baseline, 8 weeks	I: 0 C: 0	South Korea, English
Afzal et al [37]	T: 90 I: 45 (64.28) C: 45 (69.04)	I: 37.5 (12.5) C: 38.2 (11.8)	CLBP	VAS, Modified Oswestry Disability Index	Baseline, 4th, 8th 12th sessions	I: 0.07 C: 0.07	Pakistan, English
Nambi et al [38]	T: 60 I (VRE[j]): 20 I (isokinetic exercise [IKE]): 20 C: 20	I (VRE): 23.2 (1.6) I (IKE): 22.9 (1.7) C: 22.8 (1.8)	CLBP	VAS, inflammatory biomarkers	Baseline, 4 weeks	I (VRE): 5 I (IKE): 5 C: 0	Saudi Arabia, English
Nambi et al [39]	T: 36 I (VR): 12 I (combined physical rehabilitation [CPR]): 12 C: 12	I (VR): 21.3 (2.6) I (CPR): 21.8 (2.2) C: 20.9 (2.8)	CLBP	Inflammatory biomarkers	Baseline, 4 weeks	I (VR): 0 I (CPR): 0 C: 0	Saudi Arabia, English
Nambi et al [40]	T: 54 I (VR): 18 I (CPR): 18 C: 18	I (VR): 22.3 (1.6) I (CPR): 21.4 (1.8) C: 21.9 (1.8)	CLBP	VAS, TSK-17	Baseline, 4 weeks	I (VR): 0 I (CPR): 0 C: 0	Saudi Arabia, English

First author	Patient characteristics			Outcome measures	Time points	Dropout rate (%)	Country, language
	Participants, n (% female)	Age (years), mean (SD)	Diagnosis				
Matheve et al [41]	T: 84 I: 42 (64) C: 42 (64)	I: 42.1 (11.5) C: 44.2 (11.9)	CLBP	NPRS, Roland-Morris Disability Questionnaire (RMDQ), PCS	Baseline, postintervention	I: 0 C: 0	Belgium, English
Stamm et al [42]	T: 22 I: 11 (73) C: 11 (55)	I: 75.00 (5.80) C: 75.50 (4.39)	CLBP	NRS[k], Chronic Pain Grade Questionnaire (CPGQ), 12-item Short Form Health Survey (SF-12), Hannover Functional Ability Questionnaire for Measuring Back Pain–Related Disability (Ffb-H-R), TSK[l]-11	Baseline, 4 weeks	I: 0 C: 0	Germany, English
Monteiro-Junior et al [43]	T: 34 I: 17 (100) C: 17 (100)	T: 68 (4)	CLBP	NRS	Baseline, 8 weeks	I: 17.6 C: 5.8	Brazil, English
Cetin et al [44]	T: 41 I: 21 C: 20	I: 40.00 (11.88) C: 41.94 (10.76)	CNP[m]	Joint position sense error (JPSE), VAS, pressure pain threshold (PPT), SF-36	Baseline, 6 weeks	I: 19 C: 15	Turkey, English
Bahat et al [45]	T: 90 I (VR): 30 (63) I (laser): 30 (70) C: 30 (77)	I (VR): 48.00 (14.07) I (laser): 48.00 (17.41) C: 48.00 (17.76)	CNP	NDI[n], VAS, EQ-5D, TSK-17, cervical range of motion (CROM), kinematic measures	Baseline, 4 weeks	I (VR): 16.6 I (laser): 13.3 C: 16.6	Israel, English
Nusser et al [46]	T: 55 I (VR): 17 (53) I (sensorimotor group [SM]): 16 (69) C: 18 (66)	I (VR): 51.2 (8.8) I (SM): 53.1 (5.7) C: 49.8 (8.1)	CNP	NRS, active cervical range of motion (ACROM), NDI	Baseline, 3 weeks	I (VR): 0 I (SM): 11 C: 10	Germany, English
Tejera et al [47]	T: 44 I: 22 (50) C: 22 (54.5)	I: 32.72 (11.63) C: 26.68 (9.21)	CNP	VAS, conditioned pain modulation (PPT), ACROM device, NDI, PCS, 11-item Spanish version of the TSK	Baseline, 4 weeks, 1 month, 3 months	I: 0 C: 0	Spain, English

[a]T: total participants.
[b]I: intervention group.
[c]C: control group.
[d]CLBP: chronic low back pain.
[e]DVPRS: Defense and Veterans Pain Rating Scale.
[f]VR: virtual reality.
[g]NPRS: Numerical Pain Rating Scale.
[h]VAS: Visual Analogue Scale.
[i]NWE: Nintendo Wii exercise.
[j]VRE: virtual reality exercise.
[k]NRS: Numeric Rating Scale.
[l]TSK: Tampa Scale for Kinesiophobia.
[m]CNP: chronic neck pain.
[n]NDI: Neck Disability Index.

The types of VR interventions included IVR [32,35,42,44-47] and NIVR [33,34,36-41,43], which were classified based on the degree of isolation participants experienced when interacting with the virtual environment during VR therapy. NIVR uses a wall-mounted screen or a computer monitor as the vehicle for VR content, while IVR uses a headset or head-mounted display [48]. Compared to NIVR, IVR can increase the user's sense of presence by improving immersion through the addition of auditory or haptic feedback [49]. The duration of a single VR session ranged from 2 to 40 minutes, and the frequency of treatment ranged from 5 to 7 times a week; all the included studies ranged in duration from a single exercise session to 8 weeks. For the control groups, 5 (31%) studies performed conventional balance function training [33,34,38-40], 5 (31%) performed conventional physical therapy [36,37,41,46,47], 2 (13%) performed core training [43,44], and the remaining conducted treatments, including sham VR [32], conventional multimodal pain therapy [42], waiting lists [45], and standard care [35]. The intervention details are summarized in Table 2.

Table 2. Characteristics of the intervention protocols used in the included studies [32-47].

First author	Intervention group	Control group	Device	Duration
Garcia et al [32]	Ease VR[a], IVR[b], interactive, pain education, relaxation/interception, mindful escape, pain distraction games, dynamic breathing performed 56 times (2-16 minutes each time, average of 6 minutes, 1 time/day)	Sham VR, NIVR[c], not interactive, displayed 2D nature footage with neutral music, 20 videos rotated over 56 sessions, performed 56 times (2-16 minutes each time, average of 6 minutes, 1 time/day)	Pico G2 4K all-in-one head-mounted VR device	8 weeks
Nambi et al [33]	VR group: sit in the virtual platform and select firing game executed by trunk movements (flexion, extension, and lateral flexion; 30 minutes/day, 5 times/week, for 4 weeks); heat modality (20 minutes); therapeutic ultrasound (25 minutes)	Conventional balance function training, traditional active balance exercise for abdominal and back muscles (5 times/week for 4 weeks); heat modality (20 minutes); therapeutic ultrasound (25 minutes)	VR group: Pro-Kin system PK 252 N (TecnoBody)	4 weeks
Nambi et al [34]	VRT: shooting game (30 minutes, 5 days/week, for 4 weeks); home-based exercise; hot-pack therapy (20 minutes); ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2 in continuous form for 5 minutes)	Conventional balance function training: standardized conventional exercises actively involving abdominal, deep abdominal, and back muscles (30 minutes/session, 5 days/week, for 4 weeks); hydrocollator packs (20 minutes/session); continuous ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2) at the low back region (5 minutes, 5 days/week, for 4 weeks)	VRT: Pro-Kin system (TecnoBody)	4 weeks
Yalfani et al [35]	Fishing, boxing, tennis, football, bowling, beat saber, audio shield, and skiing (30 minutes, 3 times/week, for 8 weeks)	Standard care.	VR: HTC Vive virtual reality system	8 weeks
Park et al [36]	NWE[d]: using the Nintendo Wii exercise program, including the wakeboard, Frisbee dog, jet ski, and canoe games. Participants chose which Nintendo Wii sports program to perform and took a 2-minute break every 10 minutes (30 minutes/session, 3 times/week, for 8 weeks)	Conventional physical therapy: using physical agent modalities, such as a hot pack (30 minutes); interferential current therapy (15 minutes); deep heat with ultrasound (5 minutes)	VR: Nintendo	8 weeks
Afzal et al [37]	Kinetic exergames (trunk slide flexion, sitting to avoid obstacles, jumping and combined movement of arms, for 5 minutes); after 30 seconds of rest, play body ball game for 5 minutes (3 sessions/week for a total of 12 sessions); routine physical therapy	Conventional physical therapy: heat therapy for 10 minutes, hamstring stretching, back-strengthening exercises (3 sessions/week for a total of 12 sessions)	VR: nonimmersive system with a kinetic device (model V.2), incorporated with red-green-blue (RGB) cameras and time-of-flight (TOF) sensor, attached with a liquid crystal display (LCD) screen	4 weeks
Nambi et al [38]	VRE[e]: virtual training exercises performed in the upright position, a car race game chosen from the list of games, and training given to focus on the back muscles. The participant was asked to sit on the moving game chair and instructed to watch the game on the desktop monitor (30 minutes/session, 5 days/week, for 4 weeks); hydrocollator packs (20 minutes/session); continuous ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2) at the low back region (5 minutes, 5 days/week, for 4 weeks)	Conventional balance function training: standardized conventional exercises actively involving abdominal, deep abdominal, and back muscles (30 minutes/session, 5 days/week, for 4 weeks); hydrocollator packs (20 minutes/session); continuous ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2) at the low back region (5 minutes, 5 days/week, for 4 weeks)	VRE: Pro-Kin system (TecnoBody)	4 weeks

First author	Intervention group	Control group	Device	Duration
Nambi et al [39]	Virtual reality training (VRT): shooting game, sitting on a virtual platform and visualizing the game on the computer display screen (30 minutes each time, 5 times/week, for 4 weeks); heat modality (20 minutes); therapeutic ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2; 5 minutes); home-based exercise (10 repetitions, bottom-to-heel stretch, opposite arm/leg raise, back extension, bridging, knee rolling; 2 times/day for 4 weeks)	Conventional balance function training: active isotonic and isometric exercises for abdominal, deep abdominal, and back muscles (10-15 repetitions/day, 5 days/week for 4 weeks; stretching focused on each muscle group for 3 repetitions for 10 seconds per muscle group); heat modality (20 minutes); therapeutic ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2; 5 minutes); home-based exercise (10 repetitions, bottom-to-heel stretch, opposite arm/leg raise, back extension, bridging, knee rolling; 2 times/day for 4 weeks)	VR: Pro-Kin system PK 252 N (Pelvic Module balance trunk MF; TecnoBody)	4 weeks
Nambi et al [40]	VRT: shooting game, sitting on a virtual platform and visualizing the game on the computer display screen (30 minutes each time, 5 times/week, for 4 weeks); heat modality (20 minutes); therapeutic ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2; 5 minutes)	Conventional balance function training: active isotonic and isometric exercises for abdominal, deep abdominal, and back muscles (10-15 repetitions/day, 5 days/week for 4 weeks; stretching focused on each muscle group for 3 repetitions for 10 seconds per muscle group); heat modality (20 minutes); therapeutic ultrasound (frequency 1 MHz, intensity 1.5 W/cm^2; 5 minutes)	VR: Pro-Kin system PK 252 N (Pelvic Module balance trunk MF; TecnoBody)	4 weeks
Matheve et al [41]	2 different games (2 minutes each); single-session intervention, 2 × 2 minutes of pelvic tilt exercises in the sagittal plane, with 30 seconds of rest in between, through a wireless motion sensor	Conventional physical therapy: 2 different games (2 minutes each); single-session intervention, 2 × 2 minutes of pelvic tilt exercises in the sagittal plane, with 30 seconds of rest in between	VR: wireless motion sensor (Valedo Pro, Hocoma)	Single exercise session
Stamm et al [42]	Multimodal pain therapy in VR (movement therapy and psychoeducation), training session including 12 exercises, structured as follows: (1) warm-up (training of upper and lower extremities), (2) main part (strengthening of abdominal and back muscles, core stability), (3) cool-down (stretching, progressive muscle relaxation), (4) psychoeducative units (topics: physiology of pain, pain management, stress management, everyday training), 3 times/week for 30 minutes	Conventional multimodal pain therapy: chair-based group exercises and psychoeducation in a group setting), 3 times/week for 30 minutes	VR: head-mounted display headset using the ViRST VR app	4 weeks
Monteiro-Junior et al [43]	Virtual physical training (8 exercises, 30 minutes each time, with 3 weekly sessions lasting 90 minutes each), lasted 8 weeks, 3 times weekly/session	Core training: postures adopted by participants for 15-30 seconds or according to the capacity of each; 10-15 seconds between postures (ie, bridges), with each performed 3 times, lasted 8 weeks, 3 times weekly/session	VR: Wii Balance Board (WBB; Nintendo)	8 weeks
Cetin et al [44]	VR exercises: VR apps that allowed neck movements in all directions, motor control (MC) exercises (20 minutes and then VR for 20 minutes, 5 repetitions for each exercise; 40 minutes/session, 3 sessions/week, for 6 weeks, total of 18 sessions)	Core training: strengthening of deep cervical flexors (DCFs), deep cervical extensors (DCEs), and axioscapular muscles; stretching exercises; and postural correction exercises (40 minutes, 10 repetitions for each exercise, 3 sessions/week, for 6 weeks, total of 18 sessions)	VR: Oculus Go VR glasses, 2 VR apps installed: “Ocean Rift” and “Gala 360”	6 weeks

First author	Intervention group	Control group	Device	Duration
Bahat et al [45]	VR group: kinematic home training and customized software with the virtual airplane controlled by head motion (5 minutes, 4 times/day, 20 minutes/day, 4 times/week, for 4 weeks)	Waiting list	VR: customized neck VR system (hardware including Oculus Rift DK1 head-mounted display equipped with 3D motion tracking; software developed using Unity-pro, version 3.5, Unity Technologies)	4 weeks
Nusser et al [46]	VR group: neck-specific sensorimotor training (NSST)— head-repositioning test (HRT), head-to-target test (HTT), dynamic exercise including 5 different trajectories (3 minutes given between tasks), training divided into 6 20minute sessions for a total of 120 minutes); standard rehabilitation program	Conventional physical therapy: different forms of general and neck-specific exercise therapies (strengthening, mobilization, relaxation, medical training therapy, functional gymnastics, aqua therapy, physical therapy, and traditional “back school”)	VR: modified VR system (Fraunhofer Institute für Graphische Datenverarbeitung), helmet (Schutz helm uvex pheos alpine, Fürth), 3Space Fastrak System (Polhemus Inc)	3 weeks
Tejera et al [47]	VR mobile apps “Full Dive VR,” only lateral flexion movements of the neck; “VR Ocean Aquarium 3D”: flexion, extension, and rotation movements (3 series of 10 repetitions, with 30 seconds of rest between exercises)	Conventional physical therapy: flexion, extension, rotation, and tilt exercises (3 series of 10 repetitions, with 30 seconds of rest between exercises)	VR: VR Vox Play glasses with a head-mounted display clamping system (weight 330 g) with an LG Q6 smartphone attached, 2 VR mobile apps installed	4 weeks

[a]VR: virtual reality.

[b]IVR: immersive virtual reality.

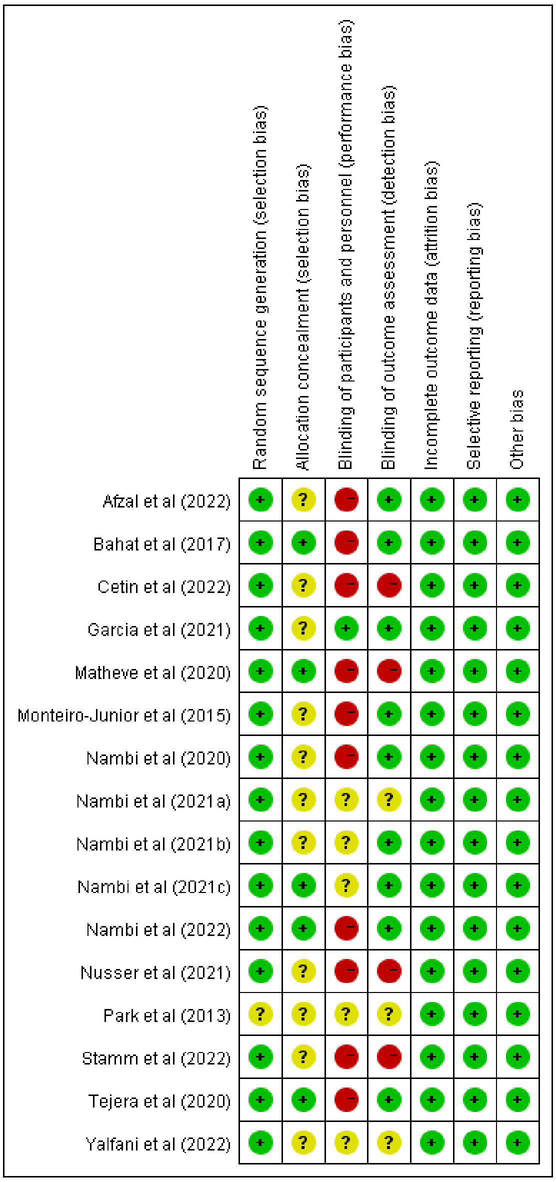
[c]NIVR: nonimmersive virtual reality.

[d]NWE: Nintendo Wii exercise.

[e]VRE: virtual reality exercise.

The risk of bias in the 16 (100%) studies included in the meta-analysis is presented in Figure 2. Overall, 10 (63%) studies showed a high risk of bias. In addition, 15 (94%) RCTs generated an adequately randomized sequence, and 9 (60%) of them were analyzed using a blinded method for outcome measurement. Ratings using the GRADE methodology for all outcome measurements were inconsistent and ranged from moderate to low quality (Multimedia Appendix 3). Therefore, the quality of evidence from most studies was classified as fair.

Figure 2. Cochrane risk-of-bias summary for included studies.



Primary Outcome

Pain Intensity

All 16 (100%) studies (800 patients) reported pain intensity: 9 (56%) used the Visual Analogue Scale (VAS) [34-38,40,44,45,47], 2 (13%) used the Numerical Pain Rating Scale (NPRS) [33,41], 3 (19%) used the Numeric Rating Scale (NRS) [42,43,46], and 1 (6%) used the Defense and Veterans Pain Rating Scale (DVPRS) [32]. The random effects model revealed that compared with the control treatment, the VR intervention significantly reduced pain intensity (WMD=–1.63, 95% CI –2.11 to –1.16, $P<.001$, I^2=90%). Clinical differences between groups were significant, and as suggested, the minimal clinically important difference (MCID) threshold on the VAS for LBP was set at a 1.5-point reduction [50]. Given the significant heterogeneity observed (I^2=90%), we performed subgroup analyses to investigate the source of heterogeneity based on the different regions, VR types, and treatment durations.

VR had a good analgesic effect on both CNP and CLBP groups compared with the control group. The results did not significantly differ among the subgroups (WMD=−1.63, 95% CI −2.11 to −1.16); see Figure 3. Moreover, a total of 7 (44%) studies demonstrated that IVR significantly improved CSP (WMD=–1.50, 95% CI –2.45 to –0.55, $P<.001$, I^2=80%) [32,35,42,44-47]. Another 8 (50%) studies showed that NIVR improved CSP substantially (WMD=–1.50, 95% CI –2.45 to –0.55, $P<.001$, I^2=90%) [33,34,36-38,40,41,43]; see Figure 4. The subgroup analyses also revealed significant differences between treatment durations of <4 weeks (WMD=–1.41, 95% CI –2.12 to –0.69, P=.001, I^2=0%) and ≥4 weeks (WMD=–1.65, 95% CI –2.16 to –1.14, $P<.001$, I^2=91%) in terms of the analgesic effect of VR treatment on CSP (Multimedia Appendix 4).

Figure 3. Forest plots of the effect of VR compared with other treatments on pain intensity in patients with CSP: subgroup analysis of posttreatment effectiveness for different regions of spinal pain. CSP: chronic spinal pain; VR: virtual reality.

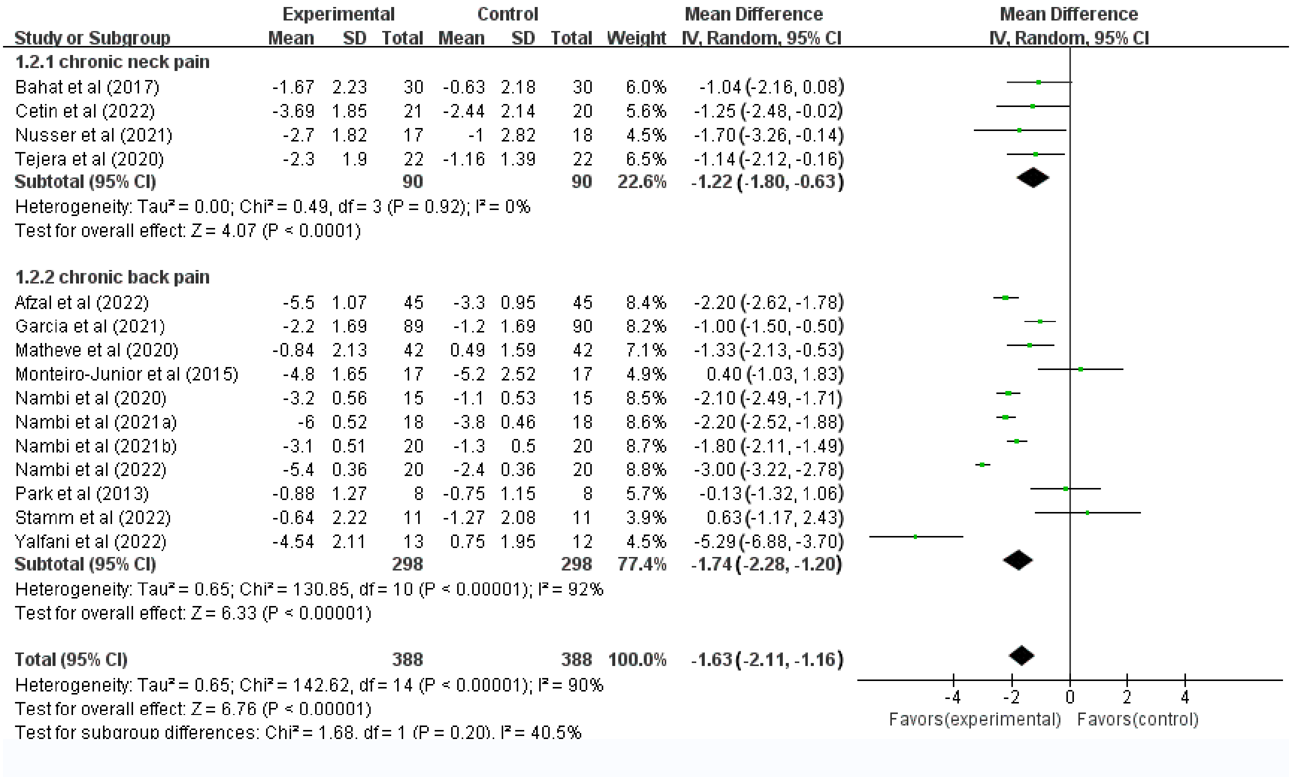
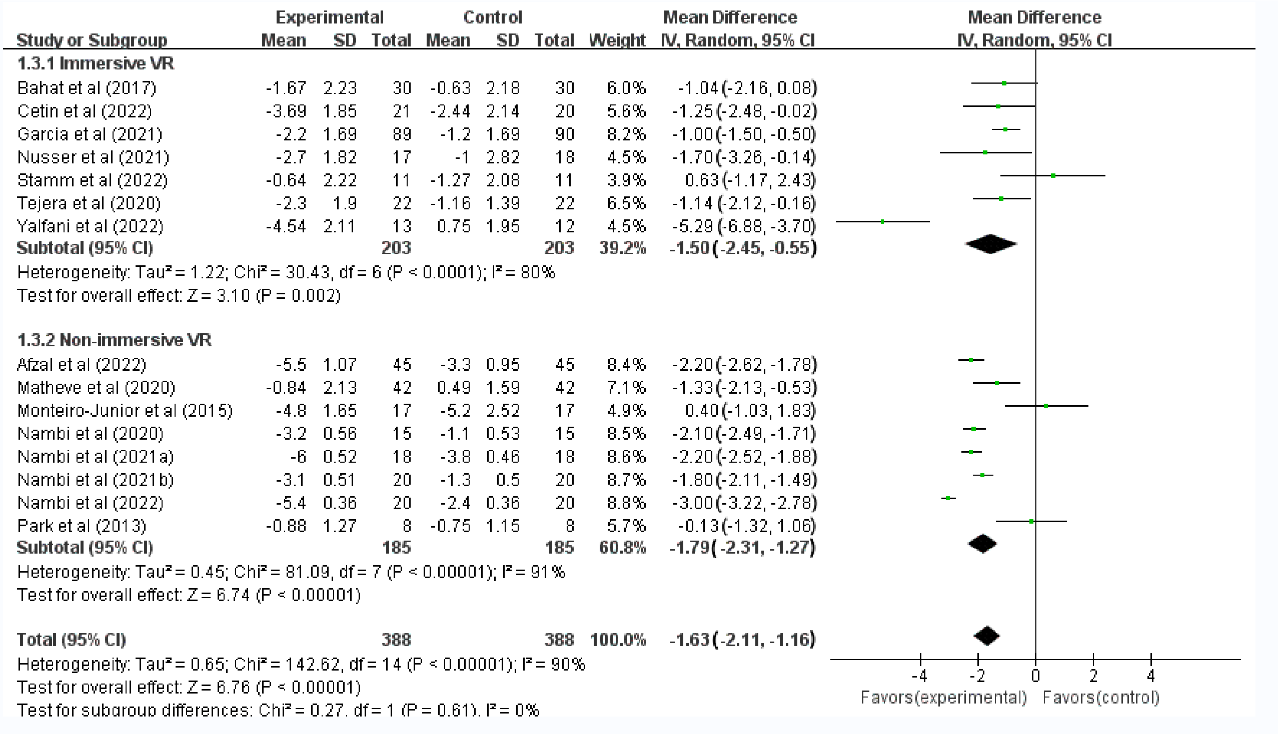


Figure 4. Forest plots of the effect of VR compared with other treatments on pain intensity in patients with CSP: subgroup analysis of posttreatment effectiveness for the VR intervention type. CSP: chronic spinal pain; VR: virtual reality.



Secondary Outcomes

Inflammatory Markers

Patients with CSP develop a systemic inflammatory response and have elevated levels of inflammatory markers in the blood [51]. Two studies (62 patients) focused on the levels of inflammatory markers (eg, CRP, TNF-α, IL-2, IL-4, and IL-6) by collecting 10 mL of venous blood [38,39]. The results showed that VR therapy significantly improved the level of CRP (WMD=–0.89, 95% CI –1.07 to –0.70, *P*<.001, I^2=0%), TNF-α (WMD=–6.60, 95% CI –8.56 to –4.64, P<.001, I^2=98%), and IL-6 (WMD=–2.76, 95% CI 2.98 to –2.53, P<.001, I^2=0%). No significant differences were found between the IL-2 and IL-4 subgroups (Figure S1 in Multimedia Appendix 5).

Fear of Movement

Four studies (162 patients) reported fear of movement according to the 11-item or 17-item Tampa Scale of Kinesiophobia (TSK-11 or TSK-17, respectively) [42,47]. No significant differences were found in either the TSK-11 (WMD=–0.81, 95% CI –4.48 to 2.86, P=.66, I^2=0%; Figure S2 in Multimedia Appendix 5) or TSK-17 (WMD=–9.66, 95% CI –22.01 to 2.68, P=.13, I^2=97%; Figure S3 in Multimedia Appendix 5).

Spinal Range of Motion

Three studies reported changes in the ROM of the neck in 4 directions before and after the intervention [45-47]. No significant differences were found between the groups in terms of flexion (WMD=2.67, 95% CI –2.31 to 7.64, P=.29, I^2=61%), extension (WMD=3.92, 95% CI –2.17 to 10.0, P=.21, I^2=48%), right rotation (WMD=–0.22, 95% CI –4.38 to 3.95, P=.92, I^2=0%), or left rotation (WMD=0.08, 95% CI –3.90 to 4.05, P=.97, I^2=42%); see Figure S4 in Multimedia Appendix 5.

Disability Level

Three studies (139 patients) reported disability levels in patients with CNP by using the Neck Disability Index (NDI) [45-47], a 10-item questionnaire that assesses self-reported disability related to CNP. Higher scores on the NDI indicate higher levels of disability. No significant differences were found in the pooled analysis of 3 (19%) studies (WMD=–2.66, 95% CI –5.47 to 0.15, P=.06, I^2=48%); see Figure S5 in Multimedia Appendix 5.

Adverse Events

One study reported that after 1 month of intervention, patients experienced nausea and motion sickness [32], two studies reported that there were no adverse events [33,37], and the remaining studies did not mention adverse events. The overall dropout rate was 4.25% (17/400) in the intervention group and 3.75% (15/400) in the control group.

Publication Bias and Sensitivity Analysis

The Egger test indicated significant publication bias in the results for pain intensity (P=.03; Figure 5). The sensitivity analysis for pain intensity revealed that removing each study separately did not significantly affect the pooled results, thus indicating that the results are robust (Figure 6). The trim-and-fill method was performed, and it was estimated that there were 4 missing studies. The pooled estimates (95% CIs) calculated for the fixed effects model and the random effects model were –2.30 (–2.42 to –2.18) and –2.06 (–2.50 to –1.61), respectively (Figure 7). No significant changes in the results were observed before or after pruning or filling, indicating that our results are robust and plausible.

Figure 5. Funnel plot of pain intensity in the VR group compared with the control group. VR: virtual reality; WMD: weighted mean difference.

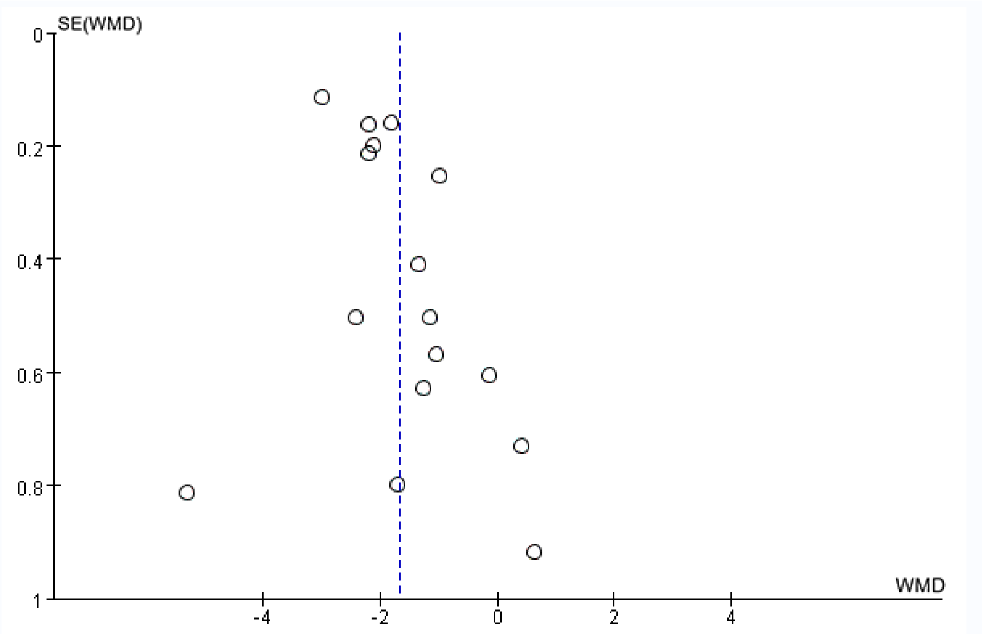


Figure 6. Sensitivity analysis of the included studies.

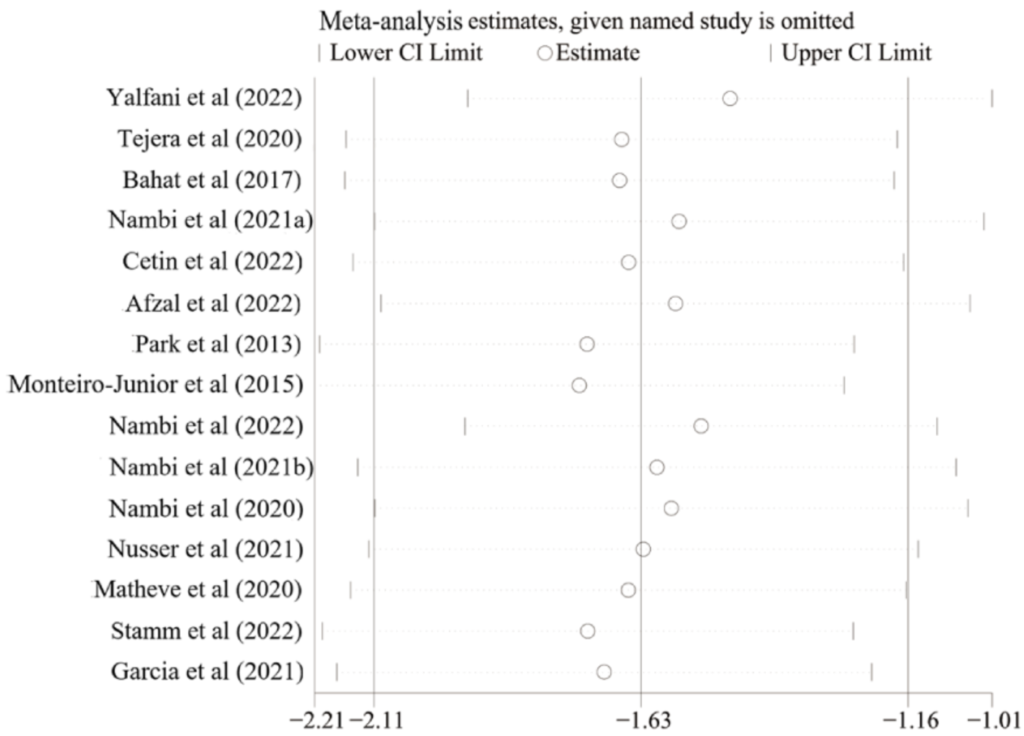
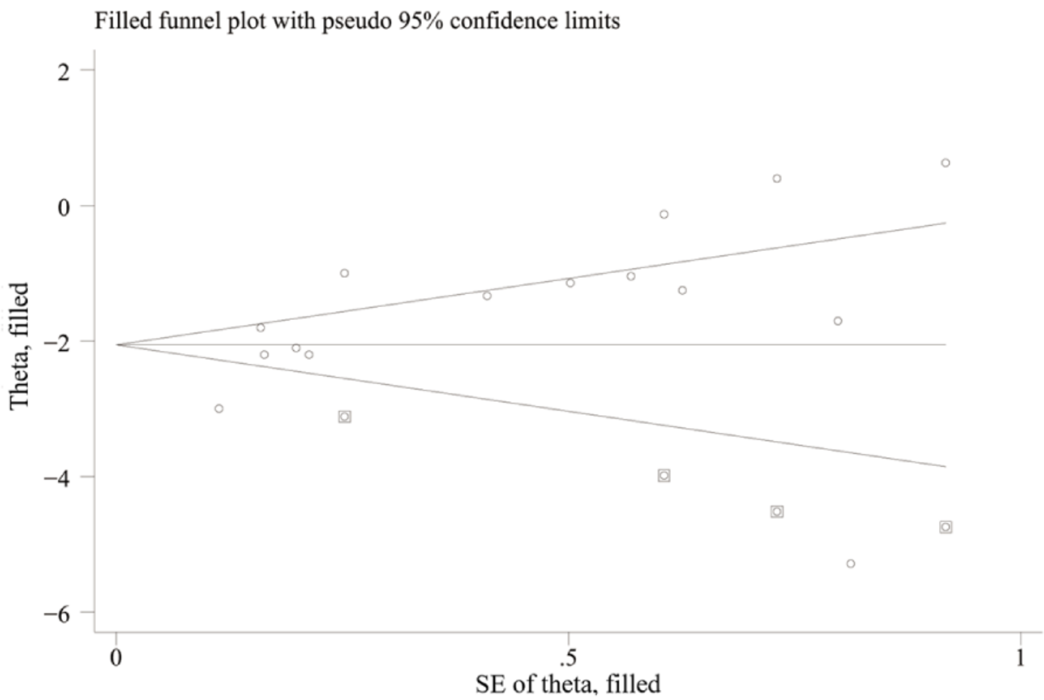


Figure 7. Trim-and-fill analysis to estimate the number of potentially missing studies about the effect of VR on CSP. Circles represent real studies, and squares represent studies estimated by this method. Circles represent real studies, and squares represent studies estimated by this method. CSP: chronic spinal pain; VR: virtual reality.



Discussion

Principal Findings

The primary purpose of this meta-analysis was to compare the relative efficacy of VR therapy and other therapies (eg, conventional therapy, sham stimulation, and standard care) for treating CSP. The results indicated that VR therapy can effectively relieve CSP. The results of subgroup analyses showed that VR is a beneficial pain management strategy for patients with CNP and CLBP. For different types of VR, subgroup analyses showed that compared to the control group, IVR and NIVR both significantly improved CSP. No statistically significant differences were found between patients who underwent VR treatments for a duration of <4 weeks and a duration of ≥4 weeks. VR was associated with a significant improvement in inflammatory marker levels but not in the fear of movement, spinal ROM, or disability level. VR was found to be well tolerated among these patients.

Discussion of the Results

The primary result suggested that VR reduces self-reported pain intensity, which might be explained by several implicating mechanisms [52,53]. A previous study reported that abdominal muscle strength is significantly lower in people with LBP [54], and a lack of strength in the core trunk muscles can lead to a decrease in intra-abdominal pressure, affecting spinal stability [55]. VR, as a novel human-computer interaction approach, can stimulate and mobilize the sensory system during training and results in changes in neuroplasticity and enhanced performance of relevant muscle groups, promoting a new motor learning process and leading to increased spinal stability [37,56], which would benefit pain relief. Furthermore, previous studies have reported that an intervertebral disc undergoes aging or pathological changes in the adjacent region in patients with CSP, exposing cells within the nucleus pulposus to macrophages, resulting in an inflammatory response that might trigger pain [7,8]. VR therapy may enhance the activity of disc fibroblasts and increase the thickness of the multifidus muscle [39,57], which is beneficial for relieving pain intensity. Furthermore, pain is an unpleasant subjective sensation associated with actual or potential tissue damage and is correlated with the degree of patient attention given to the pain area [58-61]. The various virtual game environments and real-time feedback methods are the most eye-catching features in the VR training process; these methods can be used to attract the patient's visual and auditory attention to achieve motor performance, while relatively less attention has been given to the effects of VR on pain [62,63].

Although the high heterogeneity of the primary outcome and the results of the subsequent subgroup analyses suggest that the region of CSP, VR type, and treatment duration may play a role in the heterogeneity, the results of the sensitivity analysis indicate that these differences are more likely to be caused by 6 studies [33-35,37,38,40], which included participants of different ages.

VR therapy significantly improved the levels of inflammatory markers, including CRP, TNF-α, and IL-6. Numerous studies have previously reported an association between CSP and changes in inflammatory cytokines, such as IL-1 and TNF-α, which are thought to be closely related to the pathogenesis of disc herniation and degeneration [64,65]. Similarly, Nambi et al [66] reported that 4 weeks of VR training could significantly decrease pain intensity, increase functional impairment, and improve CRP, TNF-α, IL-2, IL-4, and IL-6 levels. However, the limited number and low quality of the included studies need to be noted, and further RCTs with large samples and rigorous study designs are needed to elucidate these results.

Patients with CSP may engage in fear/avoidance behaviors to avoid pain and protect themselves by limiting spinal motion, which ultimately affects spinal mobility and the speed of movement [67,68], with the degree of pain catastrophizing being proportional to the degree of disability [69,70]. However, we found no statistically significant differences in fear avoidance beliefs after the VR intervention but at the 3-month follow-up [47]. A systematic review and meta-analysis reported that VR therapy enhances spinal ROM and physical functioning in patients with CNP [26]. We failed to observe significant differences in the spinal ROM or disability level after VR intervention compared to those in the control group, which may be attributed to the relatively short duration (0-8 weeks) of the VR intervention (the reported mean duration was 4.81 weeks).

Limitations

Several limitations need to be addressed in this meta-analysis. First, the pooled analysis of the studies may be imprecise due to the large heterogeneity and the low quality of evidence from most of the included studies, and the results should be interpreted with caution. Second, the optimal duration of treatment for CSP could not be determined. Third, the effectiveness of VR therapy in patients with CSP and its analgesic effects in long-term follow-up must be further explored in high-quality studies. Fourth, indicators related to quality of life, such as depression and anxiety, should be emphasized and investigated in depth in future studies of patients with CSP.

Conclusion

VR therapy is an innovative and effective analgesic method that has beneficial effects on inflammatory markers in patients with CSP compared to other therapies (sham stimulation, usual care, conventional treatment). However, this approach may not have significant effects on the fear of movement, spinal ROM, or disability level. Notably, the quality of the evidence from the RCTs included in this study ranged from moderate to low. Therefore, we recommend that readers interpret the results of this study with caution. Future trials with large sample sizes, rigorous designs, and long-term follow-up periods are needed to explore the clinical significance of these differences and key issues in patients with CSP and to elucidate the underlying mechanisms of VR.

Acknowledgments

The authors thank Zhengquan Chen for providing support and advice. This work was supported by the Shanghai Hospital Development Center Foundation (SHDC12023118), the Medical Innovation Research of Shanghai Science and Technology Commission (22Y21900600), and the Three-year Action Plan of Shanghai Municipality to Further Accelerate the Inheritance, Innovation and Development of Traditional Chinese Medicine (2021-2023; ZY (2021- 2023)- 0201–05).

Data Availability

All data generated or analyzed during this study are included in this published paper and its supplementary information files.

Authors' Contributions

All authors contributed to the writing and redrafting of the manuscript. QD and XZ had the original idea. TZ and FW performed the literature search, XL and ZH assessed the risk of bias; YS, YF, and LZ rated the certainty of the evidence for each outcome; and XL and FW undertook data collection. The results were analyzed, interpreted, and discussed by XZ and QD. All authors contributed to the conception and design of the study, the analysis and interpretation of data, and the drafting and revising of the manuscript and have approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The PRISMA checklist. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis.
[DOCX File , 29 KB - games_v12i1e50089_app1.docx]

Multimedia Appendix 2
Search strategy for all electronic databases.
[DOCX File , 24 KB - games_v12i1e50089_app2.docx]

Multimedia Appendix 3
The GRADE criteria. GRADE: Grading of Recommendations, Assessment, Development and Evaluation.
[DOCX File , 22 KB - games_v12i1e50089_app3.docx]

Multimedia Appendix 4
Forest plots of the effect of VR compared with other treatments for pain intensity in patients with CSP: subgroup analysis of posttreatment effectiveness for treatment duration. CSP: chronic spinal pain; VR: virtual reality.
[DOCX File , 37 KB - games_v12i1e50089_app4.docx]

Multimedia Appendix 5
Forest plots of the effect of VR compared with other treatments for inflammatory marker level, fear of movement, spinal ROM, and disability level in patients with CSP. CSP: chronic spinal pain; ROM: range of motion; VR: virtual reality.
[DOCX File , 130 KB - games_v12i1e50089_app5.docx]

References

1. Chen S, Chen M, Wu X, Lin S, Tao C, Cao H, et al. Global, regional and national burden of low back pain 1990-2019: a systematic analysis of the Global Burden of Disease study 2019. J Orthop Translat 2022 Jan;32:49-58 [FREE Full text] [doi: 10.1016/j.jot.2021.07.005] [Medline: 34934626]
2. de Campos TF, Elkins MR. Prevention of spinal pain. J Physiother 2021 Jul;67(3):158-159 [FREE Full text] [doi: 10.1016/j.jphys.2021.06.012] [Medline: 34127420]
3. Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet 2012 Mar 04;379(9814):482-491. [doi: 10.1016/S0140-6736(11)60610-7] [Medline: 21982256]
4. Izzo R, Popolizio T, D'Aprile P, Muto M. Spinal pain. Eur J Radiol 2015 May;84(5):746-756. [doi: 10.1016/j.ejrad.2015.01.018] [Medline: 25824642]
5. Ferreira ML, de Luca K. Spinal pain and its impact on older people. Best Pract Res Clin Rheumatol 2017 Apr;31(2):192-202. [doi: 10.1016/j.berh.2017.08.006] [Medline: 29224696]
6. Khan AN, Jacobsen HE, Khan J, Filippi CG, Levine M, Lehman RA, et al. Inflammatory biomarkers of low back pain and disc degeneration: a review. Ann N Y Acad Sci 2017 Dec;1410(1):68-84 [FREE Full text] [doi: 10.1111/nyas.13551] [Medline: 29265416]
7. Ota Y, Connolly M, Srinivasan A, Kim J, Capizzano AA, Moritani T. Mechanisms and origins of spinal pain: from molecules to anatomy, with diagnostic clues and imaging findings. Radiographics 2020;40(4):1163-1181. [doi: 10.1148/rg.2020190185] [Medline: 32501739]
8. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain--United States, 2016. JAMA 2016 Apr 19;315(15):1624-1645 [FREE Full text] [doi: 10.1001/jama.2016.1464] [Medline: 26977696]

9. Machado GC, Maher CG, Ferreira PH, Day RO, Pinheiro MB, Ferreira ML. Non-steroidal anti-inflammatory drugs for spinal pain: a systematic review and meta-analysis. Ann Rheum Dis 2017 Jul;76(7):1269-1278. [doi: 10.1136/annrheumdis-2016-210597] [Medline: 28153830]
10. Schjerning A, McGettigan P, Gislason G. Cardiovascular effects and safety of (non-aspirin) NSAIDs. Nat Rev Cardiol 2020 Sep;17(9):574-584. [doi: 10.1038/s41569-020-0366-z] [Medline: 32322101]
11. Chou R, Qaseem A, Snow V, Casey D, Cross JT, Shekelle P, Clinical Efficacy Assessment Subcommittee of the American College of Physicians, American College of Physicians, American Pain Society Low Back Pain Guidelines Panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 2007 Oct 02;147(7):478-491 [FREE Full text] [doi: 10.7326/0003-4819-147-7-200710020-00006] [Medline: 17909209]
12. Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM, American Physical Therapy Association. Neck pain: clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopedic Section of the American Physical Therapy Association. J Orthop Sports Phys Ther 2008 Sep;38(9):A1-A34. [doi: 10.2519/jospt.2008.0303] [Medline: 18758050]
13. Malfliet A, Kregel J, Coppieters I, De Pauw R, Meeus M, Roussel N, et al. Effect of pain neuroscience education combined with cognition-targeted motor control training on chronic spinal pain: a randomized clinical trial. JAMA Neurol 2018 Jul 01;75(7):808-817 [FREE Full text] [doi: 10.1001/jamaneurol.2018.0492] [Medline: 29710099]
14. Zhou L, Zhang R, Shang J, Kang L, Zhang Z, Zhang B, et al. Comparative effectiveness of non-pharmacological interventions in reducing psychological symptoms among patients with chronic low back pain. Int J Surg 2023 Sep 27;110(1):478-489 [FREE Full text] [doi: 10.1097/JS9.0000000000000798] [Medline: 37755380]
15. Ho EK, Chen L, Simic M, Ashton-James CE, Comachio J, Wang DXM, et al. Psychological interventions for chronic, non-specific low back pain: systematic review with network meta-analysis. BMJ 2022 Mar 30;376:e067718 [FREE Full text] [doi: 10.1136/bmj-2021-067718] [Medline: 35354560]
16. Herman PM, Edgington SE, Ryan GW, Coulter ID. Prevalence and characteristics of chronic spinal pain patients with different hopes (treatment goals) for ongoing chiropractic care. J Altern Complement Med 2019 Oct;25(10):1015-1025 [FREE Full text] [doi: 10.1089/acm.2019.0247] [Medline: 31453711]
17. Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, Zepeda-Echavarria A, Caine-Winterberger K, Kulbacka-Ortiz K, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain. Lancet 2016 Dec 10;388(10062):2885-2894. [doi: 10.1016/S0140-6736(16)31598-7] [Medline: 27916234]
18. Chi B, Chau B, Yeo E, Ta P. Virtual reality for spinal cord injury-associated neuropathic pain: systematic review. Ann Phys Rehabil Med 2019 Jan;62(1):49-57 [FREE Full text] [doi: 10.1016/j.rehab.2018.09.006] [Medline: 30312665]
19. Hoffman HG, Doctor JN, Patterson DR, Carrougher GJ, Furness TA. Virtual reality as an adjunctive pain control during burn wound care in adolescent patients. Pain 2000 Mar;85(1-2):305-309. [doi: 10.1016/s0304-3959(99)00275-4] [Medline: 10692634]
20. Mirelman A, Rochester L, Maidan I, Del Din S, Alcock L, Nieuwhof F, et al. Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (V-TIME): a randomised controlled trial. Lancet 2016 Sep 17;388(10050):1170-1182. [doi: 10.1016/S0140-6736(16)31325-3] [Medline: 27524393]
21. Collado-Mateo D, Merellano-Navarro E, Olivares PR, García-Rubio J, Gusi N. Effect of exergames on musculoskeletal pain: a systematic review and meta-analysis. Scand J Med Sci Sports 2018 Mar 22;28(3):760-771. [doi: 10.1111/sms.12899] [Medline: 28452070]
22. Turoń-Skrzypińska A, Rotter I, Przybyciński J, Szylińska A, Mińko A, Ciechanowski K, et al. Does exercising with the use of virtual reality during haemodialysis have an impact on plasma levels of interleukin 1β, interleukin 6, and interleukin 8? J Clin Med 2023 Aug 17;12(16):5358 [FREE Full text] [doi: 10.3390/jcm12165358] [Medline: 37629400]
23. Klemm P, Kleyer A, Tascilar K, Schuster L, Meinderink T, Steiger F, et al. A virtual reality-based app to educate health care professionals and medical students about inflammatory arthritis: feasibility study. JMIR Serious Games 2021 May 11;9(2):e23835 [FREE Full text] [doi: 10.2196/23835] [Medline: 33973858]
24. Guo Q, Zhang L, Gui C, Chen G, Chen Y, Tan H, et al. Virtual reality intervention for patients with neck pain: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023 Apr 03;25:e38256 [FREE Full text] [doi: 10.2196/38256] [Medline: 37010891]
25. Ahern MM, Dean LV, Stoddard CC, Agrawal A, Kim K, Cook CE, et al. The effectiveness of virtual reality in patients with spinal pain: a systematic review and meta-analysis. Pain Pract 2020 Jul;20(6):656-675. [doi: 10.1111/papr.12885] [Medline: 32196892]
26. Brea-Gómez B, Torres-Sánchez I, Ortiz-Rubio A, Calvache-Mateo A, Cabrera-Martos I, López-López L, et al. Virtual reality in the treatment of adults with chronic low back pain: a systematic review and meta-analysis of randomized clinical trials. Int J Environ Res Public Health 2021 Nov 11;18(22):11806 [FREE Full text] [doi: 10.3390/ijerph182211806] [Medline: 34831562]

27. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019 Oct 03;10(10):ED000142 [FREE Full text] [doi: 10.1002/14651858.ED000142] [Medline: 31643080]
28. Li X, Shen J, Liang J, Zhou X, Yang Y, Wang D, et al. Effect of core-based exercise in people with scoliosis: a systematic review and meta-analysis. Clin Rehabil 2021 May;35(5):669-680 [FREE Full text] [doi: 10.1177/0269215520975105] [Medline: 33356498]
29. Higgins JJPT, Savovi? J, Page M, Elbers R, Sterne J. Chapter 8: assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August). Hoboken, NJ: John Wiley & Sons; 2023:2023 URL: https://training.cochrane.org/handbook/current/chapter-08
30. Jennions MD, Møller AP. Publication bias in ecology and evolution: an empirical assessment using the 'trim and fill' method. Biol Rev Camb Philos Soc 2002 May;77(2):211-222. [doi: 10.1017/s1464793101005875] [Medline: 12056747]
31. Schünemann H, Higgins JPT, Vist G, Glasziou P, Akl E, Skoetz N, et al. Chapter 14: completing 'summary of findings' tables and grading the certainty of the evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Hoboken, NJ: John Wiley & Sons; 2023. URL: https://training.cochrane.org/handbook/current/chapter-14
32. Garcia LM, Birckhead BJ, Krishnamurthy P, Sackman J, Mackey IG, Louis RG, et al. An 8-week self-administered at-home behavioral skills-based virtual reality program for chronic low back pain: double-blind, randomized, placebo-controlled trial conducted during COVID-19. J Med Internet Res 2021 Mar 22;23(2):e26292 [FREE Full text] [doi: 10.2196/26292] [Medline: 33484240]
33. Nambi G, Abdelbasset WK, Elsayed SH, Verma A, George JS, Saleh AK. Clinical and physical efficiency of virtual reality games in soccer players with low back pain. Rev Bras Med Esporte 2021 Dec;27(6):597-602 [FREE Full text] [doi: 10.1590/1517-8692202127062021_0034]
34. Nambi G, Abdelbasset WK, Elsayed SH, Alrawaili SM, Abodonya AM, Saleh AK, et al. Comparative effects of isokinetic training and virtual reality training on sports performances in university football players with chronic low back pain-randomized controlled study. Evid Based Complement Alternat Med 2020;2020:2981273 [FREE Full text] [doi: 10.1155/2020/2981273] [Medline: 32617104]
35. Yalfani A, Abedi M, Raeisi Z. Effects of an 8-week virtual reality training program on pain, fall risk, and quality of life in elderly women with chronic low back pain: double-blind randomized clinical trial. Games Health J 2022 Apr;11(2):85-92. [doi: 10.1089/g4h.2021.0175] [Medline: 35290742]
36. Park J, Lee S, Ko D. The effects of the Nintendo Wii exercise program on chronic work-related low back pain in industrial workers. J Phys Ther Sci 2013 Aug;25(8):985-988 [FREE Full text] [doi: 10.1589/jpts.25.985] [Medline: 24259899]
37. Afzal MW, Ahmad A, Mohseni Bandpei MA, Gilani SA, Hanif A, Waqas MS. Effects of virtual reality exercises and routine physical therapy on pain intensity and functional disability in patients with chronic low back pain. J Pak Med Assoc 2022 Mar;72(3):413-417 [FREE Full text] [doi: 10.47391/JPMA.3424] [Medline: 35320216]
38. Nambi G, Alghadier M, Kashoo FZ, Aldhafian OR, Nwihadh NA, Saleh AK, et al. Effects of virtual reality exercises versus isokinetic exercises in comparison with conventional exercises on the imaging findings and inflammatory biomarker changes in soccer players with non-specific low back pain: a randomized controlled trial. Int J Environ Res Public Health 2022 Dec 28;20(1):524 [FREE Full text] [doi: 10.3390/ijerph20010524] [Medline: 36612846]
39. Nambi G, Abdelbasset WK, Alqahatani BA. Radiological (magnetic resonance image and ultrasound) and biochemical effects of virtual reality training on balance training in football players with chronic low back pain: a randomized controlled study. J Back Musculoskelet Rehabil 2021;34(2):269-277. [doi: 10.3233/BMR-191657] [Medline: 33285623]
40. Nambi G, Abdelbasset W, Alsubaie S, Saleh AK, Verma A, Abdelaziz MA, et al. Short-term psychological and hormonal effects of virtual reality training on chronic low back pain in soccer players. J Sport Rehabil 2021 Mar 16;30(6):884-893. [doi: 10.1123/jsr.2020-0075] [Medline: 33596538]
41. Matheve T, Bogaerts K, Timmermans A. Virtual reality distraction induces hypoalgesia in patients with chronic low back pain: a randomized controlled trial. J Neuroeng Rehabil 2020 May 22;17(1):55 [FREE Full text] [doi: 10.1186/s12984-020-00688-0] [Medline: 32321516]
42. Stamm O, Dahms R, Reithinger N, Ruß A, Müller-Werdan U. Virtual reality exergame for supplementing multimodal pain therapy in older adults with chronic back pain: a randomized controlled pilot study. Virtual Real 2022 Feb 11;26(4):1291-1305 [FREE Full text] [doi: 10.1007/s10055-022-00629-3] [Medline: 35194374]
43. Monteiro-Junior RS, de Souza CP, Lattari E, Rocha NBF, Mura G, Machado S, et al. Wii-workouts on chronic pain, physical capabilities and mood of older women: a randomized controlled double blind trial. CNS Neurol Disord Drug Targets 2015;14(9):1157-1164. [doi: 10.2174/1871527315666151111120131] [Medline: 26556092]
44. Cetin H, Kose N, Oge HK. Virtual reality and motor control exercises to treat chronic neck pain: a randomized controlled trial. Musculoskelet Sci Pract 2022 Dec 05;62:102636. [doi: 10.1016/j.msksp.2022.102636] [Medline: 35952621]
45. Sarig Bahat H, Croft K, Carter C, Hoddinott A, Sprecher E, Treleaven J. Remote kinematic training for patients with chronic neck pain: a randomised controlled trial. Eur Spine J 2018 Jul;27(6):1309-1323. [doi: 10.1007/s00586-017-5323-0] [Medline: 29018956]

46. Nusser M, Knapp S, Kramer M, Krischak G. Effects of virtual reality-based neck-specific sensorimotor training in patients with chronic neck pain: a randomized controlled pilot trial. J Rehabil Med 2021 Mar 10;53(2):jrm00151 [FREE Full text] [doi: 10.2340/16501977-2786] [Medline: 33369684]
47. Tejera DM, Beltran-Alacreu H, Cano-de-la-Cuerda R, Leon Hernández JV, Martín-Pintado-Zugasti A, Calvo-Lobo C, et al. Effects of virtual reality versus exercise on pain, functional, somatosensory and psychosocial outcomes in patients with non-specific chronic neck pain: a randomized clinical trial. Int J Environ Res Public Health 2020 Aug 16;17(16):5950 [FREE Full text] [doi: 10.3390/ijerph17165950] [Medline: 32824394]
48. Comparcini D, Simonetti V, Galli F, Saltarella I, Altamura C, Tomietto M, et al. Immersive and non-immersive virtual reality for pain and anxiety management in pediatric patients with hematological or solid cancer: a systematic review. Cancers (Basel) 2023 Mar 03;15(3):985 [FREE Full text] [doi: 10.3390/cancers15030985] [Medline: 36765945]
49. Gupta A, Scott K, Dukewich M. Innovative technology using virtual reality in the treatment of pain: does it reduce pain via distraction, or is there more to it? Pain Med 2018 Jan 01;19(1):151-159. [doi: 10.1093/pm/pnx109] [Medline: 29025113]
50. Ostelo RWJG, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976) 2008 Jan 01;33(1):90-94. [doi: 10.1097/BRS.0b013e31815e3a10] [Medline: 18165753]
51. Morris P, Ali K, Merritt M, Pelletier J, Macedo LG. A systematic review of the role of inflammatory biomarkers in acute, subacute and chronic non-specific low back pain. BMC Musculoskelet Disord 2020 Mar 03;21(1):142 [FREE Full text] [doi: 10.1186/s12891-020-3154-3] [Medline: 32126991]
52. Eccleston C, Fisher E, Liikkanen S, Sarapohja T, Stenfors C, Jääskeläinen SK, et al. A prospective, double-blind, pilot, randomized, controlled trial of an "embodied" virtual reality intervention for adults with low back pain. Pain 2022 Oct 01;163(9):1700-1715 [FREE Full text] [doi: 10.1097/j.pain.0000000000002617] [Medline: 35324507]
53. Lier EJ, de Vries M, Steggink EM, Ten Broek RPG, van Goor H. Effect modifiers of virtual reality in pain management: a systematic review and meta-regression analysis. Pain 2023 Aug 01;164(8):1658-1665 [FREE Full text] [doi: 10.1097/j.pain.0000000000002883] [Medline: 36943251]
54. Kato S, Murakami H, Demura S, Yoshioka K, Shinmura K, Yokogawa N, et al. Abdominal trunk muscle weakness and its association with chronic low back pain and risk of falling in older women. BMC Musculoskelet Disord 2019 Jul 03;20(1):273 [FREE Full text] [doi: 10.1186/s12891-019-2655-4] [Medline: 31159812]
55. Kitagawa R, Kato S, Demura S, Kurokawa Y, Shinmura K, Yokogawa N, et al. Efficacy of abdominal trunk muscles-strengthening exercise using an innovative device in treating chronic low back pain: a controlled clinical trial. Sci Rep 2020 Dec 14;10(1):21883 [FREE Full text] [doi: 10.1038/s41598-020-78908-9] [Medline: 33318516]
56. Mao Y, Chen P, Li L, Huang D. Virtual reality training improves balance function. Neural Regen Res 2014 Oct 01;9(17):1628-1634 [FREE Full text] [doi: 10.4103/1673-5374.141795] [Medline: 25368651]
57. Paalanne N, Niinimäki J, Karppinen J, Taimela S, Mutanen P, Takatalo J, et al. Assessment of association between low back pain and paraspinal muscle atrophy using opposed-phase magnetic resonance imaging: a population-based study among young adults. Spine (Phila Pa 1976) 2011 Dec 01;36(23):1961-1968. [doi: 10.1097/BRS.0b013e3181fef890] [Medline: 21289551]
58. Currie SR, Wang J. Chronic back pain and major depression in the general Canadian population. Pain 2004 Jan;107(1-2):54-60. [doi: 10.1016/j.pain.2003.09.015] [Medline: 14715389]
59. Janssen SA, Arntz A. Anxiety and pain: attentional and endorphinergic influences. Pain 1996 Aug;66(2-3):145-150. [doi: 10.1016/0304-3959(96)03031-x] [Medline: 8880835]
60. Turk DC. Understanding pain sufferers: the role of cognitive processes. Spine J 2004;4(1):1-7. [doi: 10.1016/s1529-9430(03)00068-8] [Medline: 14749188]
61. Larson SL, Clark MR, Eaton WW. Depressive disorder as a long-term antecedent risk factor for incident back pain: a 13-year follow-up study from the Baltimore Epidemiological Catchment Area sample. Psychol Med 2004 Mar;34(2):211-219. [doi: 10.1017/s0033291703001041] [Medline: 14982127]
62. Huang Q, Lin J, Han R, Peng C, Huang A. Using virtual reality exposure therapy in pain management: a systematic review and meta-analysis of randomized controlled trials. Value Health 2022 Mar;25(2):288-301 [FREE Full text] [doi: 10.1016/j.jval.2021.04.1285] [Medline: 35094802]
63. Gold JI, Belmont KA, Thomas DA. The neurobiology of virtual reality pain attenuation. Cyberpsychol Behav 2007 Aug;10(4):536-544. [doi: 10.1089/cpb.2007.9993] [Medline: 17711362]
64. Le Maitre CL, Hoyland JA, Freemont AJ. Catabolic cytokine expression in degenerate and herniated human intervertebral discs: IL-1beta and TNFalpha expression profile. Arthritis Res Ther 2007;9(4):R77 [FREE Full text] [doi: 10.1186/ar2275] [Medline: 17688691]
65. Phillips KLE, Jordan-Mahy N, Nicklin MJH, Le Maitre CL. Interleukin-1 receptor antagonist deficient mice provide insights into pathogenesis of human intervertebral disc degeneration. Ann Rheum Dis 2013 Dec;72(11):1860-1867. [doi: 10.1136/annrheumdis-2012-202266] [Medline: 23396662]
66. Nambi G, Abdelbasset WK, Elsayed SH, Khalil MA, Alrawaili SM, Alsubaie SF. Comparative effects of virtual reality training and sensory motor training on bone morphogenic proteins and inflammatory biomarkers in post-traumatic osteoarthritis. Sci Rep 2020 Oct 28;10(1):15864 [FREE Full text] [doi: 10.1038/s41598-020-72587-2] [Medline: 32985509]

67. Matheve T, De Baets L, Bogaerts K, Timmermans A. Lumbar range of motion in chronic low back pain is predicted by task-specific, but not by general measures of pain-related fear. Eur J Pain 2019 Jul;23(6):1171-1184. [doi: 10.1002/ejp.1384] [Medline: 30793429]
68. Rudolfsson T, Björklund M, Djupsjöbacka M. Range of motion in the upper and lower cervical spine in people with chronic neck pain. Man Ther 2012 Mar;17(1):53-59. [doi: 10.1016/j.math.2011.08.007] [Medline: 21945295]
69. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. Lancet 2018 Jul 09;391(10137):2356-2367 [FREE Full text] [doi: 10.1016/S0140-6736(18)30480-X] [Medline: 29573870]
70. Saavedra-Hernández M, Castro-Sánchez AM, Cuesta-Vargas AI, Cleland JA, Fernández-de-las-Peñas C, Arroyo-Morales M. The contribution of previous episodes of pain, pain intensity, physical impairment, and pain-related fear to disability in patients with chronic mechanical neck pain. Am J Phys Med Rehabil 2012 Dec;91(12):1070-1076. [doi: 10.1097/PHM.0b013e31827449a5] [Medline: 23159953]

Abbreviations

CLBP: chronic low back pain
CNP: chronic neck pain
CRP: C-reactive protein
CSP: chronic spinal pain
DVPRS: Defense and Veterans Pain Rating Scale
GRADE: Grading of Recommendations, Assessment, Development and Evaluation
IL: interleukin
IVR: immersive virtual reality
LBP: low back pain
NDI: Neck Disability Index
NIVR: nonimmersive virtual reality
NPRS: Numerical Pain Rating Scale
NRS: Numeric Rating Scale
NSAID: nonsteroidal anti-inflammatory drug
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCT: randomized controlled trial
ROM: range of motion
TNF-α: tumor necrosis factor-alpha
TSK: Tampa Scale of Kinesiophobia
VAS: Visual Analogue Scale
VR: virtual reality
VRE: virtual reality exercise
WMD: weighted mean difference

Edited by T Leung; submitted 19.06.23; peer-reviewed by Y Jinqing, A Ollevier; comments to author 31.10.23; revised version received 01.12.23; accepted 07.01.24; published 12.02.24.

Please cite as:
Zhang T, Li X, Zhou X, Zhan L, Wu F, Huang Z, Sun Y, Feng Y, Du Q
Virtual Reality Therapy for the Management of Chronic Spinal Pain: Systematic Review and Meta-Analysis
JMIR Serious Games 2024;12:e50089
URL: https://games.jmir.org/2024/1/e50089
doi:10.2196/50089
PMID:38345832

Review

Extended Reality–Based Mobile App Solutions for the Therapy of Children With Autism Spectrum Disorders: Systematic Literature Review

Marian-Vladut Toma[1*], MSci; Cristina Elena Turcu[2*], PhD; Corneliu Octavian Turcu[2*], PhD; Sorin Vlad[1*], PhD; Doru Eugen Tiliute[1*], PhD; Paul Pascu[1*], PhD

[1]Faculty of Economics, Administration and Business, "Stefan cel Mare" University of Suceava, Suceava, Romania

[2]Faculty of Electrical Engineering and Computer Science, University of Suceava, Suceava, Romania

[*]all authors contributed equally

Corresponding Author:
Marian-Vladut Toma, MSci
Faculty of Economics, Administration and Business
"Stefan cel Mare" University of Suceava
Universitatii Street nr.13
Suceava, 720229
Romania
Phone: 40 752425739
Email: vlad.toma@usm.ro

Abstract

Background: The increasing prevalence of autism spectrum disorder (ASD) has driven research interest on the therapy of individuals with autism, especially children, as early diagnosis and appropriate treatment can lead to improvement in the condition. With the widespread availability of virtual reality, augmented reality (AR), and mixed reality technologies to the public and the increasing popularity of mobile devices, the interest in the use of applications and technologies to provide support for the therapy of children with autism is growing.

Objective: This study aims to describe the literature on the potential of virtual reality, AR, and mixed reality technologies in the context of therapy for children with ASD. We propose to investigate and analyze the temporal distribution of relevant papers, identify the target audience for studies related to extended reality apps in ASD therapy, examine the technologies used in the development of these apps, assess the skills targeted for improvement in primary studies, explore the purposes of the proposed solutions, and summarize the results obtained from their application.

Methods: For the systematic literature review, 6 research questions were defined in the first phase, after which 5 international databases (Web of Science, Scopus, ScienceDirect, IEEE Xplore Digital Library, and ACM Digital Library) were searched using specific search strings. Results were centralized, filtered, and processed applying eligibility criteria and using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The results were refined using a technical and IT-oriented approach. The quality criteria assessed whether the research addressed ASDs, focused on children's therapy, involved targeted technologies, deployed solutions on mobile devices, and produced results relevant to our study.

Results: In the first step, 179 publications were identified in Zotero reference manager software (Corporation for Digital Scholarship). After excluding articles that did not meet the eligibility or quality assessment criteria, 28 publications were finalized. The analysis revealed an increase in publications related to apps for children with autism starting in 2015 and peaking in 2019. Most studies (22/28, 79%) focused on mobile AR solutions for Android devices, which were developed using the Unity 3D platform and the Vuforia engine. Although 68% (19/28) of these apps were tested with children, 32% (9/28) were tested exclusively by developers. More than half (15/28, 54%) of the studies used interviews as an evaluation method, yielding mostly favorable although preliminary results, indicating the need for more extensive testing.

Conclusions: The findings reported in the studies highlight the fact that these technologies are appropriate for the therapy of children with ASD. Several studies showed a distinct trend toward the use of AR technology as an educational tool for people with ASD. This trend entails multidisciplinary cooperation and an integrated research approach, with an emphasis on comprehensive empirical evaluations and technology ethics.

(JMIR Serious Games 2024;12:e49906) doi:10.2196/49906

KEYWORDS

autism; autistic; autism spectrum disorder; ASD; virtual reality; augmented reality; extended reality; mixed reality; mobile app; children; preschool; mobile phone

Introduction

Background

In recent years, there has been increased interest in using technology to address the unique challenges faced by individuals with autism spectrum disorder (ASD). Among the various technological approaches, extended reality (XR), which includes augmented reality (AR) and virtual reality (VR), has emerged as a promising solution for intervention and therapy in children with ASD. XR offers the potential to create immersive and engaging environments that can address the specific needs of individuals on the autism spectrum, assisting them with communication, social interaction, and skill development. As a result, researchers and practitioners have explored the development of XR-based mobile apps tailored to the therapy of children with ASD.

However, the rapid growth in this field has spawned a multitude of XR-based mobile app solutions, each claiming unique benefits and features. With this proliferation of interventions, it is important to comprehensively assess the current landscape of XR-based mobile apps for the therapy of children with ASD, not only to strengthen existing knowledge in the field but also to provide critical insights into the research field.

In light of these considerations, this systematic literature review aimed to explore and assess the current status of XR-based mobile app solutions for the therapy of children with ASD. By synthesizing evidence from existing studies, this review aimed to provide an updated overview of the field, identify research gaps, and provide valuable insights. In this endeavor, this review aimed to contribute to the advancement of knowledge and practice in the field of XR-based interventions for ASD therapy, ultimately aiming to improve the quality of life of children on the autism spectrum and their caregivers.

ASD is a neurological condition that has a significant negative impact on a person's social, verbal, and physical abilities. Researchers claim that ASD is typically discovered around the third year of life [1], but it can be identified and diagnosed as early as the age of 18 months [2]. According to a study from 2022, a total of 1% of infants have ASD [3]. On the basis of studies conducted over the past 50 years, the World Health Organization predicts a global increase in the prevalence of ASD [3].

Researchers consider 3 techniques that could be used to facilitate the evaluation of this condition's prevalence: providing diagnostic tools, enhancing diagnostic standards, and increasing public awareness of ASDs [4].

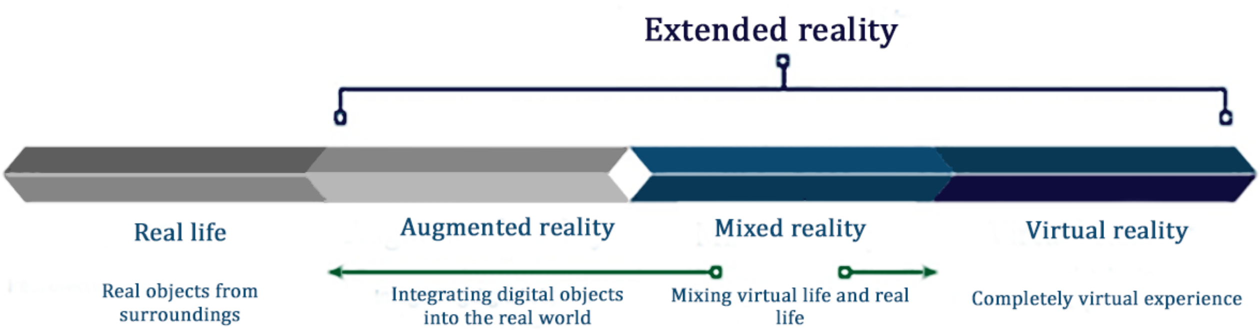
On the basis of each individual's verbal IQ and level of language delay, the diagnosis assesses the severity of the disorder (mild, moderate, or severe) [5], which can estimate the extent to which daily life is affected. Many people with ASD are timid, have difficulty communicating, or experience anxiety when engaging in casual conversation. Despite their communication and social skill deficits, individuals with ASD have demonstrated a preference for technology [6]. Furthermore, the use of technology in behavioral therapy for people with ASD has the additional advantage of being cost-effective in terms of both caregiver and treatment facility expenses [7]. In addition, various studies have shown that individuals with ASD respond better to visual stimuli than to other sensory stimuli [8]. The findings of these studies have led to various applications of technology in digital behavioral treatment.

XR, AR, VR, and Mixed Reality

As evidenced, the evolution of IT has accelerated in recent years. According to the study by Abad-Segura et al [9], rapid technological advancements have caused a significant and positive shift in how people view modern living.

The relatively new term XR refers to the entire spectrum from AR to VR, including mixed reality (MR; Figure 1 [10]).

Figure 1. The extended reality concept [10].



AR is a technology that enables real-time interaction and integration of 3D virtual models into the physical world [11]. Although the first portable AR system was developed in 2003 [12], AR did not acquire widespread acceptance and public awareness until the release of the mobile game Pokémon GO [13] in 2016. Despite the various implementation challenges, AR has many potential applications. In addition to applications in specific domains such as industry, construction, or medicine, as well as in advertising and commerce, education, and gaming [14], AR can be used by a broader audience for everyday tasks such as finding information about nearby points of interest, navigation, and assistance while following a route [14]. Most of these apps are now accessible owing to advances in mobile device technology and the spread of smart mobile phones. AR facilitates behavioral therapy by enhancing the experiences and abilities of people with ASD and establishing an integrated learning environment that enables the visualization of educational materials in 3D and engaging manipulation of real-world objects [15]. By generating "physical" structures to improve specific skills, AR fosters the imagination of patients with ASD without impairing it [16,17]. Moreover, AR can be used to create more engaging and appealing user interfaces, thereby eliminating the need for conventional input devices such as a keyboard and mouse [18]. AR technology is typically accessed using various devices and platforms. Among the widely used platforms and tools for developing AR apps are Unity, Unreal Engine, ARCore, and HP Reveal.

As described in the study by Azuma [11], VR is a computer-generated environment that simulates real-life scenarios, creating an immersive and interactive experience. This means that the users are placed in a completely virtual world, which can be similar to or different from the real one. This technology requires specialized equipment, such as VR headsets or glasses to enable users to see and interact with the virtual environment.

As seen in the studies by Bursali and Yilmaz [19] and El-Jarn and Southern [10], MR is situated between AR and VR, integrating the 2 technologies to provide the user with a unique and captivating experience in real time. It can be difficult to precisely define the limits of MR as they depend on the devices and equipment used as well as the extent to which VR or AR is incorporated into the final product. A model describing the integration of digital objects from the physical world into the virtual world is shown in the study by Milgram and Kishino [20], which also presented a taxonomy for MR, stating that it can be defined as a part of the human-computer interface field, which integrates VR and AR elements to create an environment in which virtual and real objects coexist and interact.

To be used, technologies from the XR spectrum require specific hardware with an optical sensor [19]. In addition, well-known technology companies such as Google, Facebook, Apple, Amazon, and Microsoft have significantly contributed to the development of AR tools and services [21,22], including handheld devices; holographic screens (Microsoft HoloLens); and heads-up displays, which are mainly designed for MR, tablets, and mobile devices (smartphones).

The development of collaborative XR, which enables simultaneous communication and collaboration among multiple users, is one of the research trends in the field of XR [23].

Given the benefits that AR, VR, and MR can offer as a new mode of human-computer interaction and the fact that these technologies are becoming ubiquitous and part of our daily lives, this systematic review aimed to describe how these technologies can be used in the therapy of children with ASDs.

Methods

Overview

According to Kitchenham [24], a systematic literature review is a method for identifying, evaluating, and interpreting all available research relevant to a field of study as well as answering specific research questions (RQs). We conducted this research following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [25] and the recommendations suggested by Kitchenham [24].

To conduct this literature review, several well-known scientific databases were queried, and publications containing relevant information for our analysis were filtered. We defined RQs and provided answers to each of them, thus achieving our proposed objective.

Search Strategy

According to the considered methodology, the following 7 RQs were formulated. These questions consider aspects relevant to the understanding of concepts important to this study:

1. What is the papers' distribution over time? (RQ 1)
2. What category of people are the studies aimed at? (RQ 2)
3. Which technologies are used with XR or any of its subdivisions to develop apps for ASD therapy? (RQ 3)
4. What skills were targeted for improvement in primary studies? (RQ 4)
5. What are the purposes for which the proposed solutions were used? (RQ 5)
6. What are the results obtained using the proposed solutions? (RQ 6)

To study the literature and answer the aforementioned questions, we searched for scientific publications using various academic research databases. Our study primarily focused on the technical aspects of mobile app solutions using XR for autism therapy. To comprehensively cover our research domain, we chose to use multidisciplinary scientific databases—Scopus, ScienceDirect, and Web of Science—along with 2 databases particularly relevant to computer science, namely, IEEE Xplore Digital Library and ACM Digital Library. From these sources, we only considered publications that were relevant in computer science–related categories, such as technology, engineering, and computer science, excluding categories related to medicine, chemistry, or neurosciences considering that the RQs were focused not only on the available apps but also on their technical details. The functionalities, the technologies used, and the entire process of their development also constituted an objective. Thus, the approach from a technical point of view and the development of these apps were followed. This was done using the results refinement interface available in the aforementioned databases. Initially, to view and analyze the results of queries conducted using the considered search strings, the search was not limited to a particular time.

Given the topic of this study, we aimed to query scientific databases so that the resulting list of publications would meet the following criteria:

1. Reference to ASDs
2. Consideration of one of the technologies that are part of the concept of XR (VR, AR, or MR)
3. Addressing mobile apps
4. Aim to develop solutions for the therapy of children

The literature was searched using keywords relevant to achieving the proposed objectives: *autism*, *autistic*, *ASD*, *virtual reality*, *augmented reality*, *extended reality*, *mixed reality*, *mobile application*, and *children*.

Following the analysis of these keywords, the query process was extended by including the following terms: *Autis**, *VR*, *AR*, *MR*, *XR*, *Mobile app**, *Smartphone app**, *Child**, *Infan**, *Toddler**, *Preschool**, *Kid**, and *Juvenile*. In the aforementioned list, an asterisk stands for any number of characters at the end of the current string (eg, *Preschool** refers to *Preschool*, *Preschooler*, and *Preschoolers*).

Information Sources

Depending on the search options available in each database considered, specific search strings were defined for querying the databases (Table 1). These query strings were defined using advanced search functions and appropriate operators. The search of Web of Science and Scopus publications was performed by title, abstract, and keywords, and IEEE Xplore Digital Library, ScienceDirect, and ACM Digital Library were searched using a general search. The queries were executed on December 18, 2022.

Table 1. The search strings used for querying the databases (N=219).

Item	Database	Search string	Returned results, n (%)
1	Web of Science	(TS=(Autis*) OR TS=(ASD)) AND (TS=(virtual reality OR VR) OR TS=(augmented reality OR AR) OR TS=(mixed reality OR MR) OR TS=(extended reality OR XR)) AND (TS=(mobile app* OR smartphone app*)) AND (TS=(child* OR infan* OR toddler* OR preschool* OR kid* OR juvenile))	45 (20.5)
2	Scopus	(TITLE-ABS-KEY (autis*) OR TITLE-ABS-KEY (asd)) AND (TITLE-ABS-KEY ("virtual reality" OR vr) OR TITLE-ABS-KEY ("augmented reality" OR ar) OR TITLE-ABS-KEY ("mixed reality" OR mr) OR TITLE-ABS-KEY ("extended reality" OR xr)) AND (TITLE-ABS-KEY ("mobile app*" OR "smartphone app*")) AND (TITLE-ABS-KEY (child* OR infan* OR toddler* OR preschool* OR kid* OR juvenile))	32 (14.6)
3	IEEE Xplore Digital Library	(Autis* OR ASD) AND (Augmented reality OR AR OR Mixed reality OR MR OR Extended Reality OR XR OR Virtual Reality OR VR) AND (Mobile OR Tablet OR Smartphone OR Phone OR Smartglass) AND (App* OR Solution*) AND (child* OR kid* OR infan* OR preschool* OR juvenile OR toddler*)	26 (11.9)
4	ScienceDirect	("Autism Spectrum Disorder" OR "ASD") AND ("Augmented reality" OR "Mixed reality" OR "Extended reality") AND ("App OR Application") AND ("kids OR children")	48 (21.9)
5	ACM Digital Library	[[All: "autis*"] OR [All: "asd"]] AND [[All: "augmented reality"] OR [All: "mixed reality"] OR [All: "virtual reality"] OR [All: "extended reality"]] AND [All: "mobile"] AND [[All: "app"] OR [All: application]]	68 (31.1)

Eligibility Criteria

The papers obtained by querying scientific databases had an interdisciplinary nature. However, our study took a technical and IT-focused approach to mobile app solutions using XR for autism therapy. Therefore, we needed to refine the results by considering inclusion and exclusion criteria. As previously stated, no constraints were imposed on the publication dates of the articles during the search conducted in the scientific databases. Nevertheless, considering the significant progress in mobile device capabilities and their widespread use over the last decade, which have facilitated the development and growth of the global use of XR-based mobile apps for therapeutic purposes, we focused our investigation on the period following 2012 [26,27]. In line with our technical focus on mobile app solutions using XR for the therapy of children with ASD, we refined the search results across the 5 considered databases, prioritizing computer science–related domains. We deliberately excluded categories related to medicine, chemistry, or neurosciences as our RQs focused on both the available apps and their technical details.

Before centralizing the results for analysis, they were refined according to the inclusion and exclusion criteria. The inclusion criteria were as follows:

1. Articles published in English
2. Articles published after 2012

The exclusion criteria were as follows:

1. Book chapters
2. Paper tables of contents
3. Articles published in languages other than English
4. Results on the topics of medicine, chemistry, or neurosciences

After initial processing, the database searches returned the number of results presented in Table 2.

Table 2. Results obtained after initial data processing (N=179).

Database	Results, n (%)
Web of Science	43 (24)
Scopus	22 (12.3)
IEEE Xplore Digital Library	25 (14)
ScienceDirect	41 (22.9)
ACM Digital Library	48 (26.8)

Selection Process

During the selection process, the PRISMA guidelines were considered [25]. These guidelines outline 3 steps: identification (centralizing the results and excluding duplicate publications), screening (review of titles and abstracts and testing eligibility), and inclusion (the publications identified as answering the proposed RQs). The PRISMA 2020 checklist is available in Multimedia Appendix 1. The Zotero reference manager software (Corporation for Digital Scholarship) was used to perform the specified steps. In the first step, 179 publications resulting from the search of the 5 considered databases were imported, after which duplicate publications (n=20, 11.2%) and some conference papers (n=4, 2.2%) were removed. For the next step, 86.6% (155/179) of the publications were considered. To enable author collaboration, the data were imported into Google Sheets. In total, 2 reviewers (M-VT and CET) conducted an independent screening of publications for inclusion based on title and abstract analysis. Studies meeting the eligibility criteria according to both reviewers were then considered for full-text screening. Any disagreements were discussed face-to-face between the reviewers, and a third party was involved to help reach unanimity where necessary. The same process was implemented for the full-text review with the assistance of a third reviewer (SV).

After reviewing the titles and abstracts, a total of 41.9% (65/155) of the publications were excluded as they did not address the proposed topic, focusing either on another condition or on other technologies.

Despite the high quality of the publications, as evidenced by their indexing in prestigious international databases, the analysis included a full text review (where available) of the remaining 58.1% (90/155) of the publications, and the following quality assessment criteria were applied to ensure their relevance to the RQs considered. Articles with no full text accessible were excluded. The following quality criteria (QCs) were applied to 78 publications:

1. Does the research topic address ASDs? (QC 1)
2. Does the study address children’s therapy? (QC 2)
3. Does the study include one of the technologies targeted in this review? (QC 3)
4. Is the solution deployed on a mobile device? (QC 4)
5. Are the results relevant to this review? (QC 5)

Given the nature of the research and its objectives, the 5 QCs that were developed specifically to achieve the goals of the research were used to evaluate the studies’ quality by 2 authors. Each publication was carefully reviewed and assigned a score from 0 to 2 measuring the extent to which it corresponded to the quality assessment criteria and to the subject of this study (0=no; 1=partially; 2=yes). Thus, the maximum score for a paper could be 10. After that, the data were combined for further analysis using Microsoft Excel (Microsoft Corp). After consolidation, a score with a decimal part (eg, 1.5) was rounded up to the nearest integer for better inclusion.

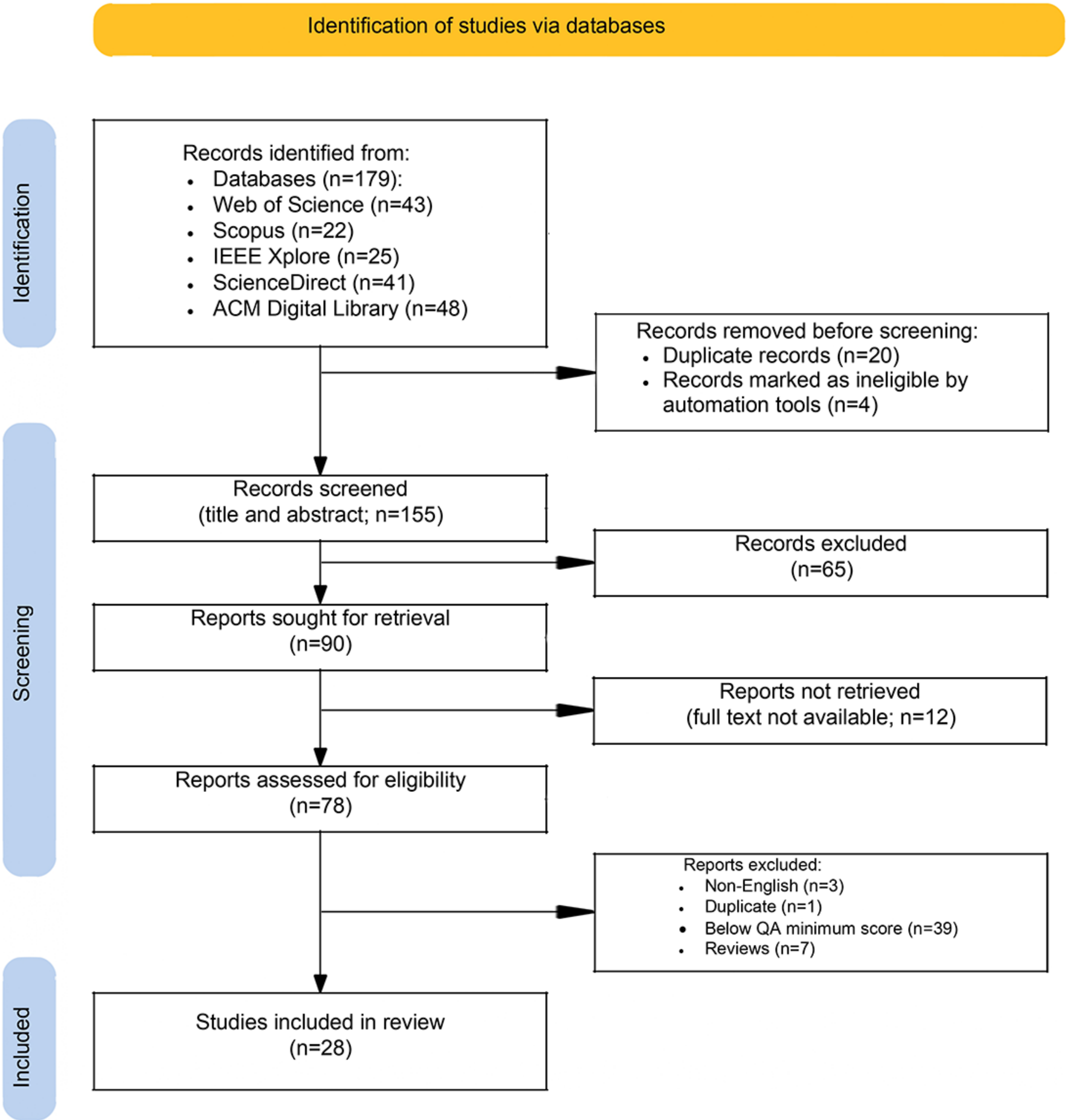
After reviewing the full texts, 3 publications were found to be not written in English, and in other publications, ASD was not addressed (only mentioned), other technologies were addressed, or the proposed solution was not clearly described and the results were inconclusive (see the sample in Figure 2; the entire table is available in Multimedia Appendix 2 [28-104]).

Following this, only publications with a score of ≥7 were evaluated as they adequately addressed the RQs. The type of publication, whether it was a review or aimed at developing an app, was also noted. Systematic literature review publications were investigated to identify any references that could be added to this study, but they were removed from the list after being reviewed. The entire publication selection process is illustrated in Figure 3.

Figure 2. Screenshot of the quality assessment of the papers.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	No.	Item Type	Contribution or review	Publication Year	Author	Title	Publication Title	Abstract Note	QC1	QC2	QC3	QC4	QC5	Total
2	1	journalArticle	review	2021	Lian, Xiaojie; Sunar, N	Mobile Augment	APPLIED SCIEN	Over the past decade, e	2	2	2	2	2	10
3	2	journalArticle	contribution	2022	Lopez-Belmonte, Jesu	The Impact of Ge	EDUCATION IN	Today's education is in t	2	2	0	0	1	5
4	3	conferencePaper	contribution	2016	Gea, M.; Alaman, X.;	TOWARDS SMAR	EDULEARN16:	Smart city is a promising	1	1	1	0	1	4
5	4	conferencePaper	contribution	2013	Vullamparthi, Annie J	Assistive Learning	2013 IEEE FIFT	Innovative computer aid	2	2	2	2	2	10
6	5	journalArticle	review	2020	Baragash, Reem Sula	Augmented reali	EUROPEAN JOU	There is a growing inter	1	1	1	1	1	5
7	6	journalArticle	contribution		Glaser, Noah; Newbu	Video-Based Virt	TECHNOLOGY	Research suggests that	2	0	2	1	1	6
8	7	conferencePaper	contribution	2020	Bouaziz, Rahma; Alhe	Using Marker Ba	2020 IEEE INTE	Autistic children suffer f	2	2	2	2	2	10
9	8	conferencePaper	contribution	2017	Pradibta, Hendra; Wi	Designing AR Dai	2017 INTERNA	An Individual with Autis	2	2	2	2	1	9
10	9	journalArticle	contribution	2019	Abou El-Seoud, M. Sa	Assisting Individu	INTERNATION	Individuals with autism	2	2	2	2	2	10
11	10	journalArticle	contribution	2022	Hashim, Haida Umier	`AReal-Vocab': A	SUSTAINABILIT	The American Psychiatri	2	2	2	2	2	10
12	11	conferencePaper	contribution	2018	Pradibta, Hendra; Wi	The evaluation o	3RD ANNUAL	ARdoa is Mobile Augme	0	0	0	0	0	0
13	12	journalArticle	contribution	2022	Wan, Guobin; Deng,	FECTS: A Facial E	COMPUTATIO	Traditional training met	2	2	0	1	2	7
14	13	conferencePaper	contribution	2015	Nubia, Rincon M.; Fal	Development of	2015 WORKSH	An augmented reality m	2	2	2	2	2	10
15	14	conferencePaper	contribution	2020	Wang, Katherine; Zha	Using Mobile Aug	CHI'20: EXTEN	Adults on the autism spe	2	0	2	2	2	8
16	15	conferencePaper	contribution	2019	Tang, Tiffany Y.; Xu, J	Automatic Object	3RD INTERNAT	A number of previous co	2	2	2	2	2	10
17	16	journalArticle	contribution	2021	Hashim, Haida Umier	Augmented Reali	ARAB WORLD	American Psychiatric As	0	0	0	0	0	0
18	17	conferencePaper	contribution	2021	Kavitha, V; Mohan, A	Immersive Learni	ICSPC'21: 2021	Autism Spectrum Disord	2	2	2	2	2	10
19	18	journalArticle	review	2018	Johnston, Daniel; Ege	Innovative compu	COGENT PSYCH	Individuals with autism	2	2	0	0	1	5
20	19	journalArticle	review	2019	Putnam, Cynthia; Har	Interactive Techn	ACM TRANSAC	Autism spectrum disord	2	2	0	1	1	6
21	20	journalArticle	review		Mukherjee, Debarati;	Digital tools for d	AUTISM	Current challenges in ea	2	0	1	1	0	4
22	21	journalArticle	review	2020	Valentine, Althea Z.;	A systematic revi	CLINICAL PSYC	Technology-based interv	1	2	2	2	1	8
23	22	conferencePaper	contribution	2022	T, Nithya Shree; Selva	An Interactive Nu	2022 Internati	Autism in children is cha	2	2	2	2	2	10
24	23	conferencePaper	review	2019	Khowaja, Kamran; Al-	Use of augmente	2019 IEEE 6th	The purpose of this revie	2	2	2	1	2	9
25	24	conferencePaper	review	2020	Almurashi, Hanin A.;	Augmented Reali	2020 IEEE Inte	Currently, augmented r	2	1	1	1	1	6
26	25	conferencePaper	contribution	2020	Xia, Mengting; Chen,	ParaShop: A Mob	2020 IEEE MIT	Approximately 1 in 160	2	2	2	2	2	10
27	26	conferencePaper	contribution	2017	Silva, Salatiel Dantas;	Knowledgemon H	2017 19th Sym	Autism Spectrum Disord	2	2	2	2	1	9
28	27	journalArticle	review	2021	Farzana, Walia; Sarke	Technological Ev	IEEE Access	Augmentative and Alter	2	2	1	2	2	9
29	28	conferencePaper	contribution	2021	Gulati, Rishi; Handa,	Towards Develop	2021 Third Inte	Children suffering from	2	2	2	0	1	7
30	29	journalArticle	contribution	2014	Escobedo, Lizbeth; Te	Using Augmented	IEEE Pervasive	Children with autism ha	2	2	2	2	2	10
31	30	conferencePaper	contribution	2022	Romadlon Junaidi, Ah	Usability Testing	2022 2nd Inter	The development of con	2	2	2	0	0	6
32	31	conferencePaper	contribution	2019	Machado, Eduardo; C	An Assistive Augn	2019 IEEE Intl	Acquiring daily living skil	2	2	2	2	2	10

Figure 3. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. QA: quality assessment.



Synthesis Methods

The data extraction process was carried out to methodically address the stated RQs. Initially, a single reviewer handled the task of data extraction, leveraging the analytical capabilities of Microsoft Excel to facilitate a structured and organized approach to data collection and analysis. The same tool was used for data organization and representation. Within this app, a comprehensive table was constructed in which the rows were designated to the considered references and data corresponding to individual RQs were entered into separate columns, fostering a systematic representation of the data obtained. Subsequently, to enhance the reliability and validity of the data integration process, a second reviewer performed a verification of the initially extracted data. This encompassing procedure ensured a high degree of accuracy and reduced potential discrepancies, thus guaranteeing the integrity of the data.

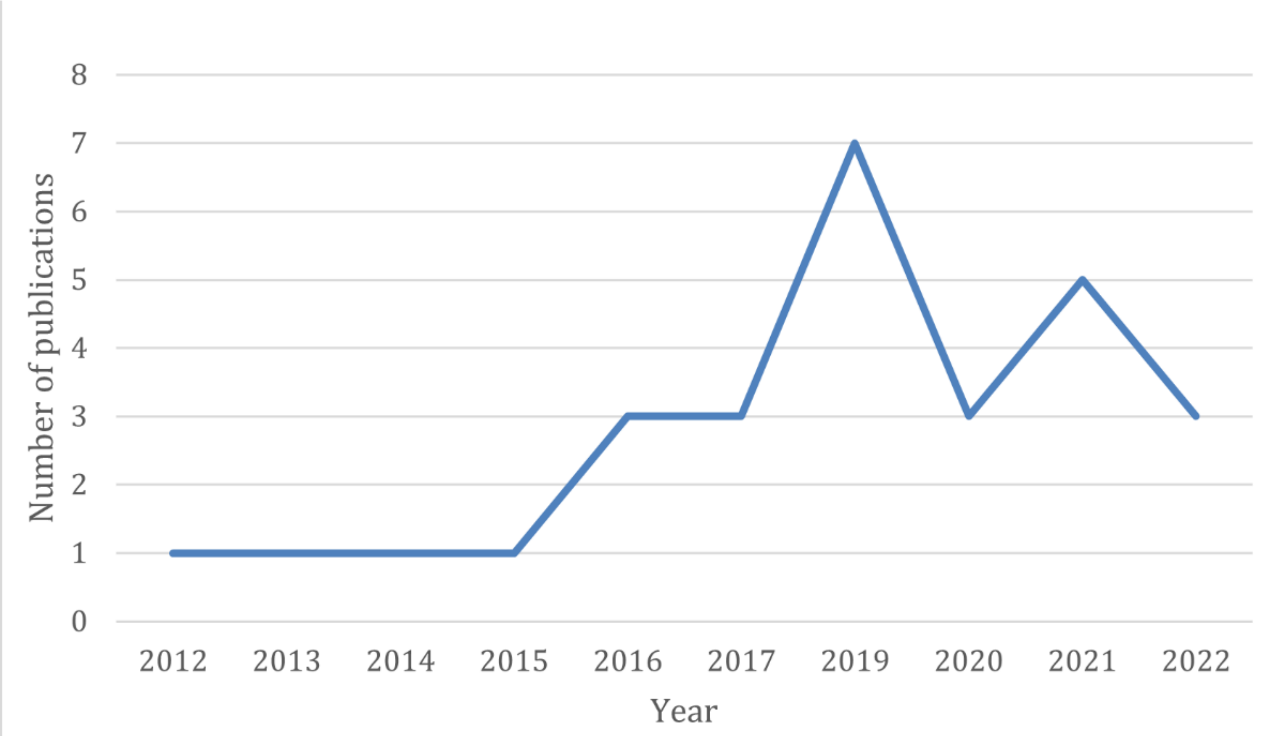
Results

As a result of applying the PRISMA guidelines, a total of 28 publications were considered in this study to address the RQs.

RQ 1: What Is the Papers' Distribution Over Time?

Considering the time range for paper analysis, Figure 4 depicts the time distribution of publications over the period of 2012 to 2022.

Figure 4. Time distribution of the reviewed publications.



Analyzing Figure 4, we can see an increasing trend in the number of publications starting in 2015, with the maximum value being reached in 2019. This denotes an increasingly high interest in these technologies. The number of publications decreased again in 2020, probably because of the pandemic, which limited human interaction and prevented the development and testing of apps dedicated to children with autism.

Regarding the types of publications, more than half (17/28, 61%) were presented at conferences, and 39% (11/28) were articles published in specialized scientific journals.

RQ 2: What Category of People Are the Studies Aimed at?

The analysis of the considered publications revealed that 43% (12/28) [29,31-33,36,39,43,44,47,50,51,53] stated that they were about children without mentioning the number of participants or their ages. Textbox 1 summarizes the data obtained.

The paper by Xia et al [47] addressed people with autism without mentioning whether they were children or adults, and the study by Wang et al [54] only involved adults but was of interest because the proposed solution can be applied to children as well.

In addition, the studies by Zheng et al [38], Escobedo et al [46], and Voss et al [49] included children both with and without ASD to compare the results and the process recorded in both cases.

Textbox 1. Number and age of the children involved in the studies.

- Hashim et al [28]: 6 children aged between 5 and 12 y
- Machado et al [29]: children—age not stated
- Tang et al [30]: children aged <4 y and between 4 and 8 y; number not mentioned
- Selvarani et al [31]: children—age not stated
- Abou El-Seoud et al [32]: children—age not stated
- Vullamparthi et al [33]: children—age not stated
- Singh et al [34]: children aged between 9 and 12 y
- Chen et al [35]: 6 teenagers aged between 11 and 13 y
- Tang et al [36]: children—age not mentioned
- Giraud et al [37]: 12 children aged between 5 and 9 y
- Zheng et al [38]: 12 children, 6 with autism spectrum disorder (ASD) and 6 with typical development
- Pradibta and Wijaya [39]: children—age not stated
- Nubia et al [40]: 6 children (5 boys and 1 girl) aged between 3 and 9 y
- Sait et al [41]: 9 children aged between 4 and 12 y
- Wan et al [42]: 10 children aged between 3 and 8 y
- Kavitha et al [43]: children—age not stated
- Silva et al [44]: children—age not stated
- Kalantarian et al [45]: 8 children aged between 6 and 12 y
- Escobedo et al [46]: unknown number of children aged between 8 and 11 y, including 3 children with autism
- Xia et al [47]: mainly people with autism
- Amado et al [48]: children aged between 7 and 9 y; number not indicated
- Voss et al [49]: 20 children with ASD and 20 children without ASD
- Washington et al [50]: 14 families
- Gulati and Handa [51]: children—age not stated
- Escobedo et al [52]: 12 children and 7 teachers
- Bouaziz et al [53]: children—age not stated
- Wang et al [54]: 4 adults, but the system was suitable for children as well
- Gelsomini et al [55]: 5 children (2 with mild ASD, 2 with medium ASD, and 1 with psychomotor retardation)

RQ 3: What Technologies Are Used With XR or Any of Its Subdivisions to Develop Apps for ASD Therapy?

AR mobile apps for therapy for children with ASD typically used a combination of the following technologies:

1. Mobile devices such as smartphones and tablets equipped with cameras; displays; and sensors such as accelerometers, gyroscopes, and GPS
2. AR software development kits such as ARKit, ARCore, and Vuforia, which provide the tools and framework needed for developing AR apps
3. Graphical and game engines such as Unity and Unreal Engine for 3D model development and creating animations and interactive environments
4. Natural language processing and speech recognition technologies for creating voice-activated AR experiences
5. Computer vision and image-processing techniques for real-time object tracking and recognition of objects, faces, and gestures
6. Machine learning algorithms for customizing the AR experience based on the child's performance and preferences
7. Cloud computing infrastructure for data storage, management, and analysis of therapy progress

Of the 28 analyzed publications, 22 (79%) addressed a solution from the spectrum of AR implemented on mobile devices such as smartphones owing to their processing power and integrated sensors that make them suitable tools for implementing apps without the need for additional and sophisticated equipment. In addition, the papers by Giraud et al [37], Sait et al [41], Gulati and Handa [51], and Gelsomini et al [55] presented solutions based on VR. Although the articles by Wan et al [42] and Kalantarian et al [45] did not present an AR or VR solution, the methodology addressed and the results obtained show the potential for research in this area. Textbox 2 summarizes information about the technologies and platforms used for app development.

Unity 3D and Vuforia were among the most common platforms used in the development of AR apps for mobile devices, with the Android operating system often mentioned. Several studies (6/28, 21%) [29,38,41,49,51,55] used wearable devices such as Google Glass, Oculus Go, Google Cardboard, Leap motion sensors, and E4 wearable sensors along with the mentioned technologies. Interactive cards were also used as markers to overlay virtual content.

Textbox 2. Technologies and platforms used.

- Hashim et al [28]: interactive cards, augmented reality, and smartphones
- Machado et al [29]: augmented reality based on smart glasses and Android, web platform, Node.js, eye tracker, sensors, and Amazon Alexa
- Tang et al [30]: augmented reality and Google TensorFlow
- Selvarani et al [31]: interactive cards, augmented reality based on markers, Vuforia, Android smartphone, and Unity 3D
- Abou El-Seoud et al [32]: augmented reality based on markers, smartphones, and the Aurasma framework
- Vullamparthi et al [33]: smartphone, Android, augmented reality, and QR codes
- Singh et al [34]: desktop app and augmented reality
- Chen et al [35]: Vuforia and smartphone or tablet PC
- Tang et al [36]: Google TensorFlow, augmented reality, and smartphone or PC
- Giraud et al [37]: virtual reality (VR) and Unity 3D
- Zheng et al [38]: augmented reality, Microsoft Kinect, and portable E4 sensor
- Pradibta and Wijaya [39]: interactive cards, augmented reality, Android smartphone, and Adobe for animation and graphic illustration
- Nubia et al [40]: augmented reality, Android tablet PC, Unity 3D, Vuforia, and Blender
- Sait et al [41]: VR, Unity 3D, and VR glasses (Oculus Go)
- Wan et al [42]: system that can be implemented on a PC, smartphones or robots; no use of augmented reality or VR
- Kavitha et al [43]: augmented reality, Android smartphone, Vuforia, and ARCore
- Silva et al [44]: augmented reality, smartphone or tablet PC, and Vuforia
- Kalantarian et al [45]: Android smartphone; no VR or augmented reality
- Escobedo et al [46]: augmented reality and Android smartphone
- Xia et al [47]: augmented reality, Android or iOS smartphone, React, Node.js, and Python for object recognition
- Amado et al [48]: augmented reality, Vuforia, Unity 3D, Android smartphone, Balsamiq Mockups 3, and Tinkercad
- Voss et al [49]: augmented reality, Android smartphone, and Google Glass
- Washington et al [50]: Google Glass and Android smartphone
- Gulati and Handa [51]: VR, Leap motion sensors, and VR camera
- Escobedo et al [52]: augmented reality, smartphone or tablet PC, PC server, MySQL database, and HTTP
- Bouaziz et al [53]: interactive cards, augmented reality, smartphone, and Vuforia
- Wang et al [54]: augmented reality, tablet PC or smartphone, Unity 3D, and Vuforia
- Gelsomini et al [55]: VR, Google Cardboard, smartphone, and Unity 3D

RQ 4: What Skills Were Targeted for Improvement in Primary Studies?

Owing to the deficiencies of children with ASD, the aim was to improve some basic skills such as the following:

1. Communication and language development
2. Social interaction and play skills
3. Fine and gross motor skills
4. Emotional regulation and collaborative strategies
5. Cognitive and problem-solving abilities
6. Attention and ability to follow instructions
7. Independence and self-help capabilities

Table 3 presents the number of publications aimed at improving basic skills.

A total of 25% (7/28) of the publications [28,30,33,36,40,49,52] focused on improving communication skills such as English vocabulary learning [28]; word learning using automatic object recognition through an app based on the TensorFlow library that can be used either when connected to the internet or offline [29]; speaking, reading, and associating images using an app that allows for customization of lessons by parents or therapists [33]; and communication and socialization by delivering certain cues through smart glasses [49]. Textbox 3 details the skills targeted in the studies.

Table 3. Targeted learning skills (n=28).

Skill	Studies, n (%)
Religious skills	1 (3)
Daily activities, meal preparation, toothbrushing, and eating	3 (11)
Cognitive or attention	2 (7)
Expressing emotions or social skills	5 (18)
Environment adaptation	1 (3)
Motor skills	2 (7)
Task training	1 (3)
General skills	3 (11)
Number learning	1 (3)
Object recognition	2 (7)
Communication or vocabulary	7 (25)

Textbox 3. Skills aimed to be improved.

- Hashim et al [28]: communication skills; learning English vocabulary, pronunciation, and articulation skills
- Machado et al [29]: daily routine activities (preparing meals)
- Tang et al [30]: word learning and object recognition
- Selvarani et al [31]: number learning
- Abou El-Seoud et al [32]: general skills; the user can choose the augmented reality (AR) content to be displayed
- Vullamparthi et al [33]: speaking abilities, reading, image associations, and activity scheduling
- Singh et al [34]: procedural task fulfillment
- Chen et al [35]: expressing emotions and social abilities
- Tang et al [36]: object recognition and vocabulary learning skills
- Giraud et al [37]: motor and social skills
- Zheng et al [38]: toothbrushing abilities
- Pradibta and Wijaya [39]: religious abilities—prayers
- Nubia et al [40]: communication abilities
- Sait et al [41]: adaptation to a new or unfamiliar environment
- Wan et al [42]: cognitive skills and practicing facial emotions
- Kavitha et al [43]: general skills; the user can choose the AR content to be displayed
- Silva et al [44]: social and general skills
- Kalantarian et al [45]: expressing emotions and social abilities
- Escobedo et al [46]: social skills in real-life situations, building and maintaining social relationships, improving conversational ability, and managing behavior and emotions
- Xia et al [47]: social and self-help abilities (shopping)
- Amado et al [48]: cognitive skills
- Voss et al [49]: social and communication abilities
- Washington et al [50]: expressing emotions
- Gulati and Handa [51]: motor, focusing, and general skills
- Escobedo et al [52]: object recognition
- Bouaziz et al [53]: self-help skills (feeding)
- Wang et al [54]: attention skills
- Gelsomini et al [55]: general skills (attention, concentration, and understanding) and narration

Some studies (3/28, 11%) focused on the development of self-help skills such as preparing meals with the help of smart glasses, receiving real-time information about the steps to follow [29], brushing teeth [38], or eating [53]. By improving the skills aimed at in the studies considered in this review and developing skills that can improve the deficiencies present in children with ASD, the social inclusion of children with ASD was pursued.

RQ 5: What Are the Purposes for Which the Proposed Solutions Were Used?

Using at least one of the technologies targeted in this review, the solutions presented in these publications were used to assist children with ASD. Developed for use by both therapists and parents at home or in specialized medical centers, these solutions aimed to improve certain fundamental aspects of the lives of children with autism. Textbox 4 presents information related to the reasons for which the apps were developed.

Hashim et al [28] created an app for the development of children's English vocabulary, which could potentially be used with other languages as well. The solution proposed in the paper by Machado et al [29] used multiple technologies to allow the therapist to model activities using a web platform and provide hints to users via smart glasses. It also works as an attention-monitoring tool via an eye tracker so that activities can be evaluated and improved. The study by Tang et al [30] addressed the problem of communication through automatic object recognition using a smartphone and display of virtual content (object names) on the screen. This app works either when connected to the internet or offline. A similar approach was observed in the study by Selvarani et al [31], in which children could learn numbers by scanning notebooks using their mobile devices, after which the relevant content was displayed on the screen.

Textbox 4. The purpose of the developed apps.

- Hashim et al [28]: the Areal-Vocab app was developed to help children with autism improve their English vocabulary.
- Machado et al [29]: the aim was to develop an assistive app using augmented reality (AR) based on smart glasses and a visual attention analysis tool to help people with autism in daily tasks by providing complementary information (eg, to pick up the knife and then cut the strawberries). The therapist can model the activity.
- Tang et al [30]: the researchers intended to develop a mobile app for children with autism that could run either connected to the internet or offline that would improve word learning skills by using object recognition.
- Selvarani et al [31]: the aim was to help children with autism learn numbers by scanning an interactive card using the app so that complementary video and audio content is displayed on the screen.
- Abou El-Seoud et al [32]: the aim was to develop a framework to help parents or educators use AR in a personalized way by choosing what type of AR educational content to display over a printed marker representing a familiar cartoon character.
- Vullamparthi et al [33]: a tool was developed that included an interface for parents or educators to scan a QR code and create various lessons and an interface for children. It used the smartphone camera, an Android apk (Android application package), a web page, a database, and Jakarta server pages.
- Singh et al [34]: the paper was a comparative study that aimed to explore the effectiveness of AR in the execution of tasks among less privileged children (who have had minimal interaction with technology), healthy but younger children, and children with autism.
- Chen et al [35]: the researchers developed a Vuforia-based AR app that can be deployed on Android or iOS devices (smartphones or tablets). This app can scan storybooks (with images captured from videos) and overlay relevant content to assist children with autism in expressing and understanding emotions and developing social skills.
- Tang et al [36]: the aim was to develop a tool for children with autism that can recognize objects and display their names.
- Giraud et al [37]: the study aimed to involve children with autism in common actions (moving furniture) by interacting with a virtual character projected on a tactile magnetized surface.
- Zheng et al [38]: the goal was to develop an AR system (Cheerbrush) that could teach children with autism how to brush their teeth considering how important this is to stay healthy and avoid dental procedures. It uses Kinect to capture the user’s movement, a 3D-printed toothbrush to assess brushing skills, a monitor to view the surroundings, and an avatar. It also uses a wristband to assess children’s stress while using the app.
- Pradibta and Wijaya [39]: the aim was to help children with autism learn daily prayers. The goal was to develop an app that contains animated learning materials in the form of daily prayers from the Islamic religion.
- Nubia et al [40]: the aim was to help children with autism communicate better using an app that can identify human-recognizable objects such as animals, fruits, or other common objects and match them with specific sounds.
- Sait et al [41]: the goal was to develop a virtual reality (VR) framework in which the teacher can enter information about the child and prepare scenes that can be watched by a child wearing VR glasses. The main objective was to familiarize children with autism with places such as school, the schoolyard, and the classroom by previously visualizing the environment.
- Wan et al [42]: the aim was to help children with autism recognize, practice, and express emotions such as happiness, sadness, fear, or anger.
- Kavitha et al [43]: the aim was to help children with autism recognize objects or animals by rendering 3D content over certain images.
- Silva et al [44]: the aim was to help reduce the isolation of children with autism by encouraging them to explore the world with the help of an app based on geolocation and AR.
- Kalantarian et al [45]: the goal was to help children with autism learn to express their emotions. Guess what? is an Android mobile app similar to Heads up!, a game in which a parent holds the smartphone with the screen facing the child, the child imitates what they see, and the parent tries to guess the simulated emotion.
- Escobedo et al [46]: the paper describes the design and development of the MOSOCO app, which is a mobile app that provides real-time support and guidance to children with autism in practicing social skills. The app uses AR technology to overlay social hints directly into the child’s real environment, allowing them to practice social skills in real-life situations.
- Xia et al [47]: the app provided step-by-step guidance for people with autism to go shopping by augmenting real shopping scenes using object recognition, barcode reading, and automatic classification.
- Amado et al [48]: the main objective was to develop an AR mobile app to be used by parents of children with autism for their therapy during the pandemic, when human interaction was limited.
- Voss et al [49]: the system described in the paper aimed to help people with autism spectrum disorder improve their social skills by providing discrete real-time social cues via wearable technology. Social cues are provided directly in the wearer’s field of vision using AR technology and are intended to help the wearer navigate social situations and improve their social interactions and communication skills.
- Washington et al [50]: the goal was to develop an app that runs on an Android smartphone (used by a parent) that is connected to a Google Glass device worn by the child. Social cues are delivered to the glasses based on emotions recognized by the mobile app, which also records the session (video only for privacy reasons). The activities are gamelike—catch the smile, guess the emotions, and unstructured activities.
- Gulati and Handa [51]: the aim was to develop a VR game to improve reading, basic math, and spelling. Motor skills are improved by reading gestures and helping coordinate them with the eyes using the Leap motion sensor.

- Escobedo et al [52]: the aim was to create an app that can identify objects that are tagged and display relevant content over them, such as text, 3D models, vibrations, video, or audio, and the user can receive a reward. The main architecture is composed of a module called therapy manager, an ambient notification system, and a tag manager.
- Bouaziz et al [53]: the aim was to develop an app dedicated to children with autism that teaches them how to eat by scanning an interactive card and displaying on top of it a 3D character depicting the targeted skill.
- Wang et al [54]: the aim was to help adults with autism be more focused by performing certain tasks, such as rearranging objects in a scene.
- Gelsomini et al [55]: the aim was to develop a VR mobile app for smartphones that can be used with Google Cardboard, helping children with autism understand activities through storytelling and allowing caregivers to customize the content using a web app, monitor children's attention, and analyze statistics.

The apps developed in the studies by Abou El-Seoud et al [32] and Vullamparthi et al [33] aimed to go through some lessons that parents and educators could customize by accessing a web platform so that they could choose which type of content to display when the app detected an object in the visual area. The study carried out by Singh et al [34] compared the effects of apps that use AR to perform certain tasks in both children with ASD and children with typical development. Religious activities were also included in one study [39], which presented an app containing animated materials that helped children learn prayers.

Given the fact that children with autism typically experience difficulties adapting to a new environment, Sait et al [41] aimed to develop an app that uses VR, VR glasses, and a web platform in which therapists can enter information about each child and set up custom scenes, such as a classroom, to be viewed virtually and get used to. In addition to the goal of conducting basic activities, Xia et al [47] developed an app to guide individuals with autism with grocery shopping step by step.

RQ 6: What Are the Results Obtained Using the Proposed Solutions?

Depending on the proposed solution and the objectives of the studies, the results were different, but in general, where the app was tested, encouraging results were obtained, with the remark that these were to be improved and tested more thoroughly. In cases in which the app was not tested with the intended audience but was proven to function, it was deemed to have potential. Textbox 5 summarizes the results obtained in each study.

Upon analyzing the results of the studies included in this review, it was found that only 68% (19/28) of the apps were tested with children with autism, whereas 32% (9/28) were tested only by the developers for functionality purposes. Regarding the methods used to quantify the results, 54% (15/28) of the studies used interviews, and only 14% (4/28) of the studies used an assessment method based on assigning a score according to the degree of skill improvement after using the apps.

Textbox 5. Summary of the results.

- Hashim et al [28]: children and their parents or educators in the study used the app and reported positive results based on interviews: "Helps listen and understand instructions, helps maintain attention longer, helps with pronunciation and enunciation, helps keep them engaged and interested to learn the vocabulary in depth."
- Machado et al [29]: the app has great potential considering the fact that smart glasses can very easily transpose the user into the world of augmented reality (AR) and help them by displaying complementary information, as well as giving feedback to the therapist. It has been tested by developers but has not been tested with children with autism, so it does not show quantifiable results.
- Tang et al [30]: the first pilot study was conducted on a university campus with neurotypical children and adults, who provided positive feedback and showed a lot of interest. The second study was conducted in a special education unit involving 2 groups: one with children aged <5 y and one with children aged between 6 and 8 y. It was noticed that the younger children had difficulty using the app, but it was well received by the older children. Positive feedback was also provided by parents and teachers, pointing out that the offline module required improvement.
- Selvarani et al [31]: the app (NUM09) is functional but has not been tested on children with autism with quantifiable results.
- Abou El-Seoud et al [32]: a total of 3 patients with autism, together with their instructors, performed a usability test. According to responses to a questionnaire, the system can improve communication, concentration, and attention and is easy to use.
- Vullamparthi et al [33]: this study developed an Android smartphone app that helps children with autism and their parents or therapists create personalized lessons to improve basic skills such as reading, writing, or picture recognition. A workshop was held, and positive feedback from parents was reported. There are no quantifiable results.
- Singh et al [34]: the main task was to complete a tangram puzzle. In the first stage, the involved children did not solve the puzzle without clues involving AR, but it was reported that solving took longer in the AR training mode. In the first study, children aged 9 to 12 y rated the desktop-based instruction mode as the least preferable, whereas the performance using the AR mode was superior. In the second study, 4 children with autism followed the same procedure but had difficulty using the AR-based solution, resulting in poorer performance on the task.
- Chen et al [35]: the app was tested in a dedicated room equipped with a computer, a 52-inch monitor, and 8 tablets. The therapist showed the children the app and asked them to look at the pictures, answer some questions, and use the tablet to access the AR content by pointing it at the picture with the app running in the background. Positive feedback was reported from the children, who were curious and eager to discover new visual cues, showing interest in the facial expressions, gestures, and related activities of the characters. The children had low scores on the initial assessment, but all 6 scores increased significantly after the app intervention. The most dramatic improvement was in one child, from 30% to 89.5%.
- Tang et al [36]: the app works, but it has not been tested on children with autism with quantifiable results.
- Giraud et al [37]: 12 children with autism spectrum disorder (ASD; including 2 girls) aged between 5 and 9 y participated in the study for a period of 3 mo. In total, 7 of the children showed little conversational language. A 3-stage experiment was conducted (familiarization, moving an object with an agent following the child, and training with an agent that the child follows). Preliminary results were encouraging: one-third of the children completed the training, another third needed device adjustments, and some had difficulty using the system.
- Zheng et al [38]: to evaluate the system, 6 children aged between 3 and 6 y (3 with ASD and 3 without ASD) were involved in an experiment comparing the results. It was noted that all the children were able to complete the training sessions, but the children with ASD were clearly more engaged and interested. After training, the most notable improvements were observed in children with autism. During an interview, both children and parents said that they liked the app and that it helped them improve their toothbrushing skills.
- Pradibta and Wijaya [39]: no proof of testing with children with autism and no quantifiable results.
- Nubia et al [40]: by playing relevant sounds in line with images, the app helped children improve their learning skills compared with traditional methods. A 14% increase in attention and a 9% increase in verbal language were reported.
- Sait et al [41]: the system was used by 9 children with autism who benefited from the help of therapists who guided them in adjusting the Oculus Go headset and using the app (AutiVE). One of the issues was the virtual reality (VR) headset itself and the VR environment, but the website provided had a video explaining them. In total, 8 of the children eventually accepted the device. There were some improvements in learning skills, but no detailed statistics were mentioned.
- Wan et al [42]: the children completed a 20-min training session each day for 4 consecutive days. A total of 6 participants showed improvement in proficiency in operating the system, 5 of 6 completed all tasks, and 4 of 6 showed improvements in expressing emotions. Children aged <5 y found the app difficult and did not perform in a satisfactory manner.
- Kavitha et al [43]: the app works, but it has not been tested with quantifiable results.
- Silva et al [44]: an app similar to Pokémon GO was developed in which users can find "monsters" in certain areas and, by clicking on them, find relevant information. The concept of gamification was used, but the system was not validated with real users with autism.
- Kalantarian et al [45]: the solution was tested with 8 children, all boys, playing up to 5 games in 1 session. In total, 94%, 81%, 92%, and 56% of the emotions were labeled correctly as disgust, neutrality, surprise, and fear, respectively.
- Escobedo et al [46]: the app was evaluated over 7 wk. Interview results revealed that the app was well received by children with autism and their therapists and that it was effective in helping children practice and improve their social skills in real-world situations. The authors reported that users were able to use the app easily and that the AR technology was effective at providing children with real-time support and feedback. The study also showed that the app was well accepted by therapists, who found it a useful tool for their patients' therapy.
- Xia et al [47]: the app, called ParaShop, was tested by a nonprofit organization that helps people with disabilities. Staff said that the app helped people with autism buy their groceries, but the number of participants or other details were not mentioned.

- Amado et al [48]: a case study was conducted using Google Forms asking parents to answer questions related to their children (eg, age, gender, and whether the parents lived together). Several studies with parents were conducted, and then the app was developed based on their responses and requirements. In the last stage of the case study, 5 questions were posed about the final prototype of the app. The survey revealed that 46.2% of parents were satisfied and 23.1% were very satisfied. Overall, the mobile app received positive feedback from respondents.
- Voss et al [49]: the research entailed a study involving 20 participants with ASD and 20 participants without ASD who used a system called Superpower Glass over a 4-mo period. The results showed that the participants found the social cues useful in situations and improved their social interactions and communication skills. The study also assessed the acceptability and usability of the system, and the results suggest that it was well received by participants and easy to use.
- Washington et al [50]: the app was tested by families, and they reported that it was useful, with some of them recording the sessions and then showing them to the children to see how they behaved for further improvement. Overall, based on interviews, parents reported positive outcomes.
- Gulati and Handa [51]: the concept of gamification was used; it has potential, but it has not been tested in children with autism. To play the game, a dedicated gaming room and specific equipment are required.
- Escobedo et al [52]: the app (Mobis) was tested with 7 teachers caring for 12 children with autism aged between 3 and 8 y. The researchers conducted weekly interviews with the teachers, keeping in mind that only 3 out of 12 children were able to properly pronounce words. The duration of the observation was 54 h. Participants were reported to find Mobis "exciting, useful, and easy to use." Students improved their motor skills by focusing the camera on the target. Mobis increased the time that students stayed on task by 20% and motivated them to use the app as they were excited to discover new objects in their environment. Selective attention improved by 62%, and sustained attention improved by 45%. Mobis also induced positive emotions and taught behavioral skills such as tolerance.
- Bouaziz et al [53]: no proof of testing with children with autism and no quantifiable results.
- Wang et al [54]: the app was developed for demonstrative purposes only; it has not been tested with quantifiable results.
- Gelsomini et al [55]: the solution (Wildcard) was tested in a special unit with 5 children with autism during 8 individual therapy sessions. Therapists reported improvements in children's attention and cognitive skills, but the paper only reported qualitative data. Therapists were excited to be able to customize each VR session and noted that patients embraced the app and found it engaging.

Discussion

Principal Findings

The analysis revealed an increasing trend in publications starting from 2015, reaching its highest point in 2019 and followed by a decline in 2020, potentially because of the pandemic. Most of the papers (17/28, 61%) were presented at conferences and largely focused on AR solutions (22/28, 79%) for mobile devices to assist children with ASD in enhancing basic skills and fundamental life aspects. Notably, Unity 3D and Vuforia emerged as popular development platforms. Although a substantial percentage of publications (13/28, 47%) did not provide details on participating children, most of the identified participants were aged between 3 and 13 years. Developed for use by both therapists and parents at home or in specialized medical centers, these solutions showed encouraging preliminary results but underscore the necessity for further, more extensive testing, particularly as a significant portion (9/28, 32%) were only developer tested.

Main Directions of Research

Upon examining the scientific publications included in our study, several main directions for the use of XR to support children with autism can be identified.

One notable topic is the use of AR in the area of language skills and vocabulary learning in children with autism. Researchers in some studies (3/28, 11%) [28,30,43] focused on the development of AR-based mobile apps that facilitate word learning and object recognition through techniques such as deep learning and automatic object recognition.

Another topic addressed in some studies (3/28, 11%) [29,39,53] was the use of smart glasses or wearable devices to support children with autism in social interactions. These devices provide real-time visual cues and information to enhance communication and social interaction skills.

The use of AR occupational therapy and the development of cognitive skills in children with autism were explored in some studies (3/28, 11%) [31,38,46], which proposed AR-based apps to aid children in learning numbers, teeth-brushing skills, or environmental adaptation skills.

Furthermore, it was stated that the apps specifically designed for children with ASD should be tested with a target group of children, and the results should be quantified in a pertinent manner given that a large part of the findings were obtained through interviews.

Personalization and adaptability are other key aspects of developing mobile apps, as shown in the studies by Wan et al [42], Kalantarian et al [45], and Washington et al [50]. These publications addressed the development of personalized systems and apps to maximize the therapeutic and educational benefits for children with autism.

Some studies (3/28, 11%) [32,34,52] also examined the use of AR to provide individual support for people with autism and cognitive impairment. These studies proposed AR-based frameworks and approaches to assist individuals with autism in various activities and tasks, such as training in procedural tasks, perception and recognition of facial emotions, or assistance in real-life situations.

In addition, some studies (3/28, 11%) [35,37,55] investigated the use of AR in the context of education and social skill development. These studies focused on the use of interactive books, serious games, or training apps to support children with autism in understanding and interpreting facial expressions, social cues, and social interactions.

The use of AR in the context of learning in a geographical environment or learning environmental coping skills was addressed in some studies (3/28, 11%) [40,41,47]. These studies proposed AR-based apps to assist children with autism in exploring and learning in a geographical environment or in developing adaptive skills applicable to different situations and contexts.

A relatively small number of studies (6/28, 21%) [34,38,46-48,52] focused on VR-based approaches for apps, indicating a shift toward the adoption of AR owing to its lower cost and greater usability. Using VR, the study by Abou El-Seoud et al [32] evaluated joint action (moving furniture) abilities using a virtual character, and the preliminary results were encouraging, with one-third of the children completing the training despite 7 of them having limited conversational language. Another issue addressed in the study by Tang et al [36] was adaptation to unfamiliar surroundings. Researchers reported improvements in learning abilities but stated that the equipment and VR environment posed the greatest challenges. The concept of gamification was integrated with VR in the study by Escobedo et al [46] to enhance reading, basic mathematics, and spelling. This app required a special room to run. Although this app has great potential, it has not yet been tested in children with ASD. In addition, an app using Google Cardboard and VR was developed to enhance the cognitive and attentional skills of children. Therapists expressed satisfaction with the outcomes as they were able to personalize each session using unique teaching methods.

User Interaction Perspectives

Overview

The interaction of children with ASD with XR devices, such as AR and VR platforms, brings forth a distinct set of considerations. The manner in which children with ASD use these devices can be influenced by their sensory sensitivities, motor skills, cognitive abilities, and preferences. Although the experiences can vary widely, the following are some ways in which children use XR devices and the challenges they may face.

Physical Interaction

Children use XR devices by interacting with touch screens, controllers, or wearable components. They may tap, swipe, or perform gestures to navigate through XR environments. However, children with fine motor difficulties may struggle with precise interactions, leading to accidental inputs or difficulties in selecting desired options.

Visual Engagement

Children engage visually with the XR content displayed on screens or through headsets. Visual stimuli can capture their attention and spark interest. Nonetheless, those with sensory sensitivities may experience sensory overload or visual discomfort if the content is excessively bright, flashy, or overwhelming.

Spatial Awareness

XR experiences often involve spatial interactions such as moving through virtual environments or manipulating virtual objects. Children's spatial awareness skills can influence their ability to navigate these environments. Some children may find it challenging to grasp the concept of a virtual space, leading to disorientation.

Auditory Response

Many XR apps incorporate auditory cues, sound effects, or voice instructions. Children may respond to auditory prompts by vocalizing or reacting physically. However, children who are sensitive to loud or sudden sounds may experience distress in XR environments with intense auditory stimuli.

Attention and Engagement

Children's level of attention and engagement with XR content can vary. Some may become deeply immersed and engaged, whereas others may have difficulty sustaining their attention because of the novelty of the experience or sensory distractions.

Preferences and Comfort

Children's preferences for certain types of interactions or content can influence their engagement. Some children may appreciate exploring virtual worlds, whereas others may prefer more structured or repetitive activities. Ensuring a variety of XR experiences allows for accommodating different preferences.

Transition Challenges

Transitioning between the real world and the XR environment can be challenging for some children. They may have trouble understanding that the virtual elements are not physically present or struggle with transitioning back to reality after prolonged XR use.

Response Variability

Children with ASD may respond to XR experiences differently across sessions. Factors such as mood, sensory sensitivities, and cognitive states can influence their interactions. Some days, children may be more receptive to XR, whereas on other days, they may be less engaged or overwhelmed.

Calibration and Setup

XR devices require proper calibration and setup for optimal interaction. Children may need assistance in adjusting headsets, ensuring proper alignment, or calibrating controllers. Technical difficulties can lead to frustration or disengagement.

Challenges

The identified challenges are as follows:

1. Individualized learning needs: a prevalent challenge across the studies in this review was catering to the diverse learning preferences and abilities of children with ASD. For instance, Hashim et al [28] faced the task of addressing the specific needs of children with mild ASD. Similarly, the studies by Abou El-Seoud et al [32] and Singh et al [34] addressed the challenge of tailoring their AR experiences to suit varying preferences and capabilities.

2. Transferability and generalization: a common limitation is the transfer of learned skills to real-world scenarios. As seen in the studies by Tang et al [36] and Giraud et al [37], researchers have encountered challenges in translating acquired skills into practical applications. In addition, studies such as those by Chen et al [35] and Kavitha et al [43] noted limitations in transferring learned skills beyond the AR context, possibly owing to the variations in real-world stimuli.
3. Technical feasibility and personalized support: technical feasibility and ongoing support emerged as challenges in some studies [29,38]. Maintaining the functionality of AR-based smart glasses and ensuring accurate real-time feedback for toothbrushing techniques required continuous technical support.
4. Sensory overload and individualization: sensory sensitivities and the need for individualized solutions were prominent challenges. Sait et al [41] encountered the challenge of designing virtual environments that cater to sensory sensitivities, whereas studies such as the one by Voss et al [49] highlighted the importance of unobtrusive cue presentation in wearables for children with ASD.
5. Cognitive adaptation and user adoption: cognitive adaptation and user adoption challenges were evident in some studies [48,52]. Designing tasks that effectively target cognitive skills and maintaining user engagement over time were key considerations.
6. Ethical implications of data handling: as AR interventions involve interactions and data collection, ethical considerations are paramount. The study by Wan et al [42] delved into recognizing facial expressions, which raises ethical concerns related to data privacy and security. Ensuring that data-handling protocols adhere to ethical standards becomes crucial, underscoring the need to protect sensitive user information while deriving meaningful insights from the interactions.
7. Cross-cultural adaptation and applicability: given the diversity of cultures and languages, ensuring the cross-cultural adaptation and applicability of AR interventions becomes a notable challenge. The study by Wang et al [54], which explored mobile AR for attention improvement in adults with ASD, highlights the importance of adapting interventions to diverse cultural contexts. This challenge emphasizes the need for cultural sensitivity and the localization of content to ensure that interventions are universally accessible and effective.
8. Long-term impact measurement: measuring the long-term impact of AR interventions and tracking the progress of children over time poses significant challenges, as pointed out in the study by Escobedo et al [46], which emphasized the importance of assessing the sustained effects of interventions beyond short-term interactions. This challenge underscores the necessity of devising reliable methodologies for gauging the lasting benefits of AR interventions and understanding how these interventions contribute to the developmental trajectory of children with ASD.

In the realm of AR apps for children with ASD, studies have striven to engage children through various interaction modes while tackling shared challenges. The diverse engagement strategies and the collective endeavor to overcome common limitations underscore the continuous efforts to create meaningful and effective AR-based interventions for this unique demographic.

Research in this area demonstrates an interdisciplinary approach involving collaboration among specialists in education, IT, and mental health. This is illustrated by the diversity of authors and publications included in this review, suggesting that integrating AR, VR, and MR into ASD pedagogy requires a comprehensive approach that considers multiple aspects—from technology design to educational and mental health psychology. It was also specified that the apps aimed at children with ASD should be tested with a target group of children and that the results should be quantified in a more relevant manner given that a large part of the reported results was obtained only through interviews. The analysis of the studies indicates a trend in research toward the use of diverse and innovative study methods, such as using both quantitative and qualitative methods to investigate the impact of AR, VR, and MR on people with ASD.

Furthermore, the analyzed publications suggest that the development and implementation of AR-, VR-, and MR-based technologies extend beyond academia or research, involving partnerships with the private sector and local communities. This demonstrates the awareness of the need to transfer research findings into practice to have a direct impact on people with autism.

Limitations

Considering the publications reviewed in this study, several limitations were identified regarding the development and testing of AR-, VR-, and MR-based mobile apps:

1. Sample size: some studies involved small samples of participants, which may have limited the generalizability of their results to a larger population of children with autism. The involvement of a limited number of participants in many studies can be attributed to the unique characteristics of the target population—children with ASD. The diversity in ASD manifestation, severity, and individualized needs necessitates careful participant selection. Moreover, recruitment challenges, ethical considerations, and the resource-intensive nature of working with children with ASD contribute to the small sample sizes. However, this limitation was often acknowledged in the papers, along with the understanding that the findings may not be easily generalizable to the broader population with ASD.
2. Study duration: the duration of the studies included in this review varied from short testing sessions to several weeks or months. The short duration of many studies was due to practical constraints and the inherent complexities of conducting research involving children with ASD. Longitudinal studies involving children with ASD can present challenges in terms of participant retention, compliance, and data collection consistency over extended periods. In addition, the rapid pace of technological advancements may affect the relevance of the findings if studies are conducted over prolonged durations. However, the researchers did recognize the limitations imposed by

short study durations and provided justifications for the chosen time frames.

3. Diversity of diagnosis and level of functioning: autism is a disorder with a wide variety of symptoms and levels of functioning, which adds complexity and variability to the research.
4. Standardized outcome assessment: some studies did not use standardized outcome assessment tools, which may have affected the comparability and validity of the obtained results.
5. Availability and accessibility of technology: although the presented studies demonstrate the potential of AR technology to support children with autism, it is important to consider the availability and accessibility of this technology in real-world settings. The cost, infrastructure, and availability of AR devices and apps may be limiting factors in the widespread adoption of this technology.

The relatively small sample sizes and short durations commonly observed in many studies involving children with ASD and XR interventions are notable aspects of the research landscape. Although solutions to these issues were not always addressed in the papers, they remain ongoing areas of consideration for researchers in the field. In the selected papers, although some discussions and considerations regarding the challenges of small sample sizes and short study durations were present, comprehensive solutions were not elaborated on. The researchers often acknowledged these limitations and offered potential insights or recommendations, but definitive solutions were not always a primary focus of the papers, the primary focus being on the technical details and potential outcomes of their approach.

In addition, this study itself has several limitations, which should be considered for further research:

1. Limited number of databases queried: despite using comprehensive search strategies, it is possible that some relevant studies were omitted because they were published in nonindexed or less accessible sources.
2. Field evolution: this field of study is rapidly evolving, and new research may have been published since the literature search was conducted. Consequently, this review may not capture the most recent evidence and emerging trends in the field.
3. No distribution of publication authors: this review did not present information on the regional distribution of authors or the origin of the apps and systems, disregarding the influence of cultural differences on the development of these types of apps.
4. Lack of security analysis: this study did not analyze the security issues associated with the proposed solutions.
5. Absence of cost information: no information regarding the cost of the presented solutions could be identified.

Conclusions

This study aimed to conduct a systematic review of the specialized scientific literature in terms of applications, devices, and technologies relevant to the development of AR-, VR-, and MR-based mobile apps dedicated to the therapy of children with ASDs, an objective that was successfully achieved. At the beginning of this paper, the general concept of ASD was presented, after which the RQs and inclusion and exclusion criteria were defined and the results of applying the PRISMA guidelines for the selection of publications to be reviewed were reported. The answers to the RQs were discussed. At the end of the paper, the limitations of the research were presented.

Although the concepts of AR, VR, and MR are not entirely new, their use in the development of therapeutic apps for children with autism has only recently gained popularity. The findings documented in various publications indexed in 5 scientific databases emphasize the suitability of these technologies for such therapy, thereby warranting further in-depth research and the future development of apps based on these technologies. The studies indicated a clear trend toward the use of AR, VR, and MR technologies as a pedagogical tool for people with ASD. This trend involves multidisciplinary collaborations and an integrated approach to research, with a focus on empirical evaluations and ethics regarding the use of technologies. As the field advances, it is essential that research and practice continue to be guided by a balanced and integrated approach that considers both the technological possibilities and the needs and rights of individuals with ASD. However, there are still many issues that require further exploration and research.

Moreover, the publications studied illustrate a wide range of research areas related to the use of AR, VR, and MR in the context of ASD, as well as a variety of methodological and theoretical approaches adopted by the researchers. This suggests that the field is in a phase of rapid growth and diversification, with a wealth of opportunities for future research and development.

Acknowledgments

This research was funded by the 119722/Centru pentru transferul de cuno tin e către întreprinderi din domeniul ICT—CENTRIC, Contract subsidiar 22080/05.10.2022/Autism/ASSIST, Autism ASSISTant—Asistent virtual pentru dezvoltarea abilită ilor cognitive ale copiilor cu patologie de spectru autist project, contract 5/AXA 1/1.2.3/G/13.06.2018, code SMIS 2014+ 119722 (ID P_40_305).

Data Availability

All data generated or analyzed during this study are included in this published article (and its multimedia appendices).

Authors' Contributions

M-VT and CET were involved in the concept and design of this review. All authors made substantial contributions to the analysis, preparation, interpretation, and organization of the data. The final manuscript was read, revised, and approved by all the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.
[PDF File (Adobe PDF File), 108 KB - games_v12i1e49906_app1.pdf]

Multimedia Appendix 2
Quality assessment of the papers.
[XLS File (Microsoft Excel File), 201 KB - games_v12i1e49906_app2.xls]

References

1. Gambrill E. The Diagnostic and Statistical Manual of Mental Disorders as a major form of dehumanization in the modern world. Res Soc Work Pract 2013 Sep 08;24(1):13-36. [doi: 10.1177/1049731513499411]
2. Hyman SL, Levy SE, Myers SM, Council on Children With Disabilities, Section on Developmental and Behavioral Pediatrics. Identification, evaluation, and management of children with autism spectrum disorder. Pediatrics 2020 Jan;145(1):e20193447 [FREE Full text] [doi: 10.1542/peds.2019-3447] [Medline: 31843864]
3. Autism. World Health Organization. 2023 Nov 15. URL: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders [accessed 2024-02-09]
4. Rice CE, Rosanoff M, Dawson G, Durkin MS, Croen LA, Singer A, et al. Evaluating changes in the prevalence of the autism spectrum disorders (ASDs). Public Health Rev 2012;34(2):1-22 [FREE Full text] [doi: 10.1007/BF03391685] [Medline: 26236074]
5. Vahia VN. Diagnostic and Statistical Manual of Mental Disorders 5: a quick glance. Indian J Psychiatry 2013 Jul;55(3):220-223 [FREE Full text] [doi: 10.4103/0019-5545.117131] [Medline: 24082241]
6. Munson J, Pasqual P. Using technology in autism research: the promise and the perils. Computer 2012 Jun;45(6):89-91. [doi: 10.1109/mc.2012.220]
7. Knapp M, Romeo R, Beecham J. The economic consequences of autism in the UK. Foundation for People With Learning Disabilities. 2007 Nov. URL: https://www.learningdisabilities.org.uk/file/2127/download?token=1kQCeb5s&__cf_chl_tk=OAY.wBxg5_KGHTftsLyp1Sn8sYw3ulK4LlemFc.1d1o-1707157072-0-gaNycGzNEHs [accessed 2024-02-09]
8. Meadan H, Ostrosky MM, Triplett B, Michna A, Fettig A. Using visual supports with young children with autism spectrum disorder. Teach Except Child 2011 Jul 01;43(6):28-35. [doi: 10.1177/004005991104300603]
9. Abad-Segura E, González-Zamar MD, Luque-de la Rosa AL, Morales Cevallos MB. Sustainability of educational technologies: an approach to augmented reality research. Sustainability 2020 May 16;12(10):4091. [doi: 10.3390/su12104091]
10. El-Jarn H, Southern G. Can co-creation in extended reality technologies facilitate the design process? J Work Appl Manag 2020 Jul 03;12(2):191-205. [doi: 10.1108/jwam-04-2020-0022]
11. Azuma RT. A survey of augmented reality. Presence (Camb) 1997 Aug;6(4):355-385. [doi: 10.1162/pres.1997.6.4.355]
12. Wagner D, Schmalstieg D. First steps towards handheld augmented reality. In: Proceedings of the Seventh IEEE International Symposium on Wearable Computers, 2003. 2003 Presented at: Seventh IEEE International Symposium on Wearable Computers, 2003; October 21-23, 2003; White Plains, NY. [doi: 10.1109/iswc.2003.1241402]
13. Pokémon GO home page. Pokémon GO. URL: https://pokemongolive.com [accessed 2024-02-09]
14. Investment in augmented and virtual reality (AR/VR) technology worldwide in 2024, by use case. Statista. URL: https://www.statista.com/statistics/1098345/worldwide-ar-vr-investment-use-case/ [accessed 2024-02-09]
15. Howorth SK, Rooks-Ellis D, Flanagan S, Ok MW. Augmented reality supporting reading skills of students with autism spectrum disorder. Interv Sch Clin 2019 Oct 14;55(2):71-77. [doi: 10.1177/1053451219837635]
16. Martin S, Diaz G, Sancristobal E, Gil R, Castro M, Peire J. New technology trends in education: seven years of forecasts and convergence. Comput Educ 2011 Nov;57(3):1893-1906. [doi: 10.1016/j.compedu.2011.04.003]
17. Dror I, Schmidt P, O'connor L. A cognitive perspective on technology enhanced learning in medical training: great opportunities, pitfalls and challenges. Med Teach 2011 Apr;33(4):291-296. [doi: 10.3109/0142159x.2011.550970]
18. Richard E, Billaudeau V, Richard P, Gaudin G. Augmented reality for rehabilitation of cognitive disabled children: a preliminary study. In: Proceeding of the 2007 Virtual Rehabilitation. 2007 Presented at: 2007 Virtual Rehabilitation; September 27-29, 2007; Venice, Italy. [doi: 10.1109/ICVR.2007.4362148]
19. Bursali H, Yilmaz RM. Effect of augmented reality applications on secondary school students' reading comprehension and learning permanency. Comput Hum Behav 2019 Jun;95:126-135. [doi: 10.1016/j.chb.2019.01.035]

20. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Transact Inf Syst 1994 Dec;E77-D(12):1321-1329 [FREE Full text]
21. Fujiuchi K, Riggie J. Academic library collections in the age of extended reality (XR). Collect Manag 2019 Feb 04;44(2-4):296-303. [doi: 10.1080/01462679.2019.1566109]
22. Eishita FZ, Stanley KG. The impact on player experience in augmented reality outdoor games of different noise models. Entertain Comput 2018 Aug;27:137-149. [doi: 10.1016/j.entcom.2018.04.006]
23. Çöltekin A, Lochhead I, Madden M, Christophe S, Devaux A, Pettit C, et al. Extended reality in spatial sciences: a review of research challenges and future directions. ISPRS Int J Geo Inf 2020 Jul 15;9(7):439. [doi: 10.3390/ijgi9070439]
24. Kitchenham B. Procedures for performing systematic reviews. Keele University. 2004 Jul. URL: https://www.researchgate.net/publication/228756057_Procedures_for_Performing_Systematic_Reviews [accessed 2024-02-09]
25. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
26. Paradis S, Roussel J, Bosson JL, Kern JB. Use of smartphone health apps among patients aged 18 to 69 years in primary care: population-based cross-sectional survey. JMIR Form Res 2022 Jun 16;6(6):e34882 [FREE Full text] [doi: 10.2196/34882] [Medline: 35708744]
27. Number of smartphones sold to end users worldwide from 2007 to 2022. Statista. URL: https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007/ [accessed 2023-12-23]
28. Hashim HU, Yunus MM, Norman H. 'AReal-Vocab': an augmented reality English vocabulary mobile application to cater to mild autism children in response towards sustainable education for children with disabilities. Sustainability 2022 Apr 18;14(8):4831. [doi: 10.3390/su14084831]
29. Machado E, Carrillo I, Saldana D, Chen F, Chen L. An assistive augmented reality-based smartglasses solution for individuals with autism spectrum disorder. In: Proceedings of the 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech). 2019 Presented at: DASC/PiCom/CBDCom/CyberSciTech; August 5-8, 2019; Fukuoka, Japan. [doi: 10.1109/dasc/picom/cbdcom/cyberscitech.2019.00052]
30. Tang TY, Xu J, Winoto P. An augmented reality-based word-learning mobile application for children with autism to support learning anywhere and anytime: object recognition based on deep learning. In: Proceedings of the 13th International Conference, UAHCI 2019, Held as Part of the 21st HCI International Conference, HCII 2019. 2019 Presented at: 13th International Conference, UAHCI 2019, Held as Part of the 21st HCI International Conference, HCII 2019; July 26-31, 2019; Orlando, FL. [doi: 10.1007/978-3-030-23563-5_16]
31. Selvarani AG, T NS. An interactive number learning augmented reality application for autistic preschool children. In: Proceedings of the 2022 International Conference on Computer Communication and Informatics (ICCCI). 2022 Presented at: ICCCI 2022; January 25-27, 2022; Coimbatore, India. [doi: 10.1109/iccci54379.2022.9740948]
32. Abou El-Seoud S, Halabi O, Geroimenko V. Assisting individuals with autism and cognitive disorders: an augmented reality-based framework. Int J Online Biomed Eng 2019 Feb 27;15(04):28-39. [doi: 10.3991/ijoe.v15i04.9835]
33. Vullamparthi AJ, Nelaturu SC, Mallaya DD, Chandrasekhar S. Assistive learning for children with autism using augmented reality. In: Proceedings of the IEEE Fifth International Conference on Technology for Education (t4e 2013). 2013 Presented at: t4e 2013; December 18-20, 2013; Kharagpur, India. [doi: 10.1109/t4e.2013.18]
34. Singh K, Shrivastava A, Achary K, Dey A, Sharma O. Augmented reality-based procedural task training application for less privileged children and autistic individuals. In: Proceedings of the 17th International Conference on Virtual-Reality Continuum and its Applications in Industry. 2019 Presented at: VRCAI '19; November 14-16, 2019; Brisbane, Australia. [doi: 10.1145/3359997.3365703]
35. Chen CH, Lee IJ, Lin LY. Serious games for people with mental disorders: state of the art of practices to maintain engagement and accessibility. Comput Hum Behav 2016 Feb;55:477-485 [FREE Full text] [doi: 10.1016/j.chb.2015.09.033]
36. Tang TY, Xu J, Winoto P. Automatic object recognition in a light-weight augmented reality-based vocabulary learning application for children with autism. In: Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence. 2019 Presented at: ICIAI '19; March 15-18, 2019; Suzhou, China. [doi: 10.1145/3319921.3319945]
37. Giraud T, Ravenet B, Dang CT, Nadel J, Prigent E, Poli G, et al. "Can you help me move this over there?": training children with ASD to joint action through tangible interaction and virtual agent. In: Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction. 2021 Presented at: TEI '21; February 14-17, 2021; Salzburg, Austria. [doi: 10.1145/3430524.3440646]
38. Zheng ZK, Sarkar N, Swanson A, Weitlauf A, Warren Z, Sarkar N. CheerBrush: a novel interactive augmented reality coaching system for toothbrushing skills in children with autism spectrum disorder. ACM Trans Access Comput 2021 Oct 15;14(4):1-20. [doi: 10.1145/3481642]
39. Pradibta H, Wijaya ID. Designing AR daily prayers for children with ASD. In: Proceedings of the International Seminar on Application for Technology of Information and Communication (iSemantic). 2017 Presented at: iSemantic; October 07-08, 2017; Semarang, Indonesia. [doi: 10.1109/isemantic.2017.8251851]

40. Nubia RM, Fabian GR, Wilson RA, Wilmer PB. Development of a mobile application in augmented reality to improve the communication field of autistic children at a Neurorehabilitar Clinic. In: Proceedings of the Workshop on Engineering Applications - International Congress on Engineering (WEA). 2015 Presented at: WEA 2015; October 28-30, 2015; Bogota, Colombia. [doi: 10.1109/wea.2015.7370154]
41. Sait M, Alattas A, Omar A, Almalki S, Sharf S, Alsaggaf E. Employing virtual reality techniques in environment adaptation for autistic children. Procedia Comput Sci 2019;163:338-344. [doi: 10.1016/j.procs.2019.12.116]
42. Wan G, Deng F, Jiang Z, Song S, Hu D, Chen L, et al. FECTS: a facial emotion cognition and training system for Chinese children with autism spectrum disorder. Comput Intell Neurosci 2022;2022:9213526 [FREE Full text] [doi: 10.1155/2022/9213526] [Medline: 35528364]
43. V K, Mohan A, S DP. Immersive learning aid for children with autism (ASD) using object recognition. In: Proceedings of the 3rd International Conference on Signal Processing and Communication (ICPSC). 2021 Presented at: ICPSC; May 13-14, 2021; Coimbatore, India. [doi: 10.1109/icspc51351.2021.9451646]
44. Silva SD, Neto FM, de Lima RM, de Macêdo FT, Santo JR, Silva WL. Knowledgemon hunter: a serious game with geolocation to support learning of children with autism and learning difficulties. In: 19th Symposium on Virtual and Augmented Reality (SVR). 2017 Presented at: SVR; November 01-04, 2017; Curitiba, Brazil. [doi: 10.1109/svr.2017.45]
45. Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, et al. Labeling images with facial emotion and the potential for pediatric healthcare. Artif Intell Med 2019 Jul;98:77-86 [FREE Full text] [doi: 10.1016/j.artmed.2019.06.004] [Medline: 31521254]
46. Escobedo L, Nguyen DH, Boyd L, Hirano S, Rangel A, Garcia-Rosas D, et al. MOSOCO: a mobile assistive tool to support children with autism practicing social skills in real-life situations. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, Texas. [doi: 10.1145/2207676.2208649]
47. Xia M, Chen N, Tang Y, Zhu Z. ParaShop: a mobile AR app in assisting people with ASD in shopping. In: Proceedings of the IEEE MIT Undergraduate Research Technology Conference (URTC). 2020 Presented at: MIT URTC; October 9-11, 2020; Cambridge, MA. [doi: 10.1109/urtc51696.2020.9668867]
48. Amado ML, Ruiz LC, Andrade-Arenas L. Prototype of an augmented reality application for cognitive improvement in children with autism using the DesingScrum methodology. Adv Sci Technol Eng Syst J 2021 Jan;6(1):587-596. [doi: 10.25046/aj060163]
49. Voss C, Washington P, Haber N, Kline A, Daniels J, Fazel A, et al. Superpower glass: delivering unobtrusive real-time social cues in wearable systems. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. 2016 Presented at: UbiComp '16; September 12-16, 2016; Heidelberg, Germany. [doi: 10.1145/2968219.2968310]
50. Washington P, Voss C, Kline A, Haber N, Daniels J, Fazel A, et al. SuperpowerGlass: a wearable aid for the at-home therapy of children with autism. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Sep 11;1(3):1-22. [doi: 10.1145/3130977]
51. Gulati R, Handa K. Towards development of a virtual reality game for children with autism spectrum disorder. In: Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). 2021 Presented at: ICICV; February 04-06, 2021; Tirunelveli, India. [doi: 10.1109/icicv50876.2021.9388374]
52. Escobedo L, Tentori M, Quintana E, Favela J, Garcia-Rosas D. Using augmented reality to help children with autism stay focused. IEEE Pervasive Comput 2014 Jan;13(1):38-46. [doi: 10.1109/mprv.2014.19]
53. Bouaziz R, Alhejaili M, Al-Saedi R, Mihdhar A, Alsarrani J. Using marker based augmented reality to teach autistic eating skills. In: Proceedings of the IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR). 2020 Presented at: AIVR 2020; December 14-18, 2020; Utrecht, The Netherlands. [doi: 10.1109/aivr50618.2020.00050]
54. Wang K, Zhang B, Cho Y. Using mobile augmented reality to improve attention in adults with autism spectrum disorder. In: Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3334480.3383090]
55. Gelsomini M, Garzotto F, Montesano D, Occhiuto D. Wildcard: a wearable virtual reality storytelling tool for children with intellectual developmental disability. In: Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016 Presented at: EMBC'16; August 16-20, 2016; Orlando, FL. [doi: 10.1109/embc.2016.7591896]
56. Lian X, Sunar MS. Mobile augmented reality technologies for autism spectrum disorder interventions: a systematic literature review. Appl Sci 2021 May 17;11(10):4550. [doi: 10.3390/app11104550]
57. López-Belmonte J, Moreno-Guerrero AJ, Marín-Marín JA, Lampropoulos G. The impact of gender on the use of augmented reality and virtual reality in students with ASD. Educ Knowl Soc 2022;23(1):e28418 [FREE Full text] [doi: 10.14201/eks.28418]
58. Gea M, Alaman X, Rodriguez P, Martinez V. Towards smart and inclusive society: building 3D immersive museum by children with cognitive disabilities. In: Proceedings of the 8th International Conference on Education and New Learning Technologies. 2016 Presented at: EDULEARN16 Proceedings; July 4-6, 2016; Barcelona, Spain.

59. Baragash RS, Al-Samarraie H, Alzahrani AI, Alfarraj O. Augmented reality in special education: a meta-analysis of single-subject design studies. Eur J Special Needs Educ 2019 Dec 20;35(3):382-397 [FREE Full text] [doi: 10.1080/08856257.2019.1703548]
60. Glaser N, Newbutt N, Palmer H, Schmidt M. Video-based virtual reality technology for autistic users: an emerging technology report. Tech Know Learn 2022 Mar 26;28(2):925-935. [doi: 10.1007/s10758-022-09594-x]
61. Pradipta H, Wijaya ID, Ronilaya F, Nurhasan U. The evaluation on a mobile augmented reality application as therapy media. In: Proceedings of the 3rd Annual Applied Science and Engineering Conference. 2018 Presented at: AASEC 2018; April 18, 2018; Bandung, Indonesia.
62. Hashim HU, Yunus MM, Norman H. Augmented reality mobile application for children with autism : stakeholders' acceptance and thoughts. Arab World Engl J 2021 Dec;12(4):132-141 [FREE Full text] [doi: 10.24093/awej/vol12no4.9]
63. Johnston D, Egermann H, Kearney G. Innovative computer technology in music-based interventions for individuals with autism moving beyond traditional interactive music therapy techniques. Cogent Psychol 2018 Dec 11;5(1):1554773. [doi: 10.1080/23311908.2018.1554773]
64. Putnam C, Hanschke C, Todd J, Gemmell J, Kollia M. Interactive technologies designed for children with autism: reports of use and desires from parents, teachers, and therapists. ACM Trans Access Comput 2019 Sep 12;12(3):1-37. [doi: 10.1145/3342285]
65. Mukherjee D, Bhavnani S, Lockwood Estrin G, Rao V, Dasgupta J, Irfan H, et al. Digital tools for direct assessment of autism risk during early childhood: a systematic review. Autism 2024 Jan;28(1):6-31 [FREE Full text] [doi: 10.1177/13623613221133176] [Medline: 36336996]
66. Valentine AZ, Brown BJ, Groom MJ, Young E, Hollis C, Hall CL. A systematic review evaluating the implementation of technologies to assess, monitor and treat neurodevelopmental disorders: a map of the current evidence. Clin Psychol Rev 2020 Aug;80:101870 [FREE Full text] [doi: 10.1016/j.cpr.2020.101870] [Medline: 32712216]
67. Khowaja K, Al-Thani D, Banire B, Salim SS, Shah A. Use of augmented reality for social communication skills in children and adolescents with autism spectrum disorder (ASD): a systematic review. In: Proceedings of the IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS). 2019 Presented at: ICETAS; December 20-21, 2019; Kuala Lumpur, Malaysia URL: https://ieeexplore.ieee.org/document/9117290/authors#authors
68. Almurashi HA, Bouaziz R. Augmented reality and autism spectrum disorder rehabilitation: scoping review. In: Proceedings of the IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR). 2020 Presented at: AIVR 2020; December 14-18, 2020; Utrecht, Netherlands URL: https://ieeexplore.ieee.org/document/9319123/authors#authors
69. Farzana W, Sarker F, Chau T, Mamun KA. Technological evolvement in AAC modalities to foster communications of verbally challenged ASD children: a systematic review. IEEE Access 2021 Jan 28:1-1. [doi: 10.1109/ACCESS.2021.3055195]
70. Junaidi AR, Irvan M, Yuwono J, Dewantoro DA, Alamsyah Y, Mulyawati NW. Usability testing of developed virtual reality environment to enhance communication skills for children with low functioning autism spectrum disorder. In: Proceedings of the 2nd International Conference on Information Technology and Education (ICIT&E). 2022 Presented at: ICIT&E 2022; January 22, 2022; Malang, Indonesia URL: https://ieeexplore.ieee.org/document/9759890/authors#authors
71. Adamu AS, Abdullahi SE. A review of autism diagnosis/screening expert system and mobile application. In: Proceedings of the IEEE 2nd International Conference on Cyberspac (CYBER NIGERIA). 2021 Presented at: CYBER NIGERIA; February 23-25, 2021; Abuja, Nigeria URL: https://ieeexplore.ieee.org/document/9428833
72. Nahavandi D, Alizadehsani R, Khosravi A, Acharya UR. Application of artificial intelligence in wearable devices: opportunities and challenges. Comput Methods Programs Biomed 2022 Jan;213:106541. [doi: 10.1016/j.cmpb.2021.106541] [Medline: 34837860]
73. Teófilo MR, Lourenço AA, Postal J, Silva YM, Lucena VFJ. The raising role of virtual reality in accessibility systems. Proc Comput Sci 2019;160:671-677. [doi: 10.1016/j.procs.2019.11.029]
74. Wattanasoontorn V, Boada I, García RJ, Sbert M. Serious games for health. Entertain Comput 2013 Dec;4(4):231-247. [doi: 10.1016/j.entcom.2013.09.002]
75. Purnama Y, Herman FA, Hartono J, Nielsen, Suryani D, Sanjaya G. Educational software as assistive technologies for children with autism spectrum disorder. Proc Comput Sci 2021 Feb;179:6-16. [doi: 10.1016/j.procs.2020.12.002]
76. de Albuquerque AP, Kelner J. Toy user interfaces: systematic and industrial mapping. J Syst Archit 2019 Aug;97:77-106. [doi: 10.1016/j.sysarc.2018.12.001]
77. Du Y, Salen Tekinbas K. Bridging the gap in mobile interaction design for children with disabilities: perspectives from a pediatric speech language pathologist. Int J Child Comput Interact 2020 Jun;23-24:100152. [doi: 10.1016/j.ijcci.2019.100152]
78. Zhu L, Cao Q, Cai Y. Development of augmented reality serious games with a vibrotactile feedback jacket. Virtual Real Intell Hardw 2020 Oct;2(5):454-470. [doi: 10.1016/j.vrih.2020.05.005]
79. Cholewa N, Wołk K, Wołk R. Precise eye-tracking technology in medical communicator prototype. Proc Comput Sci 2018;138:264-271. [doi: 10.1016/j.procs.2018.10.038]
80. Benton L, Johnson H. Widening participation in technology design: a review of the involvement of children with special educational needs and disabilities. Int J Child Comput Interact 2015 Jan;3-4:23-40. [doi: 10.1016/j.ijcci.2015.07.001]

81. Niknejad N, Ismail WB, Mardani A, Liao H, Ghani I. A comprehensive overview of smart wearables: the state of the art literature, recent advances, and future challenges. Eng Appl Artif Intell 2020 Apr;90:103529. [doi: 10.1016/j.engappai.2020.103529]
82. Theodoropoulos A, Lepouras G. Augmented Reality and programming education: a systematic review. Int J Child Comput Interact 2021 Dec;30:100335. [doi: 10.1016/j.ijcci.2021.100335]
83. Hassan A, Pinkwart N, Shafi M. Serious games to improve social and emotional intelligence in children with autism. Entertain Comput 2021 May;38:100417. [doi: 10.1016/j.entcom.2021.100417]
84. Shohieb SM, Doenyas C, Elhady AM. Dynamic difficulty adjustment technique-based mobile vocabulary learning game for children with autism spectrum disorder. Entertain Comput 2022 May;42:100495. [doi: 10.1016/j.entcom.2022.100495]
85. Soysa AI, Mahmud AA. Assessing tablet applications focused on social interactions: what functionalities do Sri Lankan practitioners want for children with ASD? In: Proceedings of the 30th Australian Conference on Computer-Human Interaction. 2018 Presented at: OzCHI '18; December 4-7, 2018; Melbourne, Australia.
86. Abdullah MH, Brereton MF. Mycalendar: supporting children on the autism spectrum to learn language and appropriate behaviour. In: Proceedings of the 29th Australian Conference on Computer-Human Interaction. 2017 Presented at: OzCHI '17; November 28-December 1, 2017; Brisbane, Australia.
87. Sharmin M, Hossain MM, Saha A, Das M, Maxwell M, Ahmed S. From research to practice: informing the design of autism support smart technology. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC.
88. Garcia-Garcia JM, Cabañero MD, Penichet VM, Lozano MD. EmoTEA: teaching children with autism spectrum disorder to identify and express emotions. In: Proceedings of the XX International Conference on Human Computer Interaction. 2019 Presented at: Interacción '19; June 25-28, 2019; Donostia, Spain.
89. Miri P, Arora M, Malhotra A, Flory R, Hu S, Lowber A, et al. FAR: end-to-end vibrotactile distributed system designed to facilitate affect regulation in children diagnosed with autism spectrum disorder through slow breathing. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA.
90. Li B, Atyabi A, Kim M, Barney E, Ahn AY, Luo Y, et al. Social influences on executive functioning in autism: design of a mobile gaming platform. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC.
91. Alessandrini A, Loux V, Serra GF, Murray C. Designing ReduCat: audio-augmented paper drawings tangible interface in educational intervention for high-functioning autistic children. In: Proceedings of the The 15th International Conference on Interaction Design and Children. 2016 Presented at: IDC '16; June 21-24, 2016; Manchester, UK.
92. Motti VG. Designing emerging technologies for and with neurodiverse users. In: Proceedings of the 37th ACM International Conference on the Design of Communication. 2019 Presented at: SIGDOC '19; October 4-6, 2019; Portland, Oregon.
93. Da Branco K, Pinheiro VM, Damian AL, Marques AB. Investigating the first user experience and accessibility of educational applications for autistic children. In: Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems. 2020 Presented at: IHC '20; October 26-30, 2020; Diamantina, Brazil.
94. Cordeiro RF, Ferreira WS, Aguiar YP, Saraiva JA, Tardif C, Galy E. The Brazilian challenge to accessibility and digital inclusion for people with autistic spectrum disorders. In: Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems. 2018 Presented at: IHC '18; October 22-26, 2018; Belém, Brazil.
95. Hachisu T, Bourreau B, Suzuki K. EnhancedTouchX: smart bracelets for augmenting interpersonal touch interactions. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, Scotland.
96. Duval J, Bertran FA, Chen S, Chu M, Subramonian D, Wang A, et al. Chasing play on TikTok from populations with disabilities to inspire playful and inclusive technology design. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan.
97. Mohd CK, Shahbodin F, Jano Z, Azni AH. Visual perception games for autistic learners: design and development. In: Proceedings of the 2019 Asia Pacific Information Technology Conference. 2019 Presented at: APIT '19; January 25-27, 2019; Jeju Island, Republic of Korea.
98. Yaneva V, Temnikova I, Mitkov R. Accessible texts for autism: an eye-tracking study. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. 2015 Presented at: ASSETS '15; October 26-28, 2015; Lisbon, Portugal.
99. Venuprasad P, Dobhal T, Paul A, Nguyen TN, Gilman A, Cosman P, et al. Characterizing joint attention behavior during real world interactions using automated object and gaze detection. In: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications. 2019 Presented at: ETRA '19; June 25-28, 2019; Denver, CO.
100. Bozgeyikli L, Bozgeyikli E, Raij A, Alqasemi R, Katkoori S, Dubey R. Vocational rehabilitation of individuals with autism spectrum disorder with virtual reality. ACM Trans Access Comput 2017 Apr 06;10(2):1-25. [doi: 10.1145/3046786]
101. Spiel K, Frauenberger C, Keyes O, Fitzpatrick G. Agency of autistic children in technology research—a critical literature review. ACM Trans Comput Hum Interact 2019 Nov 02;26(6):1-40. [doi: 10.1145/3344919]

102. Romero MR, Macas E, Harari I, Diaz J. Is it possible to improve the learning of children with ASD through augmented reality mobile applications? In: Proceedings of the Applied Technologies. 2019 Presented at: ICAT 2019; December 3-5, 2019; Quito, Ecuador.
103. Mazumdar A, Chatterjee B, Banerjee M, Bhati IH. Mobile application based modified screening and assessment tools for children with autism. Int J Interact Mob Technol 2019 Aug;13(08):159. [doi: 10.3991/ijim.v13i08.10563]
104. Signore A, Balasi P, Yuan T. You talk! – YOU vs AUTISM. In: Proceedings of the Computers Helping People with Special Needs. 2014 Presented at: ICCHP 2014; July 9-11, 2014; Paris, France.

Abbreviations

AR: augmented reality
ASD: autism spectrum disorder
MR: mixed reality
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QC: quality criterion
RQ: research question
VR: virtual reality
XR: extended reality

Edited by T de Azevedo Cardoso; submitted 13.06.23; peer-reviewed by B Bao, G Caggianese; comments to author 16.08.23; revised version received 30.08.23; accepted 21.12.23; published 19.02.24.

Please cite as:
Toma MV, Turcu CE, Turcu CO, Vlad S, Tiliute DE, Pascu P
Extended Reality–Based Mobile App Solutions for the Therapy of Children With Autism Spectrum Disorders: Systematic Literature Review
JMIR Serious Games 2024;12:e49906
URL: https://games.jmir.org/2024/1/e49906
doi:10.2196/49906
PMID:38373032

Review

Immersive Virtual Reality–Based Methods for Assessing Executive Functioning: Systematic Review

Rebecca Kirkham[1], BPsy (Hons); Lars Kooijman[2], MSc; Lucy Albertella[1], PhD; Dan Myles[1], PhD; Murat Yücel[3,4*], PhD; Kristian Rotaru[1,5*], PhD

[1]Turner Institute for Brain and Mental Health, School of Psychological Sciences and Monash Biomedical Imaging Facility, Monash University, Clayton, Australia

[2]Institute for Intelligent Systems Research and Innovation, Deakin University, Geelong, Australia

[3]Queensland Institute of Medical Research Berghofer Medical Research Institute, Herston, Australia

[4]Department of Psychiatry, School of Clinical Sciences, Monash University, Clayton, Australia

[5]Monash Business School, Monash University, Caufield, Australia

[*]these authors contributed equally

Corresponding Author:
Rebecca Kirkham, BPsy (Hons)
Turner Institute for Brain and Mental Health
School of Psychological Sciences and Monash Biomedical Imaging Facility
Monash University
770 Blackburn Road
Clayton, 3168
Australia
Phone: 61 (03) 9905 0100
Email: rebecca.kirkham@monash.edu

Abstract

Background: Neuropsychological assessments traditionally include tests of executive functioning (EF) because of its critical role in daily activities and link to mental disorders. Established traditional EF assessments, although robust, lack ecological validity and are limited to single cognitive processes. These methods, which are suitable for clinical populations, are less informative regarding EF in healthy individuals. With these limitations in mind, immersive virtual reality (VR)–based assessments of EF have garnered interest because of their potential to increase test sensitivity, ecological validity, and neuropsychological assessment accessibility.

Objective: This systematic review aims to explore the literature on immersive VR assessments of EF focusing on (1) EF components being assessed, (2) how these assessments are validated, and (3) strategies for monitoring potential adverse (cybersickness) and beneficial (immersion) effects.

Methods: EBSCOhost, Scopus, and Web of Science were searched in July 2022 using keywords that reflected the main themes of VR, neuropsychological tests, and EF. Articles had to be peer-reviewed manuscripts written in English and published after 2013 that detailed empirical, clinical, or proof-of-concept studies in which a virtual environment using a head-mounted display was used to assess EF in an adult population. A tabular synthesis method was used in which validation details from each study, including comparative assessments and scores, were systematically organized in a table. The results were summed and qualitatively analyzed to provide a comprehensive overview of the findings.

Results: The search retrieved 555 unique articles, of which 19 (3.4%) met the inclusion criteria. The reviewed studies encompassed EF and associated higher-order cognitive functions such as inhibitory control, cognitive flexibility, working memory, planning, and attention. VR assessments commonly underwent validation against gold-standard traditional tasks. However, discrepancies were observed, with some studies lacking reported a priori planned correlations, omitting detailed descriptions of the EF constructs evaluated using the VR paradigms, and frequently reporting incomplete results. Notably, only 4 of the 19 (21%) studies evaluated cybersickness, and 5 of the 19 (26%) studies included user experience assessments.

Conclusions: Although it acknowledges the potential of VR paradigms for assessing EF, the evidence has limitations. The methodological and psychometric properties of the included studies were inconsistently addressed, raising concerns about their validity and reliability. Infrequent monitoring of adverse effects such as cybersickness and considerable variability in sample sizes may limit interpretation and hinder psychometric evaluation. Several recommendations are proposed to improve the theory and practice of immersive VR assessments of EF. Future studies should explore the integration of biosensors with VR systems and the capabilities of VR in the context of spatial navigation assessments. Despite considerable promise, the systematic and validated implementation of VR assessments is essential for ensuring their practical utility in real-world applications.

(JMIR Serious Games 2024;12:e50282) doi:10.2196/50282

KEYWORDS

virtual reality; executive functioning; neuropsychological assessment; systematic review; psychometric properties; cybersickness; immersion; cognition

Introduction

Background

Executive functioning (EF) has long been a focus of neuropsychological assessment because of the significant role it plays in everyday functioning. EF is an umbrella term for higher-order cognitive skills used to control and coordinate a wide range of mental processes and everyday behaviors [1-5], including "...mentally playing with ideas; taking the time to think before acting; meeting novel, unanticipated challenges; resisting temptations; and staying focused" [6]. Although a universally accepted definition of EF does not exist [5], there is agreement on the attributes of 3 core executive functions: inhibition, cognitive flexibility, and working memory [2,4,6]. These core executive functions support other higher-order executive functions such as reasoning, planning, and problem-solving [6-8]. As EF impairment has been linked to a variety of mental disorders [9], it is often considered a transdiagnostic risk factor [10].

Although traditional methods used to assess EF are popular [11,12] and well validated [13], they have been criticized for their lack of ecological validity [14,15]. Ecological validity, within the scope of this study, is defined as the "functional and predictive relationship between the person's performance on a set of neuropsychological tests and the person's behavior in a variety of real world settings" [16]. Specifically, we interpret ecological validity as comprising 2 principal components: representativeness—the degree to which a neuropsychological test mirrors the demands of a person's daily living activities that it aims to evaluate [17], sometimes referred to as *verisimilitude* [18]—and generalizability—the extent to which test performance predicts an individual's functioning in their daily living activities [17], also known as veridicality [18].

Traditional assessments tend to take a "construct-led" approach, with each test intended to isolate a single cognitive process in an abstract measure. This process of abstraction may limit the ecological validity of the measure by resulting in poor alignment between the test outcomes and real-world functioning. In turn, this produces a large amount of variance in EF that is unaccounted for by traditional tasks. For example, Chaytor et al [19] noted that traditional EF tests accounted for only 18% to 20% of the variance in the everyday executive ability of participants. This lack of explained variance may be attributed to the nature of the testing environment, the constructs assessed in isolation, the participant's affective state, and the compensatory strategies available to the participant [19]. A related methodological issue, known as the "task impurity problem" [4,20], indicates that the score on an EF task usually reflects not only the systematic variance attributable to the specific aspect of EF targeted by that task but also the (1) systematic variance across multiple types of EF tasks, (2) systematic variance attributable to non-EF aspects of the task, and (3) nonsystematic (error) variance (see the study by Snyder et al [10] for a detailed review). Outside the testing environment, the process of making a decision or planning and eliciting goal-directed behavior in everyday life is often highly dynamic and influenced by numerous internal and external factors [13,14]. Therefore, an ecologically valid assessment tool will need to include relevant contextual, dynamic, and multidimensional features such as affect and physiological state, which traditional assessments cannot include.

Furthermore, although traditional EF assessment tools may be appropriate for clinical populations, they generate less information about functioning in relatively healthy individuals. For example, the Trail-Making Test (TMT) has routinely been administered as a neuropsychological assessment of driving performance. Although some studies have demonstrated a relationship between the two [21,22], others have shown no relationship [23], particularly in nonclinical populations [24,25]. Thus, although traditional tools are adequate for detecting more severe EF impairments, they are less effective in detecting subtle changes in EF and early decline. Increased test sensitivity to detect subtle intraindividual changes may enable better detection of the prodromal stages of cognitive decline. Early detection is important as it enables early intervention, which may in turn improve prognosis. For example, sensitive detection can identify the prodromal stages of Alzheimer disease in seemingly healthy individuals [26] and mild cognitive decline up to 12 years before clinical diagnosis [27]. Similarly, in a situation in which an individual requires a capacity assessment for an activity, traditional assessments may have limited utility for nonclinical populations. The triangulation of multiple data sources such as biosensors may increase sensitivity to better identify subtle changes in capacity.

To address the shortcomings of poor ecological validity and test sensitivity, research on psychological assessment has begun to investigate virtual reality (VR) technology as a means of providing a more naturalistic environment for evaluating EF in clinical neuropsychological assessments. VR enables the development of custom-designed simulated environments that can replicate real-life environments, potentially increasing its ecological validity through representativeness. In addition, VR could increase engagement [28,29], reduce test time, and better integrate data from biosensors with in-task events that facilitate assessment. The following sections will expand on these points and consider the importance of validating and assessing the reliability of VR for EF assessment.

Ecological Validity and Representative Tests

There is an increasing emphasis on conducting EF assessments using tasks that resemble situations experienced in everyday life [30]. For example, the Multiple Errands Test (MET) [31] requires individuals to run errands in a real environment (eg, a shopping center). Empirical assessment of the MET has demonstrated its generalizability to daily functioning [32] and carer reports of daily functioning [33]. However, given that the MET is designed to be performed in real-life locations, it is impractical for routine administration by clinicians [34,35] and susceptible to the variable features of real-world environments that are outside experimental control. VR can mitigate these difficulties by maintaining the real-world environment without requiring travel while enabling fine-tuned control and uniform presentation of environmental characteristics [36]. Several studies [37-39] have investigated and developed platforms for this purpose, commonly known as the virtual MET.

Engagement

VR has the potential to enhance individual engagement more effectively than traditional pencil-and-paper or computerized tasks by offering a fully immersive experience [40]. Recognized as a crucial aspect of cognitive assessment, engagement can be improved through gamification, thereby improving task performance [41]. “Serious games,” defined as games intended for a variety of serious purposes, such as training, learning, stimulation, or cognitive assessment [42], have been shown to be more engaging than nongamified tasks [43-45]. The unique immersive environment of VR captures increased attention, leading to reduced average response times and response time variability [46]. Notably, recent studies using electroencephalography (EEG)-based metrics have shown greater attention elicited in immersive VR paradigms than in 2D computerized assessments [46]. This heightened immersion and engagement in VR may enhance the reliability of the measures by capturing a more accurate representation of an individual’s best effort.

Cybersickness

Despite their increased engagement, VR paradigms have the potential to induce cybersickness, which can threaten the validity of the paradigm. Cybersickness (ie, dizziness and vertigo) is akin to motion sickness but occurs in response to exposure to VR [47]. Previous research suggests that there is a negative relationship between cybersickness and cognitive abilities. For example, Nalivaiko et al [47] found that reaction times were moderately correlated (r=0.5; P=.006) with subjective ratings of nausea. Similarly, Sepich et al [48] found that participants’ accuracy on n-back task performance was weakly to moderately negatively correlated (r=−0.32; P=.002) with subjective cybersickness ratings. Therefore, there is reasonable concern that the potential benefits of engagement and ecological validity may be compromised if participants experience cybersickness.

Validity, Reliability, and Sensitivity

Arguably, the biggest threat to the utility of VR platforms is that many studies do not document their validity and reliability. A meta-analysis showed that VR assessment tools are moderately sensitive to cognitive impairment across neurodevelopmental, mental health, and neurological disorders [49], demonstrating their promising application in clinical settings. Borgnis et al [50] reviewed the VR-based tools for EF assessment that are currently available, illustrating the plethora of platforms developing in this field. The works by Negu et al [49] and Borgnis et al [50] highlight the utility of VR assessment tools to detect dysfunction and present the various tools in the literature created to investigate EF. Kim et al [51] provided an overview of the research trends using VR for neuropsychological tests and documented the cognitive functions assessed in each study. However, to the best of our knowledge, there is no overview or examination of the psychometric properties of these VR tools or how they are being evaluated.

Typically, novel measures and assessments are validated against current gold-standard tasks for concurrent validity [52]. Concurrent validity can be a reliable means of determining whether two assessments measure the same construct. However, concurrent validity can also occur when two tests contain the same problems, such as inaccurately measuring a particular construct in the same way. Sequentially, many VR tasks are being created from a “function-led” perspective but validated against “construct-led” tasks [53,54]. Given their different approaches, function-led and construct-led assessments should be validated in different ways or at least using several validation approaches. If function-led VR assessments improve upon the validity of current assessment methods, validation techniques may also need to go beyond comparisons with traditional assessments. For example, function-led VR assessments may be better validated against additional alternative methods, such as carer reports, real-life performance (eg, self-care, residence, transportation, and employment), and diagnostic trajectory [49] as opposed to validation through traditional (construct-led) assessment. Without incorporating tests of ecological validity, the potential advantages of VR may go unrecognized. Given the increasingly rapid development of VR neuropsychological assessments, it will be imperative to maintain high validation standards for these tools [55].

Establishing the reliability of novel VR EF assessments is also critical to the integrity of the outcomes. Reliability ensures that the measure yields consistent and repeatable results, a foundational element for test validity. Consequently, both reliability and validity ought to be evaluated for each measurement tool. Test-retest reliability, confirming consistency over time, should be accompanied by the interval between assessments and the correlation of the results. Internal consistency, typically measured using the Cronbach α, should also be reported for each target construct or domain of assessment. Importantly, for immersive VR EF assessments that evaluate multiple EF constructs, it is essential to report the α for each distinct construct rather than a collective coefficient. This is because the coefficient is intended to evaluate item consistency within a scale measuring a single construct; applying it across disparate constructs could be confusing and potentially misleading.

Consistency of Terminology

Finally, to ensure psychometric precision and build on previous research, EF assessment paradigms must adopt consistent terminology for their target assessment constructs. The field of EF, although of significant interest to both researchers and clinicians, is marked by varied terminology for identical constructs. This issue, longstanding in EF research (see the study by Suchy [5]; for a review, see the study by Baggetta and Alexander [56]), presents challenges to VR in the EF assessment field. For instance, inconsistent terminology hinders the synthesis of research findings. Diverse labels such as "impulsivity" and "impulse control" might, upon examination, refer to the same underlying construct. Consequently, researchers aiming to extend the literature on "impulsivity" might overlook pertinent studies or exclude valuable references because of terminological discrepancies.

This literature review sought to examine and discuss the development of the VR tools used to assess EF with a specific focus on evaluating their psychometric properties. The studies selected for inclusion in this review were those that developed assessment tools for EF either holistically or in part. The aims of this review were to (1) determine the components of EF assessed using VR paradigms, (2) investigate the methods used to validate VR assessments, and (3) explore the frequency and efficacy of reporting participants' immersion in and engagement with VR for EF assessment.

Methods

Our review methodology followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [57]. In line with the literature, EF was defined as a set of executive functions, including inhibition, cognitive flexibility, and working memory [2,4,6], that support other higher-order executive functions, such as reasoning, planning, and problem-solving [6,8].

Inclusion Criteria

Before conducting the literature search, the inclusion criteria were established. First, only peer-reviewed articles and conference proceedings (complete manuscripts) written in English would be included. Second, articles that detailed an empirical, clinical, or proof-of-concept study in which an immersive virtual environment (ie, using a head-mounted display, not a 2D computer screen) was reported to broadly investigate EF or higher-order cognition or that examined EF via a selection of one or more subconstructs (eg, inhibitory control and working memory) would be included. Finally, only articles with an adult participant population published after 2013 would be included. This temporal limit was based on the release date of the Oculus Rift Development Kit 1 as it was one of the first accessible products for public use of VR. Articles were identified through the EBSCOhost, Scopus, and Web of Science (WoS) citation databases. Scopus and WoS were chosen because of their prominence as citation databases [58]. To compensate for the bias toward engineering and natural science articles found through Scopus and WoS [59], EBSCOhost was searched for articles published in fields such as (clinical) psychology and medicine.

Search Strategy

Keywords were developed by identifying 3 main components that the relevant literature should include. The 3 components were based on "Virtual Reality," "Neuropsychological Tests," and "Executive Function." It was decided not to search for specific components of EF because of the lack of consensus in the field regarding its components. Rather, it was assumed that, if an article addressed EF or a component of EF, it would include "executive functioning" as a keyword in the title, abstract, or keywords. Other reviews looking broadly at VR paradigms have used similar search strategies [49].

In this study, key terms were developed by identifying synonyms for key components and concatenating them using the "AND" Boolean operator. The final keywords used for the search were as follows: (["virtual" OR "artificial" OR "simulated"] AND ["realit*" OR "world" OR "environment"]) AND ([neuropsych* OR function* OR cognit*] AND [(executive AND function*) OR (high* AND order AND cognit*)] AND [assessment]).

Literature queries made through EBSCOhost were limited to the following databases: Academic Search Complete, AgeLine, AMED, Applied Science and Technology Source, CINAHL, E-Journals, Health Source Consumer and Nursing/Academic Edition, MEDLINE, Mental Measurements Yearbook, Psychology and Behavioral Sciences Collection, and all variations of the American Psychological Association databases. Furthermore, for the search, 3 topic fields (ie, title, abstract, and subject terms) were used to paste the keywords. The 3 topic fields were concatenated using the "OR" Boolean operator. Using the Scopus database, we implemented a basic search in the article title, abstract, or keywords using the keywords. No additional limitations were applied. Our search in WoS included all databases, and the advanced search method was used wherein keyword searches in the article title, abstract, and keyword topic fields were concatenated using the "OR" Boolean operator (ie, Title=(keywords) OR Abstract=(keywords) OR Keywords=(keywords)).

The results for each database were exported to Covidence systematic review software (Veritas Health Information) [60], which removed duplicates. All abstracts were screened independently by the first author and the senior author to determine whether the contents met the inclusion criteria. Full-text screening was also performed by the same authors. Any disagreement was discussed by the first (RK), second (LK), and senior (KR) authors.

Data Extraction

The first and second authors completed the data extraction process by manually reviewing each manuscript; data items (see the following section) were recorded in a tabular format using Microsoft Excel (Microsoft Corp).

Data Items and Synthesis

Demographic details, qualitative descriptions of the VR paradigm, user experience, cybersickness, immersion and engagement details, and comparative measures for validation purposes were extracted (Multimedia Appendix 1 [53-55,61-76]).

A qualitative evaluation of the studies included in the review was performed, meaning that the content of each manuscript was assessed based on the reported target constructs or constructs relevant to EF and the extent to which the reported VR task was related to the assessment of the target construct or constructs. To do this, studies were categorized based on the construct they targeted through their VR paradigm as reported by the authors of the respective articles. If multiple constructs were assessed in a single study, the study was included for each construct. No inferences were made about which cognitive construct or constructs was assessed based on the tasks that were reported in the manuscripts. For example, if an article indicated only that they used a VR version of the Stroop test (ST) but did not disclose which construct it assessed using this test, the study was not categorized under inhibitory control or cognitive flexibility but under the general factor "executive functioning."

Next, it was indicated whether the articles explicitly or implicitly disclosed the way in which the comparative measures (such as particular metrics) were used to validate the VR paradigm. For instance, if the article directly stated a priori that they hypothesized a correlation between a VR task measuring inhibition and a validation task such as the ST, this was recognized as providing explicit validation for inhibition. Conversely, if an article indicated that participants completed the ST, which assessed inhibition and processing speed, and mentioned that the VR paradigm evaluated inhibition, it was considered to provide implicit validation for inhibition. Furthermore, traditional construct- and function-led assessments were identified from the text.

The (quantitative) results of the studies were screened to identify (1) the direction and strength of the relationship between traditional and VR assessments and (2) whether the results from all possible and a priori–defined comparisons were reported.

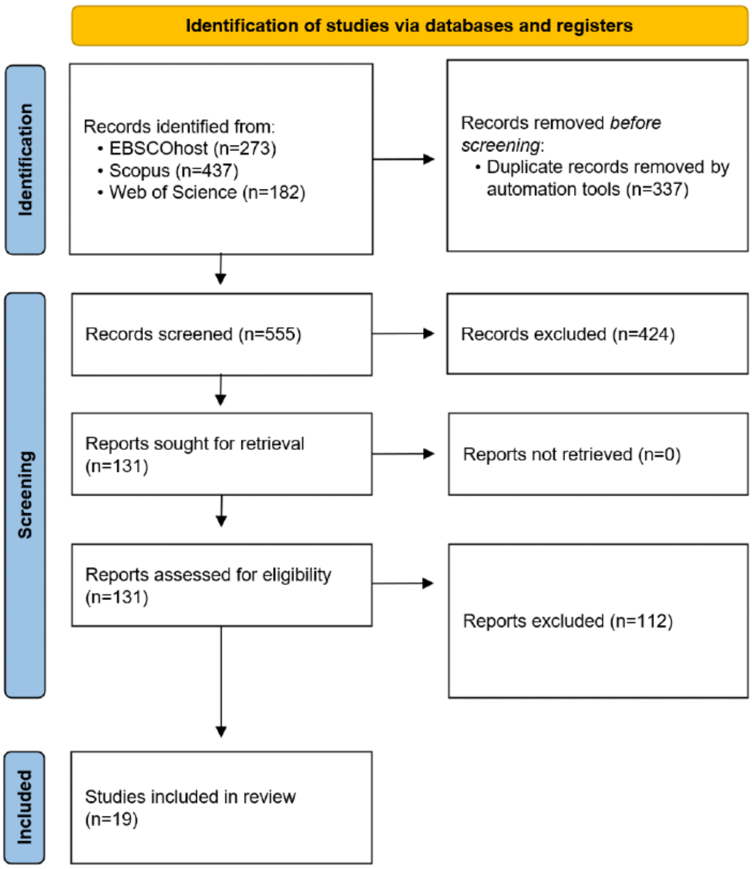
Finally, qualitative and quantitative tools used to evaluate beneficial and adverse effects of VR immersion were identified from the manuscripts and categorized in a tabulated format. The results of the studies were screened to identify whether they assessed the influence of the beneficial and adverse effects of VR immersion on task performance.

Results

Overview

Through WoS, EBSCOhost, and Scopus, 892 items were identified, from which the Covidence systematic review management platform [60] filtered 337 (37.8%) duplicates. A total of 555 unique articles remained, of which 424 (76.4%) were deemed irrelevant through abstract screening. The final 131 articles had their full texts screened, and 19 (14.5%) met the inclusion criteria. The systematic literature search process is shown in Figure 1.

Figure 1. Systematic review process and results from literature searches in EBSCOhost, Scopus, and Web of Science databases.



General EF

In total, 7 of the 19 (37%) of the reviewed studies assessed EF in general, meaning that the authors of these articles did not explicitly state which subconstruct of EF was targeted using the VR task. Table 1 shows which validation tasks were used in each study to measure EF.

Table 1. The validation tasks, authors, and total number of studies examining general executive functioning.

VR[a] target construct and validation task	Validation	Authors	Studies examining the construct, n (%)
Executive functioning: general			7 (37)
• D-KEFS[b] [77] • TMT-A[c] and TMT-B[d] • ST[e] • Modified version of the SET[f] • HTT[g] • ZMT[h]	Implicit	Banville et al [61][i]	
• ST • TMT-A • TMT-B	Implicit	Davison et al [62][j]	
• TMT-B • OTS[k] CANTAB[l] • VFT[m]	Explicit	Miskowiak et al [63]	
• TMT-A • TMT-B	Explicit	Pallavicini et al [64]	
• Groton Maze Learning Test (Cogstate)	Implicit	Porffy et al [65]	
• None specifically reported	N/A[n]	Tan et al [66]	
• None specifically reported	N/A	Tsai et al [67]	

[a]VR: virtual reality.

[b]D-KEFS: Delis-Kaplan Executive Function System.

[c]TMT-A: Trail-Making Test version A.

[d]TMT-B: Trail-Making Test version B.

[e]ST: Stroop test.

[f]SET: Six Elements Test.

[g]HTT: Tower of Hanoi test.

[h]ZMT: Zoo Map Test.

[i]The VR task was predominantly a sorting task for executive functioning assessment. The comparative assessments that validated this assessment were detailed under "executive function" broadly as the paper did not specify which components of the VR task the comparative tasks aimed to validate.

[j]The VR task was reported to assess executive functioning. The comparative assessments that validated this assessment were detailed under "executive function" broadly as the paper did not specify which components of the VR task the comparative tasks aimed to validate.

[k]OTS: One Touch Stockings of Cambridge.

[l]CANTAB: Cambridge Neuropsychological Test Automated Battery.

[m]VFT: verbal fluency test.

[n]N/A: not applicable.

Banville et al [61] immersed participants in a Virtual Multitasking Test (VMT), which was in principle designed to measure prospective memory and executive functions by having participants perform multiple tasks in a virtual apartment. However, this paper reported specifically on the task in which participants had to store groceries as fast as possible while also being attentive to other tasks, such as answering the phone or closing a window. Although the authors hypothesized that VMT scores would be correlated with neuropsychological assessments, such as mental flexibility, planning, and inhibition, it was not explicitly stated which metric of the VMT would be correlated with which neuropsychological assessment. Nonetheless, the authors identified that grocery storing time was correlated with the rule-break score on the Six Elements Test (r_{19}=−0.49; *P*=.04; *P* value as reported in the manuscript). Furthermore, the number of errors in storing fruits and vegetables was found to correlate with the perseveration score on the Zoo Map Test (r_{20}=0.53; *P*=.02; *P* value as reported in the manuscript) and reading speed during the second condition of the ST (r_{20}=0.44; *P*=.05; *P* value as reported in the manuscript).

Davison et al [62] immersed participants in a parking simulator and a chemistry laboratory where they had to park a vehicle, sort chairs, or locate items. Before immersion, participants completed the ST and the TMT versions A (TMT-A) and B (TMT-B). The authors identified that the completion time of the second level (Kendall τ=−0.32; *P*=.01; *P* value as reported in manuscript) and the number of levels completed in the parking simulator (τ=0.43; *P*<.01; *P* value as reported in manuscript) were correlated with participants' performance on the ST. In addition, the ST was correlated with seating arrangement metrics, such as time to place the first stool (τ=−0.33; *P*=.01; *P* value as reported in manuscript) and number of stools placed (τ=0.33; *P*=.02; *P* value as reported in manuscript), as well as with time to locate the first item in the chemistry laboratory (τ=−0.37; *P*=.01; *P* value as reported in manuscript). Correlations between the TMT-A or TMT-B and, for example, the number of completed parking levels (τ=−0.49; *P*<.01; *P* value as reported in the manuscript) or the number of items placed in the seating arrangement task in the chemistry laboratory (τ=−0.35; *P*=.01; *P* value as reported in the manuscript) were reported. However, reporting was limited to significant correlations only, and no a priori expectation of how performances on the VR and validation tasks were correlated was indicated in the study.

Miskowiak et al [63] assessed executive functions by letting participants complete the TMT-B, One Touch Stockings of Cambridge mean choices to correct, and verbal fluency test versions S and D. The performance on these tests was compared with participants' performance on a cooking task in VR. The authors hypothesized that the number of cooking tasks that were correctly placed on a to-do list and the latency to solve the task would be VR-equivalent measures of EF. The authors found that VR performance was correlated (r_{121}=0.26; *P*=.004) with EF, which consisted of a correlation between the average performance on the VR subtasks and the average performance on the validation tasks. The correlations between the individual performances on the VR and validation tasks were not reported in the manuscript.

Pallavicini et al [64] had participants play the Audioshield dance game, which the authors hypothesized could be closely related to EF constructs such as inhibition and working memory. However, the authors correlated participants' performance on the Audioshield game with their performance on the TMT-A and TMT-B, which measure psychomotor speed (TMT-A) and mental flexibility (TMT-B). Nonetheless, the results showed that TMT performance was negatively correlated with Audioshield performance metrics.

Porffy et al [65] had participants complete VStore, where the 2 tasks measured EF, namely the "Find" task and the "Coffee" task. Specifically, participants had to find 12 items from a list they had previously memorized. In addition, participants had to order a hot drink from the coffee shop after finding, bagging, and paying for the 12 remembered items they had found in the store. Notably, the authors indicated that the 2 VR tasks also tapped into navigation (ie, "Find" task) and processing speed (ie, "Coffee" task). Furthermore, the Groton Maze Learning Test from Cogstate, which the participants completed before the VR task, was used to evaluate general EF. Nonetheless, through their regression analysis, the authors identified that the Groton Maze Learning Test was not a predictor for the "Find" task (*B*=0.024; SE 0.029; *P*=.11; *P* value as reported in the manuscript) or the "Coffee" task (*B*=−0.003; SE 0.051; *P*=.96; *P* value as reported in the manuscript).

Tan et al [66] had 100 participants complete 13 tasks in a virtual environment that were designed to measure 6 cognitive domains, such as EF and complex attention. Although differences in performance on VR tasks related to EF between age groups were found, no comparison was made with a traditional neuropsychological assessment of EF or any subconstructs of EF.

Tsai et al [67] immersed 2 participant groups in a virtual shopping environment: one group with mild cognitive impairment (MCI) and one control group. The VR tasks assessed participants' memory, EF, and calculation by having them memorize a shopping list, search for the listed items in the shop, and subsequently pay for them. The authors trained machine learning models on features extracted from the VR tasks to predict whether participants had MCI or were healthy controls, which was achieved with high accuracy. Nonetheless, no neuropsychological assessment of EF was reported as a validation for the VR tasks.

Targeted Constructs

The following subsections elaborate on the EF constructs and subconstructs addressed in the studies under review. A range of correlation coefficients were reported in these papers; however, because of the lack of uniformity in results reporting, these coefficients were omitted from the current synthesis. Typically, the papers reported only significant correlations between metrics without presenting all potential correlations. Furthermore, only 16% (3/19) of the studies specified an α level (ie, .05), with another 16% (3/19) of the studies indicating statistical significance at a *P* value of ≤.05. A total of 21% (4/19) of the studies did not indicate an α level but mentioned applying corrections for multiple comparisons, yet they did not detail the adjusted α level. In total, 5% (1/19) of the studies adopted Bayesian statistics using a Bayesian factor of >10 for statistical inference. Nonetheless, in the reviewed studies, it was not consistently clarified which VR tasks were validated against traditional tasks, hindering the construct validity of the various EF components. Consequently, drawing consistent conclusions on how EF constructs of subconstructs were evaluated was not feasible without inferring the nature of the tests and assessment paradigms.

Core Executive Functions

Inhibition

Of the 3 "core" executive functions, 37% (7/19) of the studies included in our review investigated inhibitory control, interference control, or impulsivity either singly or combined. Table 2 details the respective validation tasks and target constructs of each of these studies. For example, Chicchi Giglioli et al [68] presented participants with 6 standardized tasks, 3 of which assessed inhibition (Table 2), before administering a serious game in which participants were required to perform tasks in outer space. In total, 10 of the 36 possible correlations between measures for the standardized tasks and the serious game tasks were reported as statistically significant and ranged from weak (0.20<*r*<0.39; relative *P* values indicated in the manuscript, eg, *P*<.05) to strong (0.60<*r*<0.79; relative *P* values indicated in the manuscript). For example, the latency metric of the dot-probe task (DPT) correlated positively (0.35<*r*<0.54; relative *P* values indicated) with the latency metric of the 3 VR tasks aimed at measuring inhibition, whereas no correlations were reported between the correct answer metric of the DPT and the correct answer metric of the 3 VR tasks aimed at measuring inhibition. None of the metrics from the ST correlated with those of the VR task (requiring participants to fight aliens); however, the correct answer and latency metrics of the ST correlated with those of the VR task (requiring participants to repair a valve).

Table 2. The validation tasks, authors, and total number of studies examining each construct.

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
Inhibition or Inhibitory control			6 (32)
• DPT[b] • GNG[c] • ST[d]	Implicit	Chicchi Giglioli et al [69]	
• DPT • GNG • ST	Explicit	Chicchi Giglioli et al [68]	
• GNG	Implicit	Marín-Morales et al [70]	
• CPT[e]	Implicit	Voinescu et al [71][f]	
• None specifically reported	N/A[g]	Parsons and Carlew [72]	
• ST	Implicit	Parsons and Barnett [73]	
Interference control			3 (16)
• ST	Implicit	Marín-Morales et al [70][h]	
• The CW-IT[i] from the D-KEFS[j] • Automated neuropsychological assessment metrics • ST	Implicit	Parsons and Carlew [72]	
• CW-IT from the D-KEFS	Implicit	Parsons and Barnett [73]	
Impulsivity			1 (5)
• None specifically reported	N/A	Chicchi Giglioli et al [68]	

[a]VR: virtual reality.

[b]DPT: dot-probe task.

[c]GNG: Go/No-Go.

[d]ST: Stroop test.

[e]CPT: continuous performance test.

[f]Some traditional tasks listed were included for divergent validity and, therefore, have been omitted from this table.

[g]N/A: not applicable.

[h]The VR task involved 42 VR mini-games that assessed various cognitive constructs. A total of 4 mini-games and their target constructs were documented and included in this table; however, the comparative assessments were not provided, and an extensive list of all 42 mini-games was not provided.

[i]CW-IT: Color-Word Interference Test.

[j]D-KEFS: Delis-Kaplan Executive Function System.

Similarly, Chicchi Giglioli et al [69] immersed participants in a virtual kitchen in which they had to cook different types of food. The activities were grouped into 4 subtasks of incremental difficulty where, in the third level, inhibition was assessed by determining whether the right dressing was added using a Go/No-Go (GNG)–type paradigm. The authors stated that the DPT, GNG, and ST were used as standard tasks to assess inhibition. The unspecified metric of “correct dressing” was shown to correlate well (*r*=0.527; *P*<.01; relative *P* value indicated in the manuscript) with the correct answer metric of the ST in one group, whereas in the second group, a moderate negative correlation (r=−0.486; P≤.05; relative P value indicated in the manuscript) was found between the execution time of the Tower of London task and the correct dressing metric. However, no other correlations between the VR task metric and those of the traditional assessments of inhibition were reported.

Marín-Morales et al [70] had participants complete neuropsychological assessments, including the GNG task, as well as 42 mini-games in VR. An undisclosed set of variables from the mini-games was used as predictors for measures of neuropsychological batteries. The mini-game predictor variables were fed into different machine learning algorithms. The authors highlighted that games related to inhibition produced worse results compared with other games but did not report any results on inhibition. The authors did find that mini-game features of planning and attention could predict GNG hit proportions and mean time with 80% and 94% accuracy, respectively.

Parsons and Carlew [72] had participants perform the ST in a virtual classroom as well as complete a computerized and paper-and-pencil version of the task. The authors found that participants' performance was lower for color naming and word reading in the VR paradigm than in the paper-and-pencil version but interference performance was better in the VR paradigm than in the paper-and-pencil version. Similarly, Parsons and Barnett [73] had participants perform the ST in a virtual apartment as well as complete a computerized and paper-and-pencil version of the task. Here, the authors found that participants were more accurate in the ST in the paper-and-pencil version than in the VR paradigm.

Voinescu et al [71] immersed participants in a virtual aquarium where they had to perform a variety of tasks. For example, participants had to respond when they saw a fish that was different from a clown fish or heard a fish name different from surgeonfish. After the VR aquarium, participants completed a variety of computerized tasks, among them a continuous performance test (CPT), which was hypothesized to measure sustained attention and inhibition. The authors found weak to moderate (0.22<r<0.49; relative P values indicated, eg, P<.05) correlations between CPT measures and VR measures.

Working Memory

Working memory was investigated in 21% (4/19) of the studies [63,65,70,74]. Table 3 details the respective validation tasks and target constructs of each of these studies. The working memory component from the study by Marín-Morales et al [70] included a mini-game wherein participants had to recall the ingredients of a recipe seen before the mini-game and collect from a range of options only those ingredients found in the recipe. However, no correlations with neuropsychological tasks were presented. Miskowiak et al [63] compared their VR paradigm with a traditional task that assessed working memory. In this study, participants were instructed to plan and cook a meal in a virtual kitchen. Performance metrics, such as the number of drawers opened and the latency until the task was completed, were used to assess working memory and were correlated with metrics from traditional tasks such as the Wechsler Adult Intelligence Scale Letter-Number Sequencing. The authors reported a significant positive correlation (r_{121}=0.31; P=.001) between the VR task metrics and the traditional task metrics that evaluated working memory.

Table 3. The validation tasks, authors, and total number of studies targeting working memory.

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
Working memory			4 (21)
• WAIS-IV[b] • The Working Memory Index (Digit Span and Arithmetic)	Implicit	Marín-Morales et al [70][c]	
• WAIS-III[d] LNS[e] • SWM[f] CANTAB[g] (error and strategy)	Explicit	Miskowiak et al [63]	
• 1-back and 2-back test (Cogstate)	Implicit	Porffy et al [65]	
• None specifically reported	N/A[h]	Robitaille et al [74][i]	

[a]VR: virtual reality.

[b]WAIS-IV: Wechsler Adult Intelligence Scale–IV.

[c]The VR task involved 42 VR mini-games that assessed various cognitive constructs. In total, 4 mini-games and their target constructs were documented and included in this table; however, the comparative assessments were not provided, and an extensive list of all 42 mini-games was not provided.

[d]WAIS-III: Wechsler Adult Intelligence Scale–III.

[e]LNS: Letter-Number Sequencing.

[f]SWM: Spatial Working Memory.

[g]CANTAB: Cambridge Neuropsychological Test Automated Battery.

[h]N/A: not applicable.

[i]Robitaille et al [74] used a VR paradigm with avatars to trial a dual-task walking protocol.

Porffy et al [65] asked participants to operate a virtual store in which the working memory component was assessed at the "Pay" step, where participants had to select and pay for their items at a self-checkout machine providing the exact amount.

The authors specified that the reaction time on the 1-back task and the accuracy of performance on the 2-back task were metrics from traditional tasks used to assess working memory. Using linear regression, the authors found that performance on the 2-back task was negatively associated (B=−0.085; SE 0.042; P=.047) with participants' performance on the "Pay" step.

Robitaille et al [74] assessed working memory during their simultaneous cognitive tasks, in which participants had to both recognize faces in windows that had been previously declared as "hostile" or "nonhostile" and complete a navigation task. However, no correlations between the traditional and VR tasks were reported.

Cognitive Flexibility

One study by Chicchi Giglioli et al [68] investigated cognitive flexibility (termed "cognitive shifting" in the paper) through 3 VR tasks. The authors specified that the TMT was used as a traditional task to assess cognitive flexibility as a comparator for the first VR task (CF1, cultivating food) and the Wisconsin Card Sorting Test was used as a traditional task to evaluate cognitive flexibility as a comparator for the other 2 VR tasks (CF2, growing plants, and CF3, fueling a turbine). The total time metric of the first VR task correlated positively with the total time of the TMT-B (r=0.396; P<.01; P value as reported in the manuscript), and multiple metrics of VR tasks 2 and 3 correlated with the performance metrics of the Wisconsin Card Sorting Test.

Higher-Order Executive Functions: Planning

In total, 26% (5/19) of the studies [62,68,69,75,76] identified planning as a target construct in their VR paradigms. Table 4 details the respective validation tasks and target constructs of each of these studies. The VR environment created by Chicchi Giglioli et al [69] used a cooking task with 4 levels of difficulty. In the 3 more difficult levels, planning was required to complete the tasks as 2 burners were used. There was no clearly specified metric for the VR task that was used to evaluate planning, but the authors specified that the Tower of London task was used as a traditional assessment to evaluate planning. A variety of VR task metrics, such as total time to complete a difficulty level, were shown to correlate with various Tower of London task metrics.

Table 4. The validation tasks, authors, and total number of studies targeting planning.

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
Planning			5 (26)
TOL-DX[b]	Implicit	Chicchi Giglioli et al [69]	
TOL[c]	Explicit	Chicchi Giglioli et al [68]	
None specifically reported	N/A[d]	Davison et al [62][e]	
The Key Search task from BADS[f] [78]	Explicit	Kourtesis et al [76]	
None specifically reported	N/A	Kourtesis and MacPherson [75]	

[a]VR: virtual reality.

[b]TOL-DX: Tower of London–Drexel test.

[c]TOL: Tower of London test.

[d]N/A: not applicable.

[e]The VR task was used to assess executive function. The comparative assessments that validated this assessment were detailed under "executive function" broadly as the paper did not specify which components of the VR task the comparative tasks aimed to validate.

[f]BADS: Behavioral Assessment of the Dysexecutive Syndrome.

In another study, Chicchi Giglioli et al [68] used a VR paradigm based on an outer-space environment. The paradigm contained 8 tasks, one of which assessed planning ability (task 7). The authors stated that the Tower of London task was the traditional assessment tool used to evaluate planning and explained that the total score, initial time, and execution time of the VR task were the outcome metrics. Moderate positive correlations were found between the execution time of the VR task and of the Tower of London task (r=0.463; P<.01; P value as reported in the manuscript) and between the initial time of the VR task and the total time of the Tower of London task (r=0.372; P<.05). Furthermore, the VR task correlated with some metrics of other traditional assessments used to assess planning ability, although these were not specified a priori.

Both the studies by Kourtesis et al [76] and Kourtesis and MacPherson [75] used the same VR environment based on a variety of everyday tasks. One task assessing planning ability required participants to draw their route around the city (eg, visiting the bakery, supermarket, and library and returning home) on a 3D board. Kourtesis et al [76] explained that the Key Search Test from the Behavioral Assessment of the Dysexecutive Syndrome was used as a traditional measure to assess planning and found a strong positive correlation between the traditional and VR tasks (r=0.80; Bayes factor=4.65 × 10^8). Furthermore, Kourtesis and MacPherson [75] noted in their results that planning explained a substantial 12% (P=.03) of the variance in time-based prospective memory, which was required in 10 of 17 tasks.

Davison et al [62] assessed planning ability using a task involving the arrangement of a table and a chair. However, they did not explicitly mention the traditional task that was used to evaluate planning. Various correlations between the performance metrics of the VR task and the traditional task were reported.

For example, the performance on the Stroop Color and Word Test was negatively correlated with the time participants took to place a blue chair in the seating arrangement task (Kendall τ=−0.39; *P*=.01; *P* value as reported in the manuscript).

Other Domains

Several studies (14/19, 74%) examined domains of functioning that did not align with the EF definition used in this review. Broadly, these domains fell under the categories of memory, attention, processing, task performance, and a variety of other uncategorized subconstructs. As the literature [1,2,4,6] does not relate these broad domains to EF, they are not discussed further but are presented in Tables 5-6.

Table 5. The validation tasks, authors, and total number of studies targeting constructs classified as uncategorized.

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
Memory			11 (58)
Memory (general)			1 (5)
• None specifically reported	N/A[b]	Tsai et al [67]	
Verbal memory and verbal learning			2 (11)
• RAVLT[c] subtests: total, immediate recall, delayed recall, and recognition	Explicit	Miskowiak et al [63]	
• International Shopping List Test (Cogstate; verbal learning)	Implicit	Porffy et al [65]	
Prospective memory			4 (21)
• None specifically reported	N/A	Banville et al [61][d]	
• CAMPROMPT[e] [79]	Explicit	Kourtesis et al [76][f]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	
• CVLT-II[g] [80]	Implicit	Parsons and McMahan [53]	
Episodic memory			3 (16)
• RBMT-III[h] [81]	Explicit	Kourtesis et al [76][f]	
• CVLT-II	Implicit	Parsons and McMahan [53]	
Immediate recognition			2 (11)
• RBMT-III [81]	Explicit	Kourtesis et al [76]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	
Delayed recognition			2 (11)
• RBMT-III [81]	Explicit	Kourtesis et al [76][f]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	
Attention			13 (68)
General attention			4 (21)
• DPT[i] • GNG[j] • ST[k]	Implicit	Chicchi Giglioli et al [69]	
• DPT • GNG • ST • TMT-A[l] • TMT-B[m]	Explicit	Chicchi Giglioli et al [68]	
• DPT—selective attention • GNG—sustained attention • ST—selective attention • TMT[n]—visual attention	Implicit	Marín-Morales et al [70][o]	
• RVP[p] CANTAB[q] (accuracy and latency) • RBANS-DS[r]	Explicit	Miskowiak et al [63]	
Divided attention			2 (11)

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
• None specifically reported	N/A	Robitaille et al [74][s]	
• CTT-B[t] [75,82]	Explicit	Wilf et al [54]	
Complex attention			1 (5)
• None specifically reported	N/A	Tan et al [66]	
Selective visual attention			2 (11)
• The map task from the Test of Everyday Attention	Explicit	Kourtesis et al [76][f]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	
Selective auditory attention			2 (11)
• The Elevator Counting With Distraction task of the Test of Everyday Attention	Explicit	Kourtesis et al [76][f]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	
Sustained visual attention			1 (5)
• CTT-A[u] [82]	Explicit	Wilf et al [54]	
Visuospatial attention			2 (11)
• The Ruff 2 and 7 Selective Attention Test	Explicit	Kourtesis et al [76][f]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	
Sustained attention			1 (5)
• CPT[v] [83]	Implicit	Voinescu et al [71]	
Processing			3 (16)
Processing speed			3 (16)
• WAIS-IV[w] Processing Speed Index (symbol search and coding)	Implicit	Marín-Morales et al [70][o]	
• RBANS-CT[x] • TMT-A	Explicit	Miskowiak et al [63]	
• Detection task (Cogstate)	Implicit	Porffy et al [65]	
Task performance			4 (21)
Dual task			1 (5)
• TMT-A • TMT-B	Implicit	Chicchi Giglioli et al [69]	
Multitask			3 (16)

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
• Modified version of the SET[y]	Implicit	Banville et al [61][d]	
• CTT[z] [82]	Explicit	Kourtesis et al [76][f]	
• None specifically reported	N/A	Kourtesis and MacPherson [75]	

[a]VR: virtual reality.

[b]N/A: not applicable.

[c]RAVLT: Rey Auditory Verbal Learning Test.

[d]The VR task was predominantly a sorting task for executive function assessment. The comparative assessments that validated this assessment were detailed under "executive function" broadly as the paper did not specify which components of the VR task the comparative tasks aimed to validate.

[e]CAMPROMPT: Cambridge Prospective Memory Test.

[f]Kourtesis et al [76] explicitly broke episodic memory down into immediate and delayed recognition. However, we gathered these two constructs under episodic memory.

[g]CVLT-II: California Verbal Learning Test–Second Edition.

[h]RBMT-III: Rivermead Behavioral Memory Test–Third Edition.

[i]DPT: dot-probe task.

[j]GNG: Go/No-Go.

[k]ST: Stroop test.

[l]TMT-A: Trail-Making Test version A.

[m]TMT-B: Trail-Making Test version B.

[n]TMT: Trail-Making Test.

[o]The VR task involved 42 VR mini-games that assessed various cognitive constructs. In total, 4 mini-games and their target constructs were documented and included in this table; however, the comparative assessments were not provided, and an extensive list of all 42 mini-games was not provided.

[p]RVP: Rapid Visual Information Processing.

[q]CANTAB: Cambridge Neuropsychological Test Automated Battery.

[r]RBANS-DS: Repeatable Battery for the Assessment of Neuropsychological Status–Digit Span.

[s]Robitaille et al [74] used a VR paradigm with avatars to trial a dual-task walking protocol.

[t]CTT-B: Color Trails Test B.

[u]CTT-A: Color Trails Test A.

[v]CPT: continuous performance test.

[w]WAIS-IV: Wechsler Adult Intelligence Scale–IV.

[x]RBANS-CT: Repeatable Battery for the Assessment of Neuropsychological Status–Coding Test.

[y]SET: Six Elements Test.

[z]CTT: Color Trails Test.

Table 6. The validation tasks, authors, and total number of studies targeting constructs classified as uncategorized.

VR[a] target construct and validation task	Validation	Authors	Studies, n (%)
Uncategorized[b]			12 (63)
Visual perception			1 (5)
• None specifically reported	N/A[c]	Marín-Morales et al [70][d]	
Verbal learning			2 (11)
• RAVLT[e] subtests: total, immediate recall, delayed recall, and recognition	Explicit	Miskowiak et al [63]	
• International Shopping List Test (Cogstate)	Implicit	Porffy et al [65]	
Navigation			2 (11)
• None specifically reported	N/A	Porffy et al [65]	
• None specifically reported	N/A	Robitaille et al [74]	
Associate learning			1 (5)
• Continuous Paired Associate Learning Test (Cogstate)	Implicit	Porffy et al [65]	
Pattern recognition			1 (5)
• Continuous Paired Associate Learning Test (Cogstate)	Implicit	Porffy et al [65]	
Perceptual motor			1 (5)
• None specifically reported	N/A	Tan et al [66]	
Social cognition			1 (5)
• None specifically reported	N/A	Tan et al [66]	
Learning and memory			1 (5)
• None specifically reported	N/A	Tan et al [66]	
Language			1 (5)
• None specifically reported	N/A	Tan et al [66]	
Calculation			1 (5)
• None specifically reported	N/A	Tsai et al [67]	

[a]VR: virtual reality.

[b]Williams et al [55] replicated the Wisconsin Card Sorting Test and multitasking task but did not explicitly state the cognitive constructs that the VR task was assessing. For this reason, the paper has not been assigned a target construct.

[c]N/A: not applicable.

[d]The VR task involved 42 VR mini-games that assessed various cognitive constructs. In total, 4 mini-games and their target constructs were documented and included in this table; however, the comparative assessments were not provided, and an extensive list of all 42 mini-games was not provided.

[e]RAVLT: Rey Auditory Verbal Learning Test.

Validity and Reliability

Tables 1-6 show details of the current validated comparator tasks against the novel VR tasks if they were explicitly provided by the authors. Where "None specifically reported" is stated, the authors of each paper did not identify or indicate a direct comparator. All but 2 studies (17/19, 89%) [72,73] set out to assess multiple constructs. In some cases, the subconstructs that were assessed were individually validated against existing validated tasks. In other cases, a suite of existing validated tasks was included in the analysis for correlation against a variety of subconstructs being assessed using the VR battery. In these cases, there was no validation at the construct level identified a priori. In 16% (3/19) of the studies, there was no reported validation of the VR paradigm.

Notably, only one study used real-life validation criteria in addition to construct-driven tests to present a validation of their VR paradigm. Specifically, Miskowiak et al [63] functionally assessed participants using the Functioning Assessment Short Test (FAST) and the brief University of California, San Diego, Performance-Based Skills Assessment (UPSA-B). Participants' scores on these assessments were correlated with their performance on the test domains of the VR paradigm, called cognition assessment in VR (CAVIR). The authors identified that participants' performance on the FAST was negatively associated (−0.17<r<−0.30; no exact or relative P values reported) with CAVIR test domains such as processing speed and working memory, whereas participants' performance on the UPSA-B was positively associated with the CAVIR test working memory (r=0.40; P value not exactly or relatively reported) and cognition composite (r_{68}=0.44; P<.001) domains. Moreover, the authors noted that lower global scores on traditional (ie, construct-led) neuropsychological tests were negatively associated with FAST scores (r_{121}=−0.45; P<.001) and positively associated with UPSA-B scores (r_{68}=0.52; P<.001), highlighting that lower CAVIR scores were associated with more functional disability, as indicated by the functional and traditional assessment tools.

The reliability of the VR paradigm was only assessed in 5% (1/19) of the studies. This was done by Kourtesis et al [76], who reported good internal reliability (Cronbach =.79) of their VR Everyday Assessment Lab (EAL) paradigm. However, this global internal consistency report did not provide a reliability estimate of the unique cognitive functions targeted by their VR EAL paradigm. Nonetheless, none of the reviewed studies included a test-retest analysis to highlight the reliability of their VR paradigm.

Evaluation of User Experience, Cybersickness, Immersion, and Engagement

An overview of the measures used to evaluate participants' experiences and well-being can be found in Multimedia Appendix 1 [53-55,61-76]. Of the 19 studies, 5 (26%) included user experience assessments. To measure participants' virtual presence, experience, and well-being, the studies administered the Igroup Presence Questionnaire [61], Presence Questionnaire [63,71,74], or Slater-Usoh-Steed questionnaire [74]. To measure participants' discomfort, the studies used the Simulator Sickness Questionnaire [61,71,74] or an adaption of it, the Virtual Reality Sickness Questionnaire [63]. To evaluate the usability of the virtual environment, the studies used the System Usability Scale [71]. To measure participants' virtual experience and comfort, 11% (2/19) of the studies used the Virtual Reality Neuroscience Questionnaire [76].

Two studies (2/19, 11%) investigated whether system usability, virtual presence, or cybersickness affected participants' task performance. For example, Porffy et al [65] measured participants' technical familiarity and found that it explained between 10% and 42% of the variability in participants' performance on the VStore outcomes "Recall", "Find", and "Select". Conversely, participants' technical familiarity appeared to influence their performance on VStore. Kourtesis et al [76] used questionnaires to evaluate the quality of the VR paradigm, participants' gaming experience, and the realism (verisimilitude) and pleasantness of the VR paradigm. The authors identified no relationship between VR experience, gaming experience, and performance on the VR EAL tasks.

Some papers (4/19, 21%) reported on cybersickness, presence, or usability scores but did not report an analysis of the relationship between task performance and measures evaluating the VR paradigm. For example, Banville et al [61] recorded participants' sickness and virtual presence but did not report any test evaluating whether sickness or presence affected task performance. Similarly, Voinescu et al [71] obtained system usability ratings from participants; however, no test was reported wherein the effect of usability on task performance was assessed. Finally, Chicchi Giglioli et al [68] recorded participants' use of technology but did not report an analysis between technology use and task performance.

Finally, some studies (2/19, 10%) evaluated participants' experiences post hoc, although it was not disclosed whether any validated scales were used. For example, Davison et al [62] measured participants' enjoyment of the VR tasks and their preference for either the VR tasks or the pencil-and-paper tasks. The authors found that younger participants rather than older ones preferred VR tasks over pencil-and-paper tasks. In addition, 11 out of 40 participants reported having experienced a mild degree of motion sickness. However, 58% (11/19) of the papers did not disclose any information about user experiences.

Discussion

Overview

The purpose of this review was to investigate the development and validation of VR assessment tools for EF. Specifically, we examined the components of EF that were assessed using VR, their validation processes, and whether immersion and cybersickness assessments were used. Although research in this domain is proliferating, the results of this review suggest that the process of development and validation varies considerably between studies.

Components of EF Assessed Using VR Paradigms

Overview

The terminology used in the papers to describe EF constructs was inconsistent. For example, the most popular construct set assessed using VR comprised the inhibition processes. "Inhibitory control" encompasses the inhibition of goal-irrelevant stimuli, cognitions, and behavioral responses [6,84]. In total, two of the key components of inhibitory control are response inhibition and attentional inhibition [85]. Response inhibition was also termed "inhibition control," "prepotent response inhibition," and "motor inhibition," whereas attentional inhibition was also termed "control of interference," "interference control," and "external interference control." Although these terms are used in the literature [85], its readability and synthesis would be improved through agreement on a particular term for the same construct. In the same way, several studies (7/19, 37%) examined "EF" broadly without specifically detailing its components. In these studies, EF was validated using different measurement tools, suggesting that, across studies, EF was defined and used differently in each VR paradigm. As the constructs that these paradigms aimed to assess were not explicitly detailed, this poses a risk of hampering researchers wishing to build upon previous findings.

Furthermore, there was a broad range of constructs that were not commonly considered as EF domains but were reported as components of EF, making it difficult for future research to replicate the findings of undefined target constructs. For example, several papers (14/19, 74%) reported on verbal learning [63], associate learning pattern recognition [65], perceptual motor, social cognition, language [66], and calculation [67]. Although many of these components rely on EF domains or underpin those domains, they exist at various levels of abstraction. Thus, although the reviewed studies investigated components at different levels and used different languages, it is possible that they overlapped. For example, "organization" may be an umbrella term for a range of EF domains, each of which uses different terminology for the same concept, such as "cognitive flexibility," "flexible updating," and "working memory." Although "organization" is not measured as a higher-order version of the subcomponents, it is difficult for the research that has examined cognitive flexibility and working memory to be extended. Thus, 2 studies assessing the same construct are not able to build on each other's progress.

Recommendation: Establish a Coherent and Consistent Framework for EF Terminology

The Research Domain Criteria (RDoC) framework developed by the National Institute of Mental Health could serve as a framework to address this recommendation. The RDoC was originally created to consolidate the research conducted in various fields of mental health [86]. The framework categorizes cognition into 6 domains and encourages the investigation of these domains via different classes of variables, such as behavioral, physiological, and self-report data. This framework encourages a common language and organizes findings in such a way that researchers can identify gaps or discrepancies in the literature and contribute to the ongoing development of the field. This framework indicates the potential benefits of using a common language for research, although it is not necessarily the only option in this field. Alternatively, researchers could engage in a Delphi study to generate expert-informed consensus on the key constructs of EF that merit investigation using VR paradigms (eg, see the study by Yücel et al [87] for a Delphi study on neuropsychological assessment for addiction). Nonetheless, the emerging area of VR development for neuropsychological assessments would benefit from using the RDoC framework to coordinate the research process.

Validation of VR for EF

Overview

Overall, there was limited reporting on the constructs that were assessed using VR paradigms and the associated validation outcome measures. In some papers, there was inadequate reporting of the constructs that the VR paradigm was intended to assess. In others, the same construct was assessed using a variety of traditional tasks. Furthermore, some VR paradigms were intended to replicate real life yet were validated against traditional tasks, none of which assessed ecological validity. In some studies, the correlations between the VR paradigm and the traditional tasks were incomplete. Finally, sample sizes varied considerably between studies, also affecting the evaluation of their psychometrics. These points are expanded upon in this section.

Several studies (5/19, 26%) examined EF as a broad category and then validated the paradigm against a variety of traditional tasks. However, some studies (3/19, 16%) detailed limited (or no) reporting of which aspect of the VR paradigm each traditional task was intended to validate. That is, no details were provided regarding which traditional task outcome measure corresponded to each component of EF within the VR paradigm. Traditional tasks, which often target one construct, were then correlated against seemingly all outcomes of the VR paradigm. Although this practice may be beneficial during the exploratory phase of VR paradigm development, failure to correct for multiple comparisons may provide misleading results whereby a correlation is found between two constructs incidentally. Conversely, some traditional tasks assessed multiple constructs, which poses a slightly different challenge. For example, if the VR paradigm broadly assessed EF but was validated against the ST, it was then unclear whether the VR paradigm aimed to assess processing speed, attention, inhibitory control, or interference control as the ST could be used to measure all four. Similarly, when these studies used multiple traditional assessments, the reader was expected to presume the target constructs of the VR paradigm as this was not clearly outlined. Poorly defined target constructs and failure to specify which traditional task validates which aspect of the VR task produces a literature that is difficult to interpret. Moreover, this general lack of clarity means that future researchers are more likely to invent a new paradigm rather than adopt or extend existing paradigms, creating inefficiency and hampering progress in the field.

Various standardized tasks were used to validate target constructs in the VR paradigm. For example, the study by Chicchi Giglioli et al [69] examined attention and inhibition control using the DPT, GNG, and ST. However, Voinescu et al [71] examined inhibition using the CPT paradigm. In addition, Marín-Morales et al [70] assessed inhibition using one mini-game of their VR paradigm. However, they neither provided details of a specific comparator task for validation purposes nor reported the statistical outcomes. Furthermore, the DPT, which is typically used to assess selective attention [88], was used to assess inhibition, although its own psychometric properties have been the subject of controversy [89,90]. Although several traditional tasks purport to measure the same construct (ie, there is not one task for one construct), the lack of consistency between studies makes it difficult to compare VR platforms. Furthermore, the traditional comparator task used to validate the VR paradigm needs to have sound psychometric properties in its own right to assess the respective construct; when two tasks are compared with one another, it is unclear which task may be responsible for discrepancies in the outcome [91]. These points are especially pertinent for studies that rely solely on traditional measures to validate tasks in the absence of other validation techniques.

Although it is promising to see that VR paradigms are being used for ecologically valid assessments, their validation remains a challenge. In the case of traditional tasks, we assume that a single construct can be assessed using a behavioral task and that the performance on that task is linear with the cognitive construct. In the case of a "function-led" VR task, there is a behavioral task that simulates real-world functioning, which is thought to deteriorate in an EF-declining population. This VR task is not a direct assessment of a target construct—it is a test of a real-world function, such as parking a car. To test convergent validity, the individual would have to park a car in real life and have their performance assessed similarly to that on the VR task and compared. However, when we use traditional measures to validate the "function-led" VR measures, we assume that EF can be reliably measured and the function-led VR task (eg, parking a car) requires the same EF. Thus, those who perform poorly on a traditional EF task are also expected to perform poorly on real-life tasks requiring EF. Critically, if our results do not show this relationship, it could be that the traditional task is a poor test of EF, the function-led assessment is a poor test of EF, or the EF at hand is not related to the functional task (eg, parking a car).

These assumptions place substantial weight on the selection of the traditional task for validating the VR paradigm for predictive validity. Davison et al [62] assessed EF using the ST and TMT. They broadly hypothesized that there would be correlations between the traditional measures and the VR paradigm, which contained tasks that replicated real life, such as car parking, arranging seating, and locating items. In the reported results, the ST and TMT were correlated with all outcome measures of the VR paradigm. For example, performance on the Stroop Color and Word Test was correlated with performance on the second parking simulator task, the number of levels completed on the parking simulator task, and the time taken to place the blue chair in the seating arrangement task. If the ST and TMT are not sufficient validators of the functional task, this may generate misleading results regarding the integrity of the VR paradigm and its ability to sensitively measure EF. Thus, the convergent validity of VR tasks would be better assessed through real-life performance on the same task, such as actually parking in a controlled environment. Although this may seem to be a resource burden to validation, it could provide integral merit to using the paradigm as a proxy for the real-life task thereafter. Alternative options are to assess convergent validity through other forms of real-life functioning (eg, self-care, residence, transportation, and employment) and diagnostic trajectory [49]. Moreover, predictive validators should be carefully chosen to ensure that their target construct aligns with that thought to be required for the function-led assessment.

Nonetheless, for novel task validation, transparent reporting of all results is crucial for advancing future research. Several papers included in this review (4/19, 21%) [61,62,68,69] reported only statistically significant correlations, leaving unanswered questions because of the omission of nonsignificant results. For instance, Chicchi Giglioli et al [69] sought to evaluate inhibition control using the GNG and ST for validation (both are common tasks for assessing inhibition) as well as the DPT yet did not report all correlational data in their results table. Such omissions hinder the comprehensive use or meta-analytic application of these findings. Conversely, Chicchi Giglioli et al [68] provided a detailed comparison between each validation task and its corresponding VR task, including the constructs assessed. However, only significant correlations were reported, some of which were between tasks intended to assess disparate constructs, such as the correlations between the Wisconsin Card Sorting Test (assessing cognitive shifting) and the VR tasks (measuring attention and inhibition control). Although these findings may indicate overlapping constructs in VR tasks, the absence of multiple-comparison correction and a detailed post hoc analysis of these correlations limits the interpretability and applicability of these results.

Finally, it is worth noting that there was significant variation in sample sizes across the studies reviewed. Although it is often accepted that pilot studies or preliminary studies have small sample sizes that often result in underpowered analysis, the utility of the VR paradigms is dependent on sound psychometric properties that require adequate sample sizes and statistical power. As detailed in Multimedia Appendix 1 [53-55,61-76], the sample sizes varied from 12 (6 per group) [74] to 103 (divided into 2 groups) [53]. Although the definition of a "sufficient" sample size may vary between studies and analyses, several of the included VR paradigms would likely require additional validation studies to provide confidence in their psychometric properties.

Recommendations

Our recommendations are as follows:

1. Papers should explicitly detail how their VR paradigms are being validated. If a paradigm has multiple components, it is essential to state how each one is being validated. A good example is the paper by Kourtesis et al [76] in this review.
2. If studies aim to validate a VR paradigm for a specific EF construct, they should identify a priori the precise outcome measures of the VR paradigm that are hypothesized to tap into various EF constructs (eg, time to completion and number of errors) and then validate them against the appropriate traditional tasks that also reliably assess those EF constructs.
3. Where appropriate, the VR paradigm's real-world task should be validated against both traditional task measures and ecologically valid measures. Ecologically valid measures may include carer reports, observation assessments, and activity of daily living assessments.
4. Multiple modes of validation should be used, including measures that provide predictive power [49], and both carer reports of daily functioning and biosensor data should be considered.
5. Papers should report all outcomes of validation data (even those in supplementary materials) to ensure the transparency of the tools' properties. A concerted effort to increase explicit and transparent reporting would greatly benefit this field.
6. To validate the VR paradigm, the psychometric properties of the traditional task must be appropriate.
7. Studies aiming to evaluate the psychometric properties of their VR paradigm should ensure that they have adequate sample sizes for a powered analysis.

Cybersickness

Overview

Although VR offers several key advantages over traditional tasks, these systems can also produce adverse effects such as cybersickness. In our review, only 21% (4/19) of the studies included an assessment of cybersickness. This is concerning as cybersickness presents a substantial confound for valid VR assessment and has been shown to negatively affect task performance [92,93]. Given that the assessment of EF involves ascertaining a participant's cognitive abilities, the recording of cybersickness is key to ensuring that common side effects such as dizziness and vertigo do not affect the participants' ability to perform at their best on the tasks. Without formal evaluation, the degree to which participants' experiences are altered is unclear. Furthermore, it is unknown at this stage whether cybersickness symptoms affect various client populations differently. For example, it is possible that, although a healthy individual may be able to continue the assessment with minor vertigo, an individual with cognitive impairment may be more affected, resulting in severely affected cognitive outcomes. Thus, caution should be exercised when using VR paradigms to ensure that the potential benefits of engagement and ecological validity are not realized at the cost of the potential negative effects of cybersickness.

Recommendations

Our recommendations are as follows:

1. Future papers should include usability data in the form of cybersickness measurements.
2. Correlations between cybersickness and participants' task performance could be included as supplementary material that should be accessible to readers, enabling them to better understand how the VR battery is performing.
3. Even when a paradigm has already assessed cybersickness, we encourage future researchers to use the same paradigm to conduct their own cybersickness assessments. This is because it is still unclear whether cybersickness will have different effects on various populations.
4. Clinical researchers and engineers should continue to investigate and report on techniques and technologies that reduce the incidence or severity of cybersickness.

Textbox 1 provides an overview of the recommendations of this review.

Textbox 1. Recommendations for future research and practice using virtual reality (VR) head-mounted display–based paradigms for executive functioning (EF) assessment.

Validate against multiple forms

- Examples include carer reports, observation assessments, ecological momentary assessments, activity of daily living assessments, physiological sensors, and in vivo studies.
- Consider longitudinal tracking of participants to establish predictive utility to initially validate the novel paradigm.

Report a priori how each assessment in the VR paradigm is being validated

- If there are multiple components to one paradigm, state how each element is being validated (a good example is the study by Kourtesis et al [76] in this review); for example: "Task 2a aims to assess inhibitory control and is validated against the traditional stop signal task and go/no-go task."

Report all validation data

- Report correlations of all aspects of a task that were identified a priori as validating the VR paradigm. In extending the previous example, show all relevant metrics from task 2a, such as errors, proportion of successful stops, reaction time, and stop signal reaction time against the relevant metrics of both the Stop Signal and Go/No-Go tasks.

Include user experience assessment

- Conduct assessments of immersion, cybersickness, usability, and engagement.

Use a common framework for defining target constructs

- The Research Domain Criteria is one option of a framework that can be applied to ensure that terminology used in the field is consistent.

Consider adding biosensors

- These provide additional objective data that may inform the VR-based EF assessment.

Limitations

We searched for articles that used the terms "executive functioning," "higher order cognition," and "functional assessment" to capture tasks that aimed to broadly assess facets of EF. This search strategy may have missed studies that examined a key construct of EF but did not specifically use the aforementioned terms (eg, used VR to assess inhibitory control alone). In addition, we did not contact the authors of the papers included in this review for further information; however, one of the key outcomes of this review was the amount of information contained in the manuscripts for future studies to extend upon.

Future Directions

The authors posit that the integration of biosensors into a VR system has significant potential. Biosensors such as pupillometry, eye gaze, EEG, and language and grammatical characteristic data can be temporally linked to the events occurring in the VR task. For example, pupillometry can offer insights into brain injury prognosis [94] and differentiate between participants with Alzheimer disease and healthy controls [95]. Eye tracking during reading aids Alzheimer disease identification [96], and linguistic attributes (eg, formation and fluency of sentences, syntax, and grammar) distinguish patients with Alzheimer disease from those with MCI [97]. The combination of these biosensor metrics and real-time function-led VR performance could increase the sensitivity of tests, enabling the detection of subtle differences such as between MCI and subjective memory complaints [98]. However, currently, biosignals are rarely evaluated alongside emerging VR paradigms for EF assessment. None of the reviewed studies used biosensors, leaving an untapped potential for VR paradigms to be frontline neuropsychological assessments.

Biosensors could also assist in modulating the cognitive load experienced by participants. Cognitive load is the cumulative working memory resources that an individual requires for a given task [99]. Similar to the gaming industry, VR paradigms could be adaptive and performance driven so that the level of challenge adjusts according to real-time individual responses [100,101]. Modulating the cognitive load adjusts the challenge of a task and enables all participants to encounter similar levels of perceived difficulty for their respective abilities. EEG, pupillometry, and cardiovascular measures are also sensitive to cognitive load capacities [99].

An additional advantage of VR is its ability to facilitate the assessment of spatial navigation. Spatial navigation is a component of cognitive functioning that can be a key factor in detecting early stages of neurodegenerative diseases. However, it cannot be assessed adequately by means of many traditional assessments. Although it is acknowledged that spatial navigation is not a component of EF, the authors of this paper consider it a generally underexamined construct when assessing cognition and general function. For example, spatial navigation is a cognitive marker used to detect early attention deficit [102,103] and offers additional relevant information beyond the traditional neuropsychological tests [103]. The environment could also be systematically manipulated to match the needs of the assessment [104] and tailored to specific populations. However, typically, spatial navigation is assessed using a real-space human analog of the Morris water maze test, which can be difficult to implement under standardized conditions. Computerized versions have been adapted, with findings comparable with those of tests conducted in real space [105], suggesting promise for translating this style of assessment to VR.

Conclusions

VR paradigms assessing EF have great potential to improve upon traditional tests. However, despite their undeniable novelty and potential, their methodological and psychometric properties must be addressed during their development to ensure their validity and reliability. Although there is no shortage of research in this area, the lack of standardized protocols to validate VR-based neuropsychological assessments hinders the progress of this field of research and practice. It is hoped that this study will be the beginning of a larger movement toward systematizing the development and validation of these paradigms.

Acknowledgments

LK received funding support from the Institute for Intelligent Systems Research and Innovation. This funding source had no role in the data analysis, presentation, or interpretation and write-up. MY received funding from government funding bodies such as the National Health and Medical Research Council; Australian Research Council; Australian Defence Science and Technology Group; Department for Industry, Innovation, and Science; and National Institutes of Health (United States), as well as philanthropic donations from the David Winston Turner Endowment Fund and Wilson Foundation; sponsored investigator-initiated trials including Incannex Healthcare Ltd; and payments in relation to court, expert witness, or expert review reports. These funding sources had no role in the data analysis, presentation, or interpretation and write-up. MY also sits on the advisory boards of the Centre for Urban Mental Health, University of Amsterdam; Monash Biomedical Imaging; and Enosis Therapeutics. The authors would also like to acknowledge Scribo Proofreading and Editing for providing their services for this manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Authors' Contributions

RK contributed to conceptualization (equal), data curation (equal), formal analysis, investigation, methodology, writing—original draft preparation (lead), and writing—review and editing. LK contributed to conceptualization (equal), formal analysis (lead), investigation, writing—original draft preparation, and writing—review and editing. KR contributed to conceptualization (equal), data curation (equal), writing—original draft preparation, and writing—review and editing. MY contributed to writing—review and editing and funding acquisition. LA and DM contributed to writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Article details including participant demographics, virtual reality (VR) paradigm, VR tasks, measures of user experience, and comparative assessments for VR paradigms.
[DOCX File , 34 KB - games_v12i1e50282_app1.docx]

Multimedia Appendix 2
PRISMA checklist.
[DOCX File , 32 KB - games_v12i1e50282_app2.docx]

References

1. Anderson P. Assessment and development of executive function (EF) during childhood. Child Neuropsychol 2002 Jun;8(2):71-82 [FREE Full text] [doi: 10.1076/chin.8.2.71.8724] [Medline: 12638061]
2. Lehto JE, Juujärvi P, Kooistra L, Pulkkinen L. Dimensions of executive functioning: evidence from children. Br J Dev Psychol 2010 Dec 23;21(1):59-80 [FREE Full text] [doi: 10.1348/026151003321164627]
3. Manchester D, Priestley N, Jackson H. The assessment of executive functions: coming out of the office. Brain Inj 2004 Nov;18(11):1067-1081 [FREE Full text] [doi: 10.1080/02699050410001672387] [Medline: 15545205]
4. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. Cogn Psychol 2000 Aug;41(1):49-100 [FREE Full text] [doi: 10.1006/cogp.1999.0734] [Medline: 10945922]
5. Suchy Y. Executive functioning: overview, assessment, and research issues for non-neuropsychologists. Ann Behav Med 2009 Apr;37(2):106-116 [FREE Full text] [doi: 10.1007/s12160-009-9097-4] [Medline: 19455377]
6. Diamond A. Executive functions. Annu Rev Psychol 2013;64:135-168 [FREE Full text] [doi: 10.1146/annurev-psych-113011-143750] [Medline: 23020641]
7. Collins A, Koechlin E. Reasoning, learning, and creativity: frontal lobe function and human decision-making. PLoS Biol 2012;10(3):e1001293 [FREE Full text] [doi: 10.1371/journal.pbio.1001293] [Medline: 22479152]
8. Lunt L, Bramham J, Morris RG, Bullock PR, Selway RP, Xenitidis K, et al. Prefrontal cortex dysfunction and 'jumping to conclusions': bias or deficit? J Neuropsychol 2012 Mar;6(1):65-78 [FREE Full text] [doi: 10.1111/j.1748-6653.2011.02005.x] [Medline: 22257612]
9. Elliott R. Executive functions and their disorders. Br Med Bull 2003;65:49-59 [FREE Full text] [doi: 10.1093/bmb/65.1.49] [Medline: 12697616]
10. Snyder HR, Miyake A, Hankin BL. Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. Front Psychol 2015 Mar 26;6:328 [FREE Full text] [doi: 10.3389/fpsyg.2015.00328] [Medline: 25859234]
11. Snyder H. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. Psychol Bull 2013 Jan;139(1):81-132 [FREE Full text] [doi: 10.1037/a0028727] [Medline: 22642228]
12. Demetriou EA, Lampit A, Quintana DS, Naismith SL, Song YJ, Pye JE, et al. Autism spectrum disorders: a meta-analysis of executive function. Mol Psychiatry 2018 May;23(5):1198-1204 [FREE Full text] [doi: 10.1038/mp.2017.75] [Medline: 28439105]
13. Chan RC, Shum D, Toulopoulou T, Chen EY. Assessment of executive functions: review of instruments and identification of critical issues. Arch Clin Neuropsychol 2008 Mar;23(2):201-216 [FREE Full text] [doi: 10.1016/j.acn.2007.08.010] [Medline: 18096360]
14. Burgess PW, Alderman N, Forbes C, Costello A, M-A.Coates L, Dawson DR, et al. The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. J Int Neuropsychol Soc 2006 Mar 22;12(2):194-209 [FREE Full text] [doi: 10.1017/s1355617706060310]
15. Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. J Int Neuropsychol Soc 1998 Nov;4(6):547-558 [FREE Full text] [doi: 10.1017/s1355617798466037] [Medline: 10050359]
16. Sbordone RJ. Ecological validity: some critical issues for the neuropsychologist. In: Sbordone RJ, Long CJ, editors. Ecological Validity of Neuropsychological Testing. Boca Raton, FL: CRC Press; 1996.
17. Kvavilashvili L, Ellis JA. Ecological validity and the real-life/laboratory controversy in memory research: a critical and historical review. History Philos Psychol 2004;6(1):59-80 [FREE Full text]
18. Spooner DM, Pachana NA. Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. Arch Clin Neuropsychol 2006 May;21(4):327-337 [FREE Full text] [doi: 10.1016/j.acn.2006.04.004] [Medline: 16769198]
19. Chaytor N, Schmitter-Edgecombe M, Burr R. Improving the ecological validity of executive functioning assessment. Arch Clin Neuropsychol 2006 Apr;21(3):217-227 [FREE Full text] [doi: 10.1016/j.acn.2005.12.002] [Medline: 16554143]
20. Gärtner A, Strobel A. Individual differences in inhibitory control: a latent variable analysis. J Cogn 2021 Feb 18;4(1):17 [FREE Full text] [doi: 10.5334/joc.150] [Medline: 33634234]

21. Choi SY, Lee JS, Oh YJ. Cut-off point for the trail making test to predict unsafe driving after stroke. J Phys Ther Sci 2016 Jul;28(7):2110-2113 [FREE Full text] [doi: 10.1589/jpts.28.2110] [Medline: 27512277]
22. Papandonatos GD, Ott BR, Davis JD, Barco PP, Carr DB. Clinical utility of the trail-making test as a predictor of driving performance in older adults. J Am Geriatr Soc 2015 Nov;63(11):2358-2364 [FREE Full text] [doi: 10.1111/jgs.13776] [Medline: 26503623]
23. Dobbs BM, Shergill SS. How effective is the Trail Making Test (Parts A and B) in identifying cognitively impaired drivers? Age Ageing 2013 Sep;42(5):577-581 [FREE Full text] [doi: 10.1093/ageing/aft073] [Medline: 23896609]
24. Duncanson H, Hollis AM, O'Connor MG. Errors versus speed on the trail making test: relevance to driving performance. Accid Anal Prev 2018 Apr;113:125-130 [FREE Full text] [doi: 10.1016/j.aap.2018.01.004] [Medline: 29407659]
25. Vaucher P, Herzig D, Cardoso I, Herzog MH, Mangin P, Favrat B. The trail making test as a screening instrument for driving performance in older drivers; a translational research. BMC Geriatr 2014 Nov 24;14:123 [FREE Full text] [doi: 10.1186/1471-2318-14-123] [Medline: 25420615]
26. Bublak P, Redel P, Sorg C, Kurz A, Förstl H, Müller HJ, et al. Staged decline of visual processing capacity in mild cognitive impairment and Alzheimer's disease. Neurobiol Aging 2011 Jul;32(7):1219-1230 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2009.07.012] [Medline: 19713001]
27. Buracchio T, Dodge HH, Howieson D, Wasserman D, Kaye J. The trajectory of gait speed preceding mild cognitive impairment. Arch Neurol 2010 Aug;67(8):980-986 [FREE Full text] [doi: 10.1001/archneurol.2010.159] [Medline: 20697049]
28. Allcoat D, von Mühlenen A. Learning in virtual reality: effects on performance, emotion and engagement. Res Learn Technol 2018 Nov 27;26 [FREE Full text] [doi: 10.25304/rlt.v26.2140]
29. Huang W, Roscoe RD, Johnson‐Glenberg MC, Craig SD. Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. Comput Assist Learn 2021 Jun;37(3):745-758 [FREE Full text] [doi: 10.1111/jcal.12520]
30. Parsons TD, Phillips AS. Virtual reality for psychological assessment in clinical practice. Practic Innov 2016;1(3):197-217 [FREE Full text] [doi: 10.1037/pri0000028]
31. Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. Brain 1991 Apr;114 (Pt 2):727-741 [FREE Full text] [doi: 10.1093/brain/114.2.727] [Medline: 2043945]
32. Dawson DR, Anderson ND, Burgess P, Cooper E, Krpan KM, Stuss DT. Further development of the Multiple Errands Test: standardized scoring, reliability, and ecological validity for the Baycrest version. Arch Phys Med Rehabil 2009 Nov;90(11 Suppl):S41-S51 [FREE Full text] [doi: 10.1016/j.apmr.2009.07.012] [Medline: 19892074]
33. Maeir A, Krauss S, Katz N. Ecological validity of the multiple errands test (MET) on discharge from neurorehabilitation hospital. OTJR (Thorofare N J) 2011;31(1):S38-S46 [FREE Full text] [doi: 10.3928/15394492-20101108-07] [Medline: 24650263]
34. Antoniak K, Clores J, Jensen D, Nalder E, Rotenberg S, Dawson DR. Developing and validating a big-store multiple errands test. Front Psychol 2019 Nov 15;10:2575 [FREE Full text] [doi: 10.3389/fpsyg.2019.02575] [Medline: 31803113]
35. Nalder EJ, Clark AJ, Anderson ND, Dawson DR. Clinicians' perceptions of the clinical utility of the Multiple Errands Test for adults with neurological conditions. Neuropsychol Rehabil 2017 Jul;27(5):685-706 [FREE Full text] [doi: 10.1080/09602011.2015.1067628] [Medline: 26235491]
36. Draschkow D. Remote virtual reality as a tool for increasing external validity. Nat Rev Psychol 2022 Jun 16;1:433-434 [FREE Full text] [doi: 10.1038/s44159-022-00082-8]
37. Pedroli E, Cipresso P, Serino S, Pallavicini F, Albani G, Riva G. Virtual Multiple Errands Test: reliability, usability and possible applications. Stud Health Technol Inform 2013;191:38-42. [Medline: 23792839]
38. Cipresso P, Albani G, Serino S, Pedroli E, Pallavicini F, Mauro A, et al. Virtual multiple errands test (VMET): a virtual reality-based tool to detect early executive functions deficit in Parkinson's disease. Front Behav Neurosci 2014 Dec 5;8:405 [FREE Full text] [doi: 10.3389/fnbeh.2014.00405] [Medline: 25538578]
39. Riva G, Gaggioli A, Grassi A, Raspelli S, Cipresso P, Pallavicini F, et al. NeuroVR 2--a free virtual reality platform for the assessment and treatment in behavioral health care. Stud Health Technol Inform 2011;163:493-495. [Medline: 21335845]
40. Mount NJ, Chambers C, Weaver D, Priestnall G. Learner immersion engagement in the 3D virtual world: principles emerging from the DELVE project. Innov Teach Learn Inf Comput Sci 2009;8(3):40-55 [FREE Full text] [doi: 10.11120/ital.2009.08030040]
41. Lumsden J, Edwards EA, Lawrence NS, Coyle D, Munafò MR. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games 2016 Jul 15;4(2):e11 [FREE Full text] [doi: 10.2196/games.5888] [Medline: 27421244]
42. Krath J, Schürmann L, von Korflesch HF. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. Comput Hum Behav 2021 Dec;125:106963 [FREE Full text] [doi: 10.1016/j.chb.2021.106963]
43. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DM. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. JMIR Serious Games 2020 Aug 10;8(3):e18644 [FREE Full text] [doi: 10.2196/18644] [Medline: 32773374]

44. Llobera J, Jacquat V, Calabrese C, Charbonnier C. Playing the mirror game in virtual reality with an autonomous character. Sci Rep 2022 Dec 09;12(1):21329 [FREE Full text] [doi: 10.1038/s41598-022-25197-z] [Medline: 36494394]
45. Mahtab EA, Egorova AD. Current and future applications of virtual reality technology for cardiac interventions. Nat Rev Cardiol 2022 Dec;19(12):779-780 [FREE Full text] [doi: 10.1038/s41569-022-00789-4] [Medline: 36195685]
46. Li G, Anguera JA, Javed SV, Khan MA, Wang G, Gazzaley A. Enhanced attention using head-mounted virtual reality. J Cogn Neurosci 2020 Aug;32(8):1438-1454 [FREE Full text] [doi: 10.1162/jocn_a_01560] [Medline: 32286132]
47. Nalivaiko E, Davis SL, Blackmore KL, Vakulin A, Nesbitt KV. Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time. Physiol Behav 2015 Nov 01;151:583-590 [FREE Full text] [doi: 10.1016/j.physbeh.2015.08.043] [Medline: 26340855]
48. Sepich NC, Jasper A, Fieffer S, Gilbert SB, Dorneich MC, Kelly JW. The impact of task workload on cybersickness. Front Virtual Real 2022 Aug 19;3:1-19 [FREE Full text] [doi: 10.3389/frvir.2022.943409]
49. Negu A, Matu SA, Sava FA, David D. Virtual reality measures in neuropsychological assessment: a meta-analytic review. Clin Neuropsychol 2016 Feb;30(2):165-184 [FREE Full text] [doi: 10.1080/13854046.2016.1144793] [Medline: 26923937]
50. Borgnis F, Baglio F, Pedroli E, Rossetto F, Uccellatore L, Oliveira JA, et al. Available virtual reality-based tools for executive functions: a systematic review. Front Psychol 2022 Apr 11;13:833136 [FREE Full text] [doi: 10.3389/fpsyg.2022.833136] [Medline: 35478738]
51. Kim E, Han J, Choi H, Prié Y, Vigier T, Bulteau S, et al. Examining the academic trends in neuropsychological tests for executive functions using virtual reality: systematic literature review. JMIR Serious Games 2021 Nov 24;9(4):e30249 [FREE Full text] [doi: 10.2196/30249] [Medline: 34822341]
52. Urbina S. Essentials of Psychological Testing. Hoboken, NJ: John Wiley & Sons; Aug 4, 2014.
53. Parsons TD, McMahan T. An initial validation of the virtual environment grocery store. J Neurosci Methods 2017 Nov 01;291:13-19 [FREE Full text] [doi: 10.1016/j.jneumeth.2017.07.027] [Medline: 28782630]
54. Wilf M, Ben Yair N, Wright WG, Plotnik M. The trail less traveled: analytical approach for creating shortened versions for virtual reality-based color trails test. Appl Neuropsychol Adult 2022 May 06:1-10 [FREE Full text] [doi: 10.1080/23279095.2022.2065204] [Medline: 35522843]
55. Williams RM, Alikhademi K, Gilbert JE. Design of a toolkit for real-time executive function assessment in custom-made virtual experiences and interventions. Int J Hum Comput Stud 2022 Feb;158:102734 [FREE Full text] [doi: 10.1016/j.ijhcs.2021.102734]
56. Baggetta P, Alexander PA. Conceptualization and operationalization of executive function. Mind Brain Educ 2016 Feb 18;10(1):10-33 [FREE Full text] [doi: 10.1111/mbe.12100]
57. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
58. Zhu J, Liu W. A tale of two databases: the use of Web of Science and Scopus in academic papers. Scientometrics 2020 Feb 22;123(1):321-335 [FREE Full text] [doi: 10.1007/s11192-020-03387-8]
59. Mongeon P, Paul-Hus A. The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics 2015 Oct 19;106(1):213-228 [FREE Full text] [doi: 10.1007/s11192-015-1765-5]
60. Babineau J. Product review: covidence (systematic review software). J Can Health Libr Assoc 2014;35(2) [FREE Full text] [doi: 10.5596/c14-016]
61. Banville F, Lussier C, Massicotte E, Verhulst E, Couture JF, Allain P, et al. Validation of a sorting task implemented in the virtual multitasking task-2 and effect of aging. In: Proceedings of the Human Interface and the Management of Information. Information in Applications and Services. 2018 Presented at: Human Interface and the Management of Information. Information in Applications and Services; July 15-20, 2018; Las Vegas, NV URL: https://doi.org/10.1007/978-3-319-92046-7_4 [doi: 10.1007/978-3-319-92046-7_4]
62. Davison SM, Deeprose C, Terbeck S. A comparison of immersive virtual reality with traditional neuropsychological measures in the assessment of executive functions. Acta Neuropsychiatr 2018 Apr;30(2):79-89 [FREE Full text] [doi: 10.1017/neu.2017.14] [Medline: 28482936]
63. Miskowiak KW, Jespersen AE, Kessing LV, Aggestrup AS, Glenthøj LB, Nordentoft M, et al. Cognition assessment in virtual reality: validity and feasibility of a novel virtual reality test for real-life cognitive functions in mood disorders and psychosis spectrum disorders. J Psychiatr Res 2021 Dec 12;145:182-189 [FREE Full text] [doi: 10.1016/j.jpsychires.2021.12.002] [Medline: 34923359]
64. Pallavicini F, Pepe A, Minissi ME. Taking neuropsychological test to the next level: commercial virtual reality video games for the assessment of executive functions. In: Proceedings of the Universal Access in Human-Computer Interaction. Multimodality and Assistive Environments. 2019 Presented at: UAHCI 2019; July 26-31, 2019; Orlando, FL URL: https://doi.org/10.1007/978-3-030-23563-5_12 [doi: 10.1007/978-3-030-23563-5_12]
65. Porffy LA, Mehta MA, Patchitt J, Boussebaa C, Brett J, D'Oliveira T, et al. A novel virtual reality assessment of functional cognition: validation study. J Med Internet Res 2022 Jan 26;24(1):e27641 [FREE Full text] [doi: 10.2196/27641] [Medline: 35080501]

66. Tan NC, Lim JE, Allen JCJ, Wong WT, Quah JH, Muthulakshmi P, et al. Age-related performance in using a fully immersive and automated virtual reality system to assess cognitive function. Front Psychol 2022;13:847590 [FREE Full text] [doi: 10.3389/fpsyg.2022.847590] [Medline: 35360611]
67. Tsai CF, Chen CC, Wu EH, Chung CR, Huang CY, Tsai PY, et al. A machine-learning-based assessment method for early-stage neurocognitive impairment by an immersive virtual supermarket. IEEE Trans Neural Syst Rehabil Eng 2021;29:2124-2132 [FREE Full text] [doi: 10.1109/tnsre.2021.3118918]
68. Chicchi Giglioli IA, de Juan Ripoll C, Parra E, Alcañiz Raya M. Are 3D virtual environments better than 2D interfaces in serious games performance? An explorative study for the assessment of executive functions. Appl Neuropsychol Adult 2021;28(2):148-157 [FREE Full text] [doi: 10.1080/23279095.2019.1607735] [Medline: 31070055]
69. Chicchi Giglioli IA, Pérez Gálvez B, Gil Granados A, Alcañiz Raya M. The virtual cooking task: a preliminary comparison between neuropsychological and ecological virtual reality tests to assess executive functions alterations in patients affected by alcohol use disorder. Cyberpsychol Behav Soc Netw 2021 Oct;24(10):673-682 [FREE Full text] [doi: 10.1089/cyber.2020.0560] [Medline: 33761276]
70. Marín-Morales J, Carrasco-Ribelles LA, Alcañiz M, Giglioli IA. Applying machine learning to a virtual serious game for neuropsychological assessment. In: Proceedings of the IEEE Global Engineering Education Conference (EDUCON). 2021 Presented at: IEEE Global Engineering Education Conference (EDUCON); April 21-23, 2021; Vienna, Austria URL: https://doi.org/10.1109/EDUCON46332.2021.9454138 [doi: 10.1109/educon46332.2021.9454138]
71. Voinescu A, Petrini K, Stanton Fraser D, Lazarovicz RA, Papavă I, Fodor LA, et al. The effectiveness of a virtual reality attention task to predict depression and anxiety in comparison with current clinical measures. Virtual Reality 2021 Apr 17;27(1):119-140 [FREE Full text] [doi: 10.1007/s10055-021-00520-7]
72. Parsons TD, Carlew AR. Bimodal virtual reality stroop for assessing distractor inhibition in autism spectrum disorders. J Autism Dev Disord 2016 Apr 27;46(4):1255-1267. [doi: 10.1007/s10803-015-2663-7] [Medline: 26614084]
73. Parsons TD, Barnett M. Virtual apartment-based stroop for assessing distractor inhibition in healthy aging. Appl Neuropsychol Adult 2019;26(2):144-154 [FREE Full text] [doi: 10.1080/23279095.2017.1373281] [Medline: 28976213]
74. Robitaille N, Jackson PL, Hébert LJ, Mercier C, Bouyer LJ, Fecteau S, et al. A Virtual Reality avatar interaction (VRai) platform to assess residual executive dysfunction in active military personnel with previous mild traumatic brain injury: proof of concept. Disabil Rehabil Assist Technol 2017 Oct;12(7):758-764 [FREE Full text] [doi: 10.1080/17483107.2016.1229048] [Medline: 27677827]
75. Kourtesis P, MacPherson SE. An ecologically valid examination of event-based and time-based prospective memory using immersive virtual reality: the influence of attention, memory, and executive function processes on real-world prospective memory. Neuropsychol Rehabil 2023 Mar;33(2):255-280 [FREE Full text] [doi: 10.1080/09602011.2021.2008983] [Medline: 34856886]
76. Kourtesis P, Collina S, Doumas LA, MacPherson SE. Validation of the virtual reality everyday assessment lab (VR-EAL): an immersive virtual reality neuropsychological battery with enhanced ecological validity. J Int Neuropsychol Soc 2020 Aug 10;27(2):181-196 [FREE Full text] [doi: 10.1017/s1355617720000764]
77. Delis DC, Kaplan E, Kramer JH. Delis-Kaplan Executive Function System:Technical Manual. San Antonio, TX: Harcourt Assessment Company; 2001.
78. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. BADS: Behavioural Assessment of the Dysexecutive Syndrome. London, UK: Pearson Education; 1996.
79. Wilson BA. The Cambridge Prospective Memory Test: CAMPROMPT. London, UK: Pearson Assessment; 2005.
80. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test-Second Edition (CVLT-II). San Antonio, TX: Psychological Corporation; 2000.
81. Wilson BA, Greenfield E, Clare L, Baddeley A, Cockburn J, Watson P, et al. Rivermead Behavioural Memory Test - Third Edition. London, UK: Pearson Education; 2008.
82. D'Elia LF, Satz P, Uchiyama CL, White T. Color Trails Test. Lutz, FL: Psychological Assessment Resources; 1996.
83. Mueller ST, Piper BJ. The psychology experiment building language (PEBL) and PEBL test battery. J Neurosci Methods 2014 Jan 30;222:250-259 [FREE Full text] [doi: 10.1016/j.jneumeth.2013.10.024] [Medline: 24269254]
84. Nigg J. Annual research review: on the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. J Child Psychol Psychiatry 2017 Apr;58(4):361-383 [FREE Full text] [doi: 10.1111/jcpp.12675] [Medline: 28035675]
85. Tiego J, Testa R, Bellgrove MA, Pantelis C, Whittle S. A hierarchical model of inhibitory control. Front Psychol 2018;9:1339 [FREE Full text] [doi: 10.3389/fpsyg.2018.01339] [Medline: 30123154]
86. Cuthbert BN. Research Domain Criteria (RDoC): Progress and Potential. Curr Dir Psychol Sci 2022 Apr;31(2):107-114 [FREE Full text] [doi: 10.1177/09637214211051363] [Medline: 35692384]
87. Yücel M, Oldenhof E, Ahmed SH, Belin D, Billieux J, Bowden-Jones H, et al. A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. Addiction 2019 Jun;114(6):1095-1109 [FREE Full text] [doi: 10.1111/add.14424] [Medline: 30133930]

88. Kappenman ES, Farrens JL, Luck SJ, Proudfit GH. Behavioral and ERP measures of attentional bias to threat in the dot-probe task: poor reliability and lack of correlation with anxiety. Front Psychol 2014;5:1368 [FREE Full text] [doi: 10.3389/fpsyg.2014.01368] [Medline: 25538644]
89. Schmukle SC. Unreliability of the dot probe task. Eur J Pers 2020 Dec 02;19(7):595-605 [FREE Full text] [doi: 10.1002/per.554]
90. Chapman A, Devue C, Grimshaw GM. Fleeting reliability in the dot-probe task. Psychol Res 2019 Mar;83(2):308-320 [FREE Full text] [doi: 10.1007/s00426-017-0947-6] [Medline: 29159699]
91. Abma IL, Rovers M, van der Wees PJ. Appraising convergent validity of patient-reported outcome measures in systematic reviews: constructing hypotheses and interpreting outcomes. BMC Res Notes 2016 Apr 19;9:226 [FREE Full text] [doi: 10.1186/s13104-016-2034-2] [Medline: 27094345]
92. Pöhlmann KM, Li G, Mcgill M, Markoff R, Brewster SA. You spin me right round, baby, right round: examining the impact of multi-sensory self-motion cues on motion sickness during a VR reading task. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany. [doi: 10.1145/3544548.3580966]
93. Kim YY, Kim HJ, Kim EN, Ko HD, Kim HT. Characteristic changes in the physiological components of cybersickness. Psychophysiology 2005 Sep;42(5):616-625 [FREE Full text] [doi: 10.1111/j.1469-8986.2005.00349.x] [Medline: 16176385]
94. Solari D, Miroz JP, Oddo M. Opening a window to the injured brain: non-invasive neuromonitoring with quantitative pupillometry. In: Vincent JL, editor. Annual Update in Intensive Care and Emergency Medicine 2018. Cham, Switzerland: Springer; 2018.
95. Pavisic IM, Firth NC, Parsons S, Rego DM, Shakespeare TJ, Yong KX, et al. Eyetracking metrics in young onset Alzheimer's disease: a window into cognitive visual functions. Front Neurol 2017;8:377 [FREE Full text] [doi: 10.3389/fneur.2017.00377] [Medline: 28824534]
96. Fernández G, Manes F, Politi LE, Orozco D, Schumacher M, Castro L, et al. Patients with mild Alzheimer's disease fail when using their working memory: evidence from the eye tracking technique. J Alzheimer's Dis 2016 Feb 02;50(3):827-838 [FREE Full text] [doi: 10.3233/jad-150265]
97. Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. J Alzheimer's Dis 2015 Nov 27;49(2):407-422 [FREE Full text] [doi: 10.3233/jad-150520]
98. Iliadou P, Paliokas I, Zygouris S, Lazarou E, Votis K, Tzovaras D, et al. A comparison of traditional and serious game-based digital markers of cognition in older adults with mild cognitive impairment and healthy controls. J Alzheimer's Dis 2021 Feb 16;79(4):1747-1759 [FREE Full text] [doi: 10.3233/jad-201300]
99. Ranchet M, Morgan JC, Akinwuntan AE, Devos H. Cognitive workload across the spectrum of cognitive impairments: a systematic review of physiological measures. Neurosci Biobehav Rev 2017 Sep;80:516-537 [FREE Full text] [doi: 10.1016/j.neubiorev.2017.07.001] [Medline: 28711663]
100. Mishra J, Anguera JA, Gazzaley A. Video games for neuro-cognitive optimization. Neuron 2016 Apr 20;90(2):214-218 [FREE Full text] [doi: 10.1016/j.neuron.2016.04.010] [Medline: 27100194]
101. LaConte SM. Decoding fMRI brain states in real-time. Neuroimage 2011 May 15;56(2):440-454 [FREE Full text] [doi: 10.1016/j.neuroimage.2010.06.052] [Medline: 20600972]
102. Benke T, Karner E, Petermichl S, Prantner V, Kemmler G. Neuropsychological deficits associated with route learning in Alzheimer disease, MCI, and normal aging. Alzheimer Dis Assoc Disord 2014;28(2):162-167 [FREE Full text] [doi: 10.1097/WAD.0000000000000009] [Medline: 24077018]
103. Laczó J, Andel R, Nedelska Z, Vyhnalek M, Vlcek K, Crutch S, et al. Exploring the contribution of spatial navigation to cognitive functioning in older adults. Neurobiol Aging 2017 Mar;51:67-70 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2016.12.003] [Medline: 28039765]
104. Cogné M, Taillade M, N'Kaoua B, Tarruella A, Klinger E, Larrue F, et al. The contribution of virtual reality to the diagnosis of spatial navigation disorders and to the study of the role of navigational aids: a systematic literature review. Ann Phys Rehabil Med 2017 Jun;60(3):164-176 [FREE Full text] [doi: 10.1016/j.rehab.2015.12.004] [Medline: 27017533]
105. Schoenfeld R, Foreman N, Leplow B. Ageing and spatial reversal learning in humans: findings from a virtual water maze. Behav Brain Res 2014 Aug 15;270:47-55 [FREE Full text] [doi: 10.1016/j.bbr.2014.04.036] [Medline: 24815214]

Abbreviations

CAVIR: cognition assessment in virtual reality
CPT: continuous performance test
DPT: dot-probe task
EAL: Everyday Assessment Lab
EEG: electroencephalography
EF: executive functioning
FAST: Functioning Assessment Short Test
GNG: Go/No-Go

MCI: mild cognitive impairment
MET: Multiple Errands Test
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RDoC: Research Domain Criteria
ST: Stroop test
TMT: Trail-Making Test
TMT-A: Trail-Making Test version A
TMT-B: Trail-Making Test version B
UPSA-B: brief University of California, San Diego, Performance-Based Skills Assessment
VMT: Virtual Multitasking Test
VR: virtual reality
WoS: Web of Science

Edited by G Eysenbach; submitted 03.07.23; peer-reviewed by S Turkay, E Seesjarvi; comments to author 05.10.23; revised version received 26.11.23; accepted 29.12.23; published 26.02.24.

Please cite as:
Kirkham R, Kooijman L, Albertella L, Myles D, Yücel M, Rotaru K
Immersive Virtual Reality–Based Methods for Assessing Executive Functioning: Systematic Review
JMIR Serious Games 2024;12:e50282
URL: https://games.jmir.org/2024/1/e50282
doi:10.2196/50282
PMID:38407958

Review

Digital Gamification Tools to Enhance Vaccine Uptake: Scoping Review

Hina Hakim[1], PhD; S Michelle Driedger[2], PhD; Dominique Gagnon[3], MSc; Julien Chevrier[4], MLIS; Geneviève Roch[5,6,7], MSc, RN, PhD; Eve Dubé[3,8], PhD; Holly O Witteman[1,6,7], PhD

[1]Department of Family and Emergency Medicine, Université Laval, Québec City, QC, Canada
[2]Department of Community Health Sciences, University of Manitoba, Winnipeg, MB, Canada
[3]Direction des risques biologiques, Institut national de santé publique du Québec, Quebec City, QC, Canada
[4]Bibliothèque Louise-Lalonde-Lamarre, Polytechnique Montréal, Montréal, QC, Canada
[5]Faculty of Nursing, Université Laval, Quebec City, QC, Canada
[6]Centre hospitalier universitaire (CHU) de Québec-Université Laval, Université Laval, Quebec City, QC, Canada
[7]VITAM Research Centre for Sustainable Health, Université Laval, Quebec City, QC, Canada
[8]Département d'anthropologie, Université Laval, Quebec City, QC, Canada

Corresponding Author:
Hina Hakim, PhD
Department of Family and Emergency Medicine
Université Laval
1050 avenue de la Médecine, Pavillon Ferdinand-Vandry
Québec City, QC, G1V 0A6
Canada
Phone: 1 418 656 2131
Email: hina.hakim.1@ulaval.ca

Abstract

Background: Gamification has been used successfully to promote various desired health behaviors. Previous studies have used gamification to achieve desired health behaviors or facilitate their learning about health.

Objective: In this scoping review, we aimed to describe digital gamified tools that have been implemented or evaluated across various populations to encourage vaccination, as well as any reported effects of identified tools.

Methods: We searched Medline, Embase, CINAHL, the Web of Science Core Collection, the Cochrane Database of Systematic Reviews, the Cochrane Central Register of Controlled Trials, Academic Search Premier, PsycInfo, Global Health, and ERIC for peer-reviewed papers describing digital gamified tools with or without evaluations. We also conducted web searches with Google to identify digital gamified tools lacking associated publications. We consulted 12 experts in the field of gamification and health behavior to identify any papers or tools we might have missed. We extracted data about the target population of the tools, the interventions themselves (eg, type of digital gamified tool platform, type of disease/vaccine, type and design of study), and any effects of evaluated tools, and we synthesized data narratively.

Results: Of 1402 records, we included 28 (2%) peer-reviewed papers and 10 digital gamified tools lacking associated publications. The experts added 1 digital gamified tool that met the inclusion criteria. Our final data set therefore included 28 peer-reviewed papers and 11 digital gamified tools. Of the 28 peer-reviewed papers, 7 (25%) explained the development of the tool, 16 (57%) described evaluation, and 2 (7%) reported both development and evaluation of the tool. The 28 peer-reviewed papers reported on 25 different tools. Of these 25 digital gamified tools, 11 (44%) were web-based tools, 8 (32%) mobile (native mobile or mobile-enabled web) apps, and 6 (24%) virtual reality tools. Overall, tools that were evaluated showed increases in knowledge and intentions to receive vaccines, mixed effects on attitudes, and positive effects on beliefs. We did not observe discernible advantages of one type of digital gamified tool (web based, mobile, virtual reality) over the others. However, a few studies were randomized controlled trials, and publication bias may have led to such positive effects having a higher likelihood of appearing in the peer-reviewed literature.

Conclusions: Digital gamified tools appear to have potential for improving vaccine uptake by fostering positive beliefs and increasing vaccine-related knowledge and intentions. Encouraging comparative studies of different features or different types of digital gamified tools could advance the field by identifying features or types of tools that yield more positive effects across populations and contexts. Further work in this area should seek to inform the implementation of gamification for vaccine acceptance and promote effective health communication, thus yielding meaningful health and social impacts.

(JMIR Serious Games 2024;12:e47257) doi:10.2196/47257

KEYWORDS

digital gamified tools; digital game; vaccination; gamification; vaccine uptake; scoping review; review method; vaccine; gamified; COVID-19; COVID; SARS-CoV-2; health behaviour; health behavior; health promotion; behavior change; behaviour change

Introduction

Vaccination is one of the most cost-effective methods of preventing the spread of vaccine-preventable diseases. If vaccination coverage falls below the thresholds that are safe for the prevention of epidemic transmission, the incidence of vaccine-preventable diseases increases [1,2]. For example, measles returned over the past 2 decades, and the incidence of measles in the European Union increased in 2017-2018 [3].

In 2019, prior to the COVID-19 pandemic, the World Health Organization identified vaccine hesitancy (ie, the reluctance or refusal to be vaccinated despite the availability of vaccination services) as 1 of the top 10 threats to worldwide health [4]. Vaccine hesitancy is one of the several reasons some people are un- or undervaccinated [5-9]. Interventions addressing vaccine hesitancy are therefore necessary to promote vaccine acceptance and uptake. As the contributors of vaccine acceptance are diverse, no single intervention will solve this issue [10]. Multicomponent interventions tailored to local barriers to vaccine acceptance and uptake are known to be the most effective [11,12]. Misinformation and conspiracy theories spread online, where extensive antivaccine content is shared [13-15], potentially negatively influencing views about vaccines [16,17]. Efforts have been made to counter vaccine misinformation and mistrust by targeting various groups, such as parents, non–health care workers [18,19], and adolescents [20], and delivering information about the risks and benefits of different types of vaccines, for instance, human papillomavirus (HPV) vaccination [21] and measles, mumps, and rubella vaccines [22,23]. Along with traditional communication strategies, the use of other strategies to inform and educate about immunizations, for example, with digital gamified tools, may help encourage vaccine uptake.

Gamification is defined as the use of game design elements in nongame contexts [24]. It includes several aspects and features, such as fun interfaces, immediate success or feedback, reward systems (levels, point scores, badges), challenges and competitions, team playing, avatars, and quizzes. Previous studies have used gamification to achieve desired health behaviors [25-27] or facilitate their learning about health [28]. Gamification draws on elements from serious games, meaning fully developed digital games used to train and educate players [29,30]. For example, a serious game “Land of Secret Gardens” facilitates conversations about HPV with preteens. In the game, preteens need to protect their bodies with a “potion,” which offers a metaphor for the HPV vaccine [31]. However, serious games and digital gamified tools are distinct but related concepts. Serious games use gaming as a central and primary medium [32]. In contrast, digital gamified tools (eg, apps) or gamified interventions are not complete game experiences but have gaming features, such as rewards systems, scoring of points, or engaging users in different challenges [33]. In this study, we defined digital gamified tools as digital apps with the aforementioned gaming features. Our definition includes serious games that meet the criteria, that is, they must include such gaming features. This scoping review provides insight into the reported effects of digital gamified tools to increase vaccine uptake. Our review built upon existing reviews in the field by including a comprehensive search of both published literature and online tools, as well as an examination of both the characteristics and the reported effects of these digital tools. This review was distinct in that it focused specifically on digital gamified tools and their effects, rather than simply the effectiveness of gamification in general. In doing so, this review aimed to fill a gap in the literature by providing evidence-based answers to the question of whether gamification “works” to increase vaccine uptake. Therefore, the objectives of this scoping review [34] were, first, to review digital gamified tools that have been implemented or evaluated across various populations to encourage vaccine uptake and, second, to describe any reported effects of the identified tools in terms of influence on users’ knowledge or behavior toward vaccination. Our research questions can therefore be summarized as follows:

- What digital gamified tools intended to encourage vaccination exist and have been described in the literature?
- Do these tools demonstrate any effects on knowledge, attitudes, beliefs, and behaviors about vaccination?

Methods

Search Strategy

For peer-reviewed papers, we searched Medline (Ovid), Embase (Ovid), CINAHL (EBSCO), the Web of Science Core Collection, the Cochrane Database of Systematic Reviews (Ovid), the Cochrane Central Register of Controlled Trials (Ovid), Academic Search Premier (EBSCO), PsycInfo (Ovid), Global Health (Ovid), and ERIC (Ovid) with no language or date restrictions. The proposed search terms were, for example, “vaccine,” “vaccination,” “immunization,” “video games,” “gamification,” “application,” and “virtual reality” (see Multimedia Appendix 1 for the full search strategy). The search was conducted on January 26 and 27, 2022.

We also conducted an online Google search on May 5, 2022, for any digital tools with gamified features that deliver informative or educative messages on vaccination. The search terms we used were “vaccination,” “immunization,” “electronic game,” “computer game,” “mobile game,” “interactive game,” and “digital game” (see Multimedia Appendix 1 for the full search strategy). We reviewed the first 30 results for each search, as it is rare for users to click past the third page of 10 search results per page, and therefore, researchers analyzing medical content available on the web often use 30 as a threshold [35-37]. On May 6, 2022, we conducted the same searches in private browsing mode to ascertain whether our results had been affected by a "filter bubble" [38], that is, the way Google search results are adapted to one's previous browsing activity. Our search strategy was constructed and reviewed by 2 librarians. Following the librarians' advice, we expanded our search strategy to include ERIC and Global Health databases.

Study Selection and Screening Process

We used PICO (Population, Intervention, Comparison, and Outcome) to structure study inclusion and exclusion criteria (Table 1). Our population of interest was the general public or any subgroup, including health care professionals and students. We sought studies describing tools with gamification techniques or gamified elements, including gamified web-based quizzes to deliver informative or educative messages on vaccination. Posters, preprints, editorials, conference proceedings, news bulletins, and paper-based or board games were excluded. Our comparator was any control, including offering no education on vaccination or comparing participants before and after an intervention. Our outcomes of interest included common outcomes associated with vaccine uptake, namely knowledge (comprehension, understanding), attitudes (for or against vaccination), beliefs (perceived benefits, perceived risks), and behaviors toward vaccines (vaccination intention [ie, intention to get vaccinated or not get vaccinated] and vaccine uptake [ie, receiving or not receiving a vaccine]). We excluded papers that did not present the description or evaluation of a concrete digital gamified tool.

Table 1. Inclusion and exclusion criteria.

Component	Inclusion criteria	Exclusion criteria	Question related to the criteria
Type of report	• Original paper • Evaluated intervention or digital gamified tool	• Posters, preprints, and conference proceedings • Modeling or simulation study • Brochures • Editorials • Bulletins	Has the study or research described the development of the tool and evaluated it?
Population	• General public (any subgroup) • Professionals • Students	N/A[a]	Who is the audience for whom the key message was intended?
Intervention	• Tools with gamification technique or gamified elements, including gamified web-based quizzes to deliver informative or educative messages on vaccination	• Any study or gamification tools not intended for vaccination/vaccine uptake • Studies or apps to reduce vaccine pain and fears and to report immunization status or record keeping, surveillance or vaccine coverage apps, contact-tracing or early detection apps • Paper games, board games (not digital) • Videos with no gamified element included	Does the study or tool aim to deliver an informative or educative message on vaccination?
Comparator	• Any control, including offering no education or no digital gamified tool	N/A	N/A
Outcome	• Common outcomes that encourage vaccine uptake: knowledge (comprehension, understanding), attitudes (for or against vaccination, beliefs (risk perception, etc), behaviors toward vaccines (vaccination intention [ie, intention to get vaccinated or not get vaccinated] and vaccine uptake [ie, receiving or not receiving a vaccine])	• Outcomes not related to the encouragement of vaccine uptake	Has the study or tool been evaluated for the outcomes that encourage vaccine uptake?

[a]N/A: not applicable.

For Google-searched digital gamified tools, our inclusion and exclusion criteria used the same specifications regarding population and intervention. We did not apply comparison and outcome criteria to web-based tools because we did not expect these to report evaluation studies.

We reported this review according to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (see the PRISMA checklist in Multimedia Appendix 2) [39]. We registered our protocol on the Open Science Framework [40].

Expert Consultations

After extracting information from peer-reviewed papers and tools identified via a Google search, we contacted experts in the field of digital gamified tools (eg, developers and researchers working on the topic in Canada and worldwide who were already known to the research team) to complement our online searches and ensure completeness. Specifically, we sent emails to 12 experts about the results of our searches and asked them to alert us to any games or papers we might have missed.

Data Charting

We developed a form in Microsoft Excel to guide the charting of data. We pretested and reviewed the form with team members to ensure we were accurately and adequately capturing relevant data. Data charting occurred independently with verification. Specifically, a reviewer (author HH) identified and screened all studies and digital gamified tools for their eligibility. Screening results were verified by a second reviewer (author DG). The data charting was then performed by a reviewer (HH) and again verified by a second reviewer (DG). Any conflicts throughout screening or data charting were resolved by a third reviewer (author ED). From the included papers, we charted data about (1) the type and design of study (developmental or evaluation study, user testing, randomized controlled trial, etc), (2) the vaccine(s) addressed (COVID-19, HPV, etc), (3) the purpose of the study or intervention, (4) the digital gamified tool platform (web based, native mobile app, mobile-enabled web app, virtual reality), and (5) the characteristics of study participants. For the evaluated interventions, we charted data about preselected outcomes that are widely used to predict health-related behaviors and to assess outcomes in studies of interventions about vaccination and immunization [11-14]. Specifically, we extracted data about the tools' effects on knowledge, attitudes, beliefs (perceived benefits, perceived risks), and behavioral intentions. Emotional, cultural, and social factors can also influence a decision about vaccination [29,30]. Therefore, we also extracted data about other outcomes that the studies may have evaluated. Because we sought to understand all possible effects, we did not prespecify any of these as a primary outcome.

We organized the extracted data in tables and synthesized them descriptively.

Quality Assessment

To assess the quality of the studies that evaluated their interventions, we used the Mixed Methods Appraisal Tool (MMAT) developed by Pluye et al [41]. Two reviewers independently conducted the quality assessment, resolving disagreements through discussion until reaching a consensus. A third and a fourth reviewer (authors HH and HW) intervened to settle any remaining conflicts.

Data Synthesis

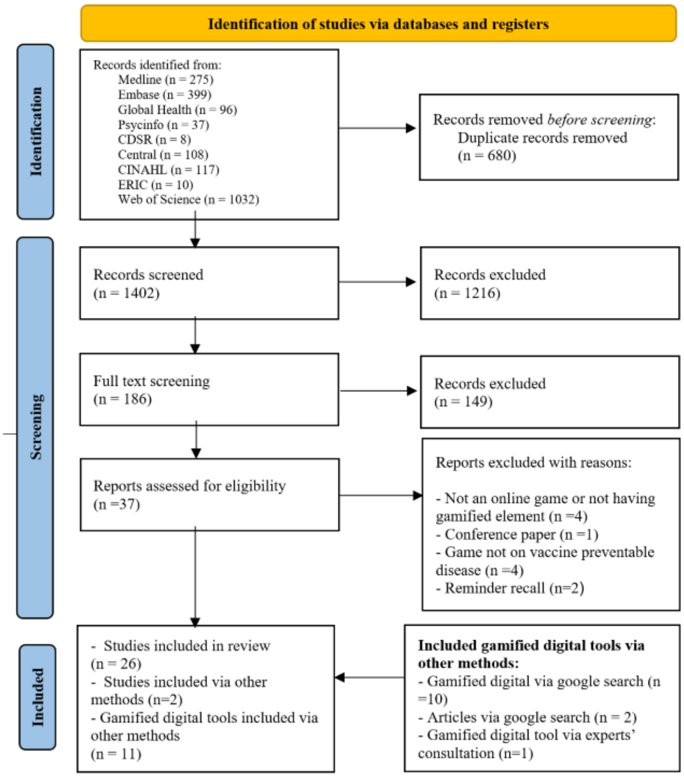
We summarized data using a narrative approach involving framework and content analysis. We classified each digital gamified tool platform using the 4 types of digital gamified tools: web-based tool, native mobile app, mobile-enabled web app, virtual reality tool. For the type of digital gamified tool, we classified web-based tools that explicitly noted their suitability for mobile use (eg, by smartphone or tablet) as mobile-enabled apps. We classified web-based tools without such an explicit statement as web based only, even though they may be functional on mobile devices. For the type and design of study, we grouped randomized designs together, including traditional randomized controlled trials with only 2 study arms and factorial designs with more than 2 study arms. Although these methods are not exactly the same, they all use randomization to minimize potential biases and are therefore functionally equivalent for our purposes of understanding what kinds of evaluations have been undertaken [42]. We summarized the main characteristics of tools, including PICO elements, in a tabular display. We used the PRISMA 2020 flowchart to describe the process of study selection [43].

Results

Papers Identified and Scope of Literature

We identified a total of 2082 records through database searches. After removing duplicates, we screened 1402 (67.3%) database records. Through Google searches, we identified 10 digital gamified tools and 2 papers. In a private browsing mode search, there was no change in search results. Of the 12 experts contacted, 2 (17%) responded and suggested 2 papers and 2 links, of which 1 (50%) digital gamified tool met the inclusion criteria and was included in our review. Through these methods, our final data set included 28 (2%) peer-reviewed papers and 11 digital gamified tools. Figure 1 shows our PRISMA diagram.

Figure 1. PRISMA flow diagram. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Of the 28 peer-reviewed papers, 7 (25%) explained the development of the tool, 16 (57%) described evaluation, and 2 (7%) reported both development and evaluation of the tool (Table 2). To report our results, we grouped studies together that reported the same tool, meaning 28 peer-reviewed papers reporting on 25 different tools. Of these 25 digital gamified tools, 11 (44%) were web-based tools, 7 (28%) mobile (native mobile or mobile-enabled web) apps, 6 (24%) virtual reality tools, and 1 (4%) offered in both mobile and web-based versions (for details, see Table 2). The most common single vaccines addressed in the tools were influenza (n=6, 24%, tools) and HPV (n=6, 24%, tools). Other tools addressed COVID-19 (n=2, 8%); measles, mumps, influenza, and smallpox (n=2, 8%); a hypothetical disease (n=2, 8%); other vaccine-preventable diseases (n=6, 24%); and the role vaccines play in preventing the spread of disease with no particular vaccine specified (n=1, 4%). Of the 10 digital gamified tools identified via a Google search and 1 suggested by the expert (a total of 11 digital gamified tools; see Table 3), the largest group (n=5, 45%) addressed COVID-19, and the rest were about other vaccine-preventable diseases. The 11 gamified elements identified in the Google search and expert feedback identified 6 types of gamified elements: reward points, serious games, physical trading cards, certificates, role-playing, and quizzes (see Table 3). The most common type was reward points, which appeared in 5 (45%) cases. Two cases used serious games, one case used physical trading cards and reward points, one case used certificates, one case used role-playing, and one case used quizzes. Additional characteristics of the studies included (eg, country of origin, sample size, participant characteristics) are detailed in Multimedia Appendix 3 [31,44-70]. The expanded versions of Table 2 [31,44-70] and Table 3 [71-81] are provided in Multimedia Appendix 4.

Table 2. General information about the studies.

Type of study and author(s)	Type of digital gamified tool platform	Type of disease/vaccine	Type and design of study (development or evaluation, iterative design, randomized controlled trial, etc)
Evaluation studies			
Betsch and Böhm [44]	Web-based tool	Hypothetical	Evaluation: online experiment
Carolan et al [45]	Web-based tool	Measles, mumps, influenza, and smallpox	Evaluation: pre-post study
Cates et al [31]	Web-based tool	HPV[a]	Evaluation: pilot randomized controlled trial
Dale et al [46]	Native mobile app	Influenza	Evaluation: nonrandomized trial
Darville et al [47]	Web-based tool	HPV	Evaluation: randomized controlled trial
Eley et al [48], McNulty et al [49]	Web-based tool	Bacteria, vaccine-preventable disease	Evaluation: quantitative followed by qualitative research design
Fadda et al [50], Fadda et al [51]	Native mobile app	MMR vaccines	Evaluation: mixed methods research design
Ibuka et al [52]	Web-based tool	Hypothetical disease	Evaluation: experimental design
Kaufman and Flanagan [53]	Web-based tool	Not reported	Evaluation: experimental design
Lee et al [54]	Native mobile app	Influenza	Evaluation: randomized controlled trial
Mitchell et al [55], Laplana [56]	Web-based tool	Influenza	Evaluation: pre-post study
Mottelson et al [57]	Virtual reality tool	COVID-19	Evaluation: randomized controlled trial (2×2 factorial design)
Nowak et al [58]	Virtual reality tool	Influenza	Evaluation: one-way between-subjects design with random assignment
Real et al [59]	Virtual reality tool	Influenza	Evaluation: quasi- randomized controlled trial[b]
Woodall et al [60]	Mobile-enabled web app	HPV	Evaluation: clinic-cluster randomized trial
Vandeweerdt et al [61]	Virtual reality tool	COVID-19	Evaluation: randomized controlled trial
Development studies			
Amresh et al [62]	Web-based tool	HPV	Development: iterative design
Bertozzi et al [63] (data extracted for the game related to vaccines)	Web-based tool	Influenza	Development: iterative design
Carolan et al [64]	Web-based tool	Measles, mumps, influenza, and smallpox	Development: iterative design
de Araujo Lima et al [66]	Native mobile app	Vaccine-preventable diseases	Development: heuristic evaluation by users, content evaluation by experts
Kafai et al [65]	Virtual reality	Dragon swooping cough virus to reflect real-life features of infectious viruses, such as Ebola.	Development: user feedback via surveys (asking users questions) and log files (observing user behaviors)
Real et al [67]	Native mobile app	HPV	Development: usability testing
Streuli et al [68]	Virtual reality	Pediatric vaccines	Development: Community-based participatory research and co-design
Development and evaluation studies			
Davies et al [69]	Mobile or web app (multiple formats available)	Hepatitis B	Development and evaluation: Participatory Action Research
Ruiz-López et al [70]	Native mobile app	HPV	Development and evaluation: Iterative design and evaluation via questionnaire

[a]HPV: human papillomavirus.

[b]Allocation to a study arm was performed according to work schedules, which are often arbitrary. We therefore considered this quasi-randomization.

Table 3. Tools from Google search and expert suggestions.

Digital gamified tool	Type of disease/vaccine	Type of digital gamified tool platform	Gamification elements (eg, rewards, role-playing, leaderboard, serious game)
Antidote COVID-19 [71]	COVID-19	Native mobile app	Reward points
The Vaccination Game [72]	H11N7 and influenza	Web-based tool	Serious game
Help take down COVID-zilla! [73]	COVID-19	Web-based tool	Role-playing
Just the Vax! [74]	Vaccine-preventable disease	Web-based tool	Reward points
COVID Invaders [75]	COVID-19	Web-based tool	Reward points
Vax Pack Hero [76]	Vaccine-preventable disease	Web-based tool	Reward points and physical trading cards
Flu's Clues [77]	Influenza	Web-based tool	Certificate of completion for solving the influenza mystery
Virus Fighter [78]	COVID-19, influenza, Ebola, measles	Web-based tool	Serious game
Immunization411: for preteens and teens' online training [79]	Tdap meningococcal vaccine, varicella, HPV[a], influenza	Web-based tool	Reward points
COVID Chronicles [80]	COVID-19	Web-based tool	Reward points
I Boost[b] [81]	Vaccine-preventable disease	Web-based tool	Quiz

[a]HPV: human papillomavirus.

[b]Suggested by an expert.

The studies were conducted in 26 different countries, with the majority of studies coming from the United States (n=13, 46%, studies) and the United Kingdom (n=5, 18%, studies). Study populations included students at various levels (elementary school to college, specialty programs, eg, nursing and pediatric residency), parents of vaccine-eligible children, adults from the general population, members of particular sociocultural communities (eg, immigrants, Indigenous peoples), and convenience samples, such as players of a game, attendees of a conference, and employees of an organization. Sample sizes ranged from 8 to 50,286. Whenever papers reported study participant characteristics such as age, sex, gender, ethnocultural identity, or socioeconomic levels, we extracted summary data, as shown in Multimedia Appendix 3.

Reported Effects of Evaluated Interventions

In total, 18 (64%) of 28 studies evaluated at least 1 of our outcomes of interest, while 11 (39%) studies reported the effects of the evaluated interventions on more than 1 outcome of interest. Summarized outcomes and their MMAT quality assessments are shown in Table 4. Multimedia Appendix 5 provides full details.

Table 4. Outcomes of evaluation studies included.

Type of digital gamified tool platform and study	Knowledge (comprehension/understanding, etc)	Attitudes (for/against vaccination, etc)	Beliefs (risk perceptions, etc)	Behavioral intentions (getting vaccinated or not, etc)	Others (eg, emotions)	MMAT[a] quality score
Web-based tool						
Betsch and Böhm [44]	—[b]	Negative vaccine attitudes with compulsory vaccination	—	Decreased vaccine uptake with compulsory vaccination	Increased level of anger with compulsory vaccination	60% quality criteria met
Carolan et al [45]	—	No significant effect on attitudes towards vaccination	—	—	Increased confidence in information needs	80% quality criteria met
Cates et al [31]	Increase in knowledge about immunization	—	—	Positive increase in intentions to vaccinate	Increase in vaccination self-efficacy, decisional balance towards vaccination	100% quality criteria met
Darville et al [47]	—	—	Positive effects on beliefs towards vaccination	Increase in intentions to vaccinate	—	60% quality criteria met
Eley et al [48], McNulty et al [49]	Improvements in knowledge about immunization	—	—	—	—	100% quality criteria met
Ibuka et al [52]	—	—	—	—	Free riding in vaccination decisions decreases vaccine acceptance	80% quality criteria met
Kaufman and Flanagan [53]	The digital version of the game was less effective at facilitating learning	The digital version of the game was less effective at attitude change	—	—	The digital version of the game was perceived to be complicated to use	20% quality criteria met
Mitchell et al [55], Laplana [56]	Increase in knowledge	Positive increase in attitudes for vaccination	—	Increase in vaccine uptake after accessing the game	—	80% quality criteria met (Mitchell et al [55])
Mobile app						
Dale et al [46]	—	—	—	Positive increase in intentions to vaccinate	—	80% quality criteria met
Fadda et al [50], Fadda et al [51]	Improvements in knowledge about immunization	—	—	Increase in intentions to vaccinate	Increase in psychological empowerment and confidence in the decision	80% quality criteria met (Fadda et al [50], Fadda et al [51])
Lee et al [54]	—	—	—	Increase in intentions to vaccinate	—	80% quality criteria met
Woodall et al [60]	—	—	Increase in beliefs towards vaccination	Increase in intentions to vaccinate	Increase in vaccine confidence	40% quality criteria met
Ruiz-López et al [70]	Increase in knowledge after playing the game	—	—	—	—	100% quality criteria met
Virtual reality tool						

Type of digital gamified tool platform and study	Knowledge (comprehension/understanding, etc)	Attitudes (for/against vaccination, etc)	Beliefs (risk perceptions, etc)	Behavioral intentions (getting vaccinated or not, etc)	Others (eg, emotions)	MMAT[a] quality score
Mottelson et al [57]	—	—	—	Increase in vaccination intention when both the personal and collective benefit of COVID-19 vaccination was communicated	Increase in COVID-19 empathy, vaccination recommendation, and vaccination readiness	80% quality criteria met
Nowak et al [58]	—	—	Positive effects on beliefs towards vaccination	Increase in intentions to vaccinate	—	100% quality criteria met
Real et al [59]	—	Increase in attitudes in favour of vaccination	—	—	—	60% quality criteria met
Vandeweerdt et al [61]	—	—	—	Increase in intentions to vaccinate	Virtual reality intervention increases a sense of collective responsibility	100% quality criteria met
Mobile or web app (multiple formats available)						
Davies et al [69]	Increase in knowledge about immunization	—	—	—	—	80% quality criteria met

[a]MMAT: Mixed Methods Appraisal Tool.

[b]Not reported.

Effects on Knowledge (Includes Comprehension/Understanding, etc)

Overall, the 28 included studies suggested that digital gamified tools may positively influence knowledge. Of 7 (25%) studies that assessed knowledge, 6 (86%) showed an increase in knowledge about immunization in general [31,48,51,55,69,70]. All these 6 (86%) studies were of high quality (≥80%). One study of low quality (≤25%) reported that a digital game is less effective at increasing knowledge compared to its original board game format [53]. When considering only the high-quality (≥80%) studies, we observed that digital gamified tools are associated with increased knowledge.

Effects on Attitudes (for or Against Vaccination)

Overall, digital gamified tools appeared to have mixed effects on attitudes toward vaccination. Of 5 (18%) of 28 studies that assessed attitudes, 2 (40%), one of high quality (≥80%) and the other of medium quality (60%), showed an increase in positive attitudes toward vaccination [55,59]. In addition, 2 (40%) studies, one of high quality (≥80%) and the other of low quality (20%), reported no or less effect on attitudes toward vaccination [45,53], and 1 (20%) study comparing voluntary and compulsory vaccines in a game context showed negative attitudes regarding compulsory vaccination [44]. When considering only the high-quality (≥80%) studies, we observed inconsistent effects on attitudes.

Effects on Beliefs (Perceived Benefits, Perceived Risks)

Overall, digital gamified tools demonstrated positive effects on beliefs toward vaccination. In total, 3 (11%) of 28 studies, 1 (33%) of high quality (100%) and 2 (67%) of medium quality (60% and 40%), evaluated the effects of digital gamified tools on beliefs toward vaccination. All 3 (100%) studies showed positive effects on beliefs toward vaccination [47,58,60]. When considering only the high-quality (≥80%) studies, we observed that digital gamified tools are associated with positive beliefs about vaccines.

Effects on Behavioral Intentions

Overall, the 28 included studies suggested that digital gamified tools may positively influence intentions to receive vaccines. In total, 11 (39%) studies evaluated the effects of digital gamified tools on behavioral intentions with regard to vaccines. Of these 11 studies, 1 (9%) of medium quality (60%) showed a decrease in vaccination intention when compulsory vaccination was introduced within a game context [44], whereas 10 (91%) studies, 3 (30%) of medium quality (60% and 40%) and 7 (70%) of high quality (≥75%), showed increased intentions to vaccinate [31,46,47,51,54,55,57,58,60,61]. When considering only the high-quality (≥80%) studies, digital gamified tools appeared to be consistently associated with increased vaccination intention.

Other Outcomes

In total, 9 (32%) of 28 studies have also evaluated the effects of digital gamified tools on other outcomes. Of these, 4 (44%) studies reported an increase in confidence in vaccines (medium quality=40%) [60], confidence in information needs (high quality=80%) [45], decisional balance in support of vaccination (high quality=100%) [31], and confidence in vaccine decisions (high quality=80%) [50]. In addition, 1 (11%) study of high quality (80%) reported an increase in empathy toward those vulnerable to COVID-19 and vaccination recommendations [57], and 2 (22%) studies of high quality (100% and 80%) reported an increase in vaccination self-efficacy and readiness [31,57]. An increase in psychological empowerment (high quality=80%) [51] and in emotions such as anger toward compulsory vaccination (medium quality=60%) [44] was also reported by 2 (22%) studies. One study of high quality (80%) reported that the concept of free riding decreases vaccine acceptance [52], whereas another study of high quality (100%) reported that virtual reality intervention increases collective responsibility [61]. When considering only the high-quality (≥80%) studies, we observed a variety of positive effects associated with digital gamified tools, including confidence in vaccines, confidence in decisions about vaccines, empathy toward vulnerable people, collective responsibility, psychological empowerment, and vaccination self-efficacy and readiness.

Effects of the Platform (Web Based, Mobile, Virtual Reality)

The study designs of the 28 included papers did not permit us to formally compare the effects of different platforms in a robust way. Upon inspection, there did not appear to be a strong effect of the platform. In other words, we did not observe evidence in favor of web-based, mobile, or virtual reality apps over the other 2 types of apps.

Discussion

Principal Findings

The broad objective of this scoping review was to map the state of the science regarding digital gamified tools and their effects. In other words, we wished to answer a common question at the intersection of public health and digital health: does gamification encourage vaccination and influence knowledge, attitudes, beliefs, and behaviors related to vaccination? By mapping both published literature and tools currently available online, we observed 2 principal findings.

First, our results suggest that gamification can increase predictors of vaccine uptake, such as knowledge, attitudes, beliefs, behaviors, and vaccination intention. This finding is similar to the findings of a previous review by Montagni et al [82] suggesting that gamification can contribute to changed behaviors and improved knowledge of vaccination. Similarly, other reviews have suggested the potential benefits of gamification for non–vaccination-related behavior change, such as a systematic review suggesting that gamification interventions could be a feasible way to improve health-related outcomes among cancer survivors [83] and another review suggesting their effectiveness in improving physical activity [84]. Such previous work became even more relevant during the COVID-19 pandemic, as many jurisdictions sought to optimize vaccine uptake in the context of an “infodemic” (ie, overabundance of information, true, false, and misleading, about the pandemic and recommended preventive behaviors) [85]. Half of the digital gamified tools identified in our web search addressed COVID-19, suggesting an active interest in using a gamified approach in the pandemic context. Recent research by Plechatá et al [86] published after our data extraction steps were complete suggested that explaining the concept of herd immunity with gamification has a positive impact on the COVID-19 vaccination intention.

Second, our review suggests that although gamification has the potential to enhance the impact of education strategies, gamified tools alone may not wholly address gaps in vaccine acceptance and uptake. Although some of the identified tools did increase vaccination, the increases did not fully close gaps between previous and desired vaccine uptake. This finding aligns with those of Tozzi et al [87], which suggested that promising results could be achieved by combining gamification with educative and informative tools to improve immunization programs. This finding also aligns with previous reviews suggesting the use of digital gamified interventions as a public health tool of interest in enhancing vaccine uptake [82,88]. Further research published by Real et al [89] after our systematic search similarly observed that integrating gamification, such as virtual reality, in training modules enhances uptake of the HPV vaccine. Integrating gamified features may work because they make digital tools acceptable and more fun to use and may reduce the chances of people feeling pushed toward vaccination. In parallel, gamification may be a promising strategy for increasing knowledge, skills, and confidence among health professionals engaging in discussions about vaccines with their patients [90,91].

In addition to these findings drawn directly from our review of the included tools, we offer a broader observation based on the contents of this scoping review, along with the larger landscape of vaccine acceptance research: context is key. Although an engaging approach may work for some groups or in some situations, it may be less well accepted among other groups and in other situations. For instance, a casual and approachable style of communication will work for the younger audience to convey vaccine information but might be deemed insufficient to health care professionals in a more formal setting, such as hospitals. A good understanding of the factors associated with low vaccine acceptance at the local level is needed prior to developing gamified tools [92]. Future research in this area should consider possible contextual factors, such as local culture, social and demographic characteristics of users, and different influences on vaccine hesitancy and acceptance in different regions. To help better match games to the context(s) in which they will be played, when developing games, developers and researchers may wish to consider involving potential players from different contexts early and often. This aligns with previous work [93,94] suggesting that involving users earlier in developing tools may help in designing interventions suitable for a targeted context. One of the examples in our review was an intervention by Cates et al [31] designed to explain HPV vaccines to teenagers using a “secret garden” theme. Involving potential game players early in the development of the game may have contributed toward its positive effects on vaccination intention.

The implications of this research extend beyond the immediate reported effects of gamified tools and delve into the strategic dimensions of public health policy and communication efforts. Considering the insights gleaned from the findings, this study supports a comprehensive and well-informed approach to integrating gamification into strategies for promoting vaccination. As gamification continues to demonstrate its potential in enhancing vaccine uptake, it is crucial to navigate this terrain thoughtfully, considering the various factors that influence its impact. This includes not only the technological and behavioral aspects but also the larger sociocultural context in which vaccination decisions are made. Therefore, our study emphasizes the importance of a comprehensive approach that fosters a mutually beneficial relationship between technological innovation, evidence-based strategies, and an intricate understanding of local contexts. This approach has the potential to make gamification a sustainable and adaptable tool in the arsenal of public health interventions, rather than just a passing trend.

The review does not find a clear advantage for any platform in terms of reported effects. It was challenging to measure the impact of the platforms on behavioral outcomes and calls for more focused research to better understand the specific elements within each platform that drive behavior change. In essence, our study suggests that the reported effects of an app may not be solely determined by its platform but rather by the strategic incorporation of mechanics and elements that facilitate the desired behavior change.

Gamification can influence knowledge, attitudes, and beliefs about vaccines, which can affect vaccine uptake. This is consistent with theories of change proposing that cognitive changes can lead to behavioral outcomes. Although our study mainly examines the immediate effects of gamification on these cognitive aspects, it also offers some implications for using gamification as a potentially viable strategy to improve vaccine acceptance.

Strengths and Limitations

Our study has 5 main limitations. First, because we aimed to capture all relevant evidence and examples, as is typical in a scoping review, we included a broad range of study designs and did not draw conclusions about the relative advantages or disadvantages of different game platforms and features. Given the rapid growth within this field of research, it would be difficult to truly prioritize evidence according to quality criteria at this point. In the future, it may be possible to conduct a systematic review and meta-analysis, restricting included studies to randomized experiments or randomized controlled trials. Such future work may include approaches such as a network meta-analysis to allow for comparison of the effects of different game types or game features. Based on the existing literature, it is difficult to conclude whether certain games are more or less likely to achieve their aims. Second, our results may be influenced by publication bias. It is possible that groups that have developed digital gamified tools that showed disappointing results simply did not publish their studies. This bias could lead to an overestimation of the reported effects of these tools. This highlights the importance of further research to fully understand the real impact of these tools and thus accurately inform policy decisions about the development and use of these tools. Third, and related to the previous 2 points, the rapid growth in this area may mean that we missed more recent evidence in literature published after January 2022 and web searches after May 2022. Fourth, the majority of digital gamified tools on vaccination represented in publications and online were developed in high-income countries. This finding aligns with the findings of previous work by Ohannessian et al [88], who also reported a predominance of high-income countries. This may reflect more widespread internet access and resources for developing digital gamified tools in high-income countries. It may also reflect publication bias in the scientific literature (ie, there may be fewer papers written about digital gamified tools in lower-income countries) and online (ie, tools developed and published in lower-income countries may not be ranked highly by search engines and therefore may not have appeared in our web searches). Tools developed in lower-income countries may also take different forms; for example, they may be text message–based interventions (with or without gamification) rather than web-based tools and therefore would be less likely to be identified in web searches. Analog games from high-income countries were similarly excluded from the scope of our study [95]. Nondigital games, such as board and card games, have demonstrated positive impacts on educational knowledge, cognitive function, and social interactions [96,97]. Such games can support diverse learning across subjects and settings, fostering interactions that develop skills, such as computational thinking and teamwork, and have positive impacts on academic achievement and vocabulary acquisition compared to digital games [97-99]. We restricted our scoping review to digital gamified tools because the review was intended to provide an evidence base for digital game development. Although nondigital games are also potentially useful interventions, the implementation and distribution of such interventions is more challenging, especially in a geographically dispersed country, such as Canada. Fifth, and finally, as we used Google and private browsing in Google, there may be a possibility that different search engines would provide different results.

This study also has 2 main strengths. First, by systematically examining the current literature and currently available tools online, we were able to offer an updated overview of the potential effects of including gamification in digital tools about vaccination. Second, by conducting a scoping review to broadly map the literature, future work can more easily identify and select key outcomes for systematic reviews and meta-analyses in this domain.

Conclusion

Digital gamified tools have the potential to improve vaccine uptake by increasing knowledge and promoting positive attitudes, beliefs, behaviors, and vaccination intention. Further evaluations of these innovative digital tools, including head-to-head comparisons of different features and different platforms, will add more knowledge about what works and what does not in order to achieve public health goals more efficiently. In the wider context of health policy, digital gamified tools may be useful components of multifaceted strategies to improve vaccination rates throughout society.

Acknowledgments

This work was supported by a research grant from the Canadian Institutes of Health Research (GA3177725). The authors gratefully acknowledge the assistance of Frédéric Bergeron (librarian) for reviewing the search strategy. We appreciate the assistance of Charles Racine and Créscence Joëlle Mefou Tasong in conducting the quality appraisal of the studies that evaluated their interventions.

Authors' Contributions

All authors provided substantial contributions to this paper’s conception and edits and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 20 KB - games_v12i1e47257_app1.docx]

Multimedia Appendix 2
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 498 KB - games_v12i1e47257_app2.pdf]

Multimedia Appendix 3
Characteristics of the studies included in the review.
[DOCX File , 25 KB - games_v12i1e47257_app3.docx]

Multimedia Appendix 4
Expanded version of Table 2 (general information about the studies) and Table 3 (tools from Google search and expert suggestions).
[DOCX File , 28 KB - games_v12i1e47257_app4.docx]

Multimedia Appendix 5
MMAT quality assessment. MMAT: Mixed Methods Appraisal Tool.
[XLSX File (Microsoft Excel File), 92 KB - games_v12i1e47257_app5.xlsx]

References

1. Omer SB, Salmon DA, Orenstein WA, deHart MP, Halsey N. Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. N Engl J Med 2009 May 07;360(19):1981-1988. [doi: 10.1056/nejmsa0806477]
2. Phadke VK, Bednarczyk RA, Salmon DA, Omer SB. Association between vaccine refusal and vaccine-preventable diseases in the United States: a review of measles and pertussis. JAMA 2016 Mar 15;315(11):1149-1158 [FREE Full text] [doi: 10.1001/jama.2016.1353] [Medline: 26978210]
3. Plans-Rubió P. Low percentages of measles vaccination coverage with two doses of vaccine and low herd immunity levels explain measles incidence and persistence of measles in the European Union in 2017-2018. Eur J Clin Microbiol Infect Dis 2019 Sep 10;38(9):1719-1729. [doi: 10.1007/s10096-019-03604-0] [Medline: 31289949]
4. Ten health issues WHO will tackle this year. World Health Organization. URL: http://tinyurl.com/39cd7tuf [accessed 2021-10-04]
5. Attwell K, Betsch C, Dubé E, Sivelä J, Gagneur A, Suggs LS, et al. Increasing vaccine acceptance using evidence-based approaches and policies: insights from research on behavioural and social determinants presented at the 7th Annual Vaccine Acceptance Meeting. Int J Infect Dis 2021 Apr;105:188-193 [FREE Full text] [doi: 10.1016/j.ijid.2021.02.007] [Medline: 33578012]
6. MacDonald NE, Dube E, Comeau JL. Have vaccine hesitancy models oversimplified a complex problem to our detriment? The Adapted Royal Society of Canada vaccine uptake framework. Vaccine 2022 Jun 23;40(29):3927-3930 [FREE Full text] [doi: 10.1016/j.vaccine.2022.05.052] [Medline: 35637069]
7. Dubé E, Gagnon D, MacDonald N, Bocquier A, Peretti-Watel P, Verger P. Underlying factors impacting vaccine hesitancy in high income countries: a review of qualitative studies. Expert Rev Vaccines 2018 Nov;17(11):989-1004. [doi: 10.1080/14760584.2018.1541406] [Medline: 30359151]
8. MacDonald NE, Comeau J, Dubé, Graham J, Greenwood M, Harmon S, et al. Royal society of Canada COVID-19 report: enhancing COVID-19 vaccine acceptance in Canada. FACETS 2021 Jan 01;6(1):1184-1246. [doi: 10.1139/facets-2021-0037]

9. Dolu İ, Turhan Z, Yalnız Dilcen H. COVID-19 vaccine acceptance is associated with vaccine hesitancy, perceived risk and previous vaccination experiences. Disaster Med Public Health Prep 2021 Dec 23;17:e97 [FREE Full text] [doi: 10.1017/dmp.2021.370] [Medline: 34937599]
10. Behavioural and social drivers of vaccination: tools and practical guidance for achieving high uptake. World Health Organization. URL: http://tinyurl.com/3uteanyc [accessed 2023-01-10]
11. Dubé E, Gagnon D, MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Strategies intended to address vaccine hesitancy: review of published reviews. Vaccine 2015 Aug 14;33(34):4191-4203 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.041] [Medline: 25896385]
12. Omer SB, Benjamin RM, Brewer NT, Buttenheim AM, Callaghan T, Caplan A, et al. Promoting COVID-19 vaccine acceptance: recommendations from the Lancet Commission on Vaccine Refusal, Acceptance, and Demand in the USA. Lancet 2021 Dec 11;398(10317):2186-2192 [FREE Full text] [doi: 10.1016/S0140-6736(21)02507-1] [Medline: 34793741]
13. Mills M, Rahal C, Brazel D, Yan J, Gieysztor S. COVID-19 vaccine deployment: behaviour, ethics, misinformation and policy strategies. The Royal Society; the British Academy. URL: http://tinyurl.com/2bc87r5f [accessed 2024-02-10]
14. Oehler RL. On measles, vaccination, social media activism and how to win back our role as our patients' best advocates. Clin Infect Dis 2020 Jan 02;70(2):338-340. [doi: 10.1093/cid/ciz656] [Medline: 31309980]
15. Tasnim S, Hossain M, Mazumder H. Impact of rumors or misinformation on coronavirus disease (COVID-19) in social media. SocArXiv preprint posted online 2020 [FREE Full text] [doi: 10.31235/osf.io/uf3zn]
16. Witteman HO, Zikmund-Fisher BJ. The defining characteristics of Web 2.0 and their potential influence in the online vaccination debate. Vaccine 2012 May 28;30(25):3734-3740. [doi: 10.1016/j.vaccine.2011.12.039] [Medline: 22178516]
17. Haase N, Schmid P, Betsch C. Impact of disease risk on the narrative bias in vaccination risk perceptions. Psychol Health 2020 Mar;35(3):346-365. [doi: 10.1080/08870446.2019.1630561] [Medline: 31480866]
18. Kaufman J, Ryan R, Walsh L, Horey D, Leask J, Robinson P, et al. Face-to-face interventions for informing or educating parents about early childhood vaccination. Cochrane Database Syst Rev 2018 May 08;5(5):CD010038 [FREE Full text] [doi: 10.1002/14651858.CD010038.pub3] [Medline: 29736980]
19. Lendacki FR, Forst LS, Mehta SD, Kerins JL. COVID-19 vaccination requirements, encouragement and hesitancy among non-health care, non-congregate workers in Chicago: results from the WEVax survey. BMC Public Health 2023 May 25;23(1):951 [FREE Full text] [doi: 10.1186/s12889-023-15781-x] [Medline: 37231367]
20. Golden SD, Moracco KE, Feld AL, Turner KL, DeFrank JT, Brewer NT. Process evaluation of an intervention to increase provision of adolescent vaccines at school health centers. Health Educ Behav 2014 Dec;41(6):625-632. [doi: 10.1177/1090198114531773] [Medline: 24786792]
21. Kharbanda EO, Stockwell MS, Fox HW, Andres R, Lara M, Rickert VI. Text message reminders to promote human papillomavirus vaccination. Vaccine 2011 Mar 21;29(14):2537-2541. [doi: 10.1016/j.vaccine.2011.01.065] [Medline: 21300094]
22. Shourie S, Jackson C, Cheater F, Bekker H, Edlin R, Tubeuf S, et al. A cluster randomised controlled trial of a web based decision aid to support parents' decisions about their child's measles mumps and rubella (MMR) vaccination. Vaccine 2013 Dec 05;31(50):6003-6010 [FREE Full text] [doi: 10.1016/j.vaccine.2013.10.025] [Medline: 24148574]
23. Andriani Y, Rusmil K, Akbar I. Measles-rubella immunization health education using animated videos and text messages via WhatsApp. Midwifery. 2020. URL: https://www.academia.edu/download/84881544/52501-208272-1-PB.pdf [accessed 2024-02-10]
24. Groh F. Gamificationtate of the art definition and utilization. 2012 Presented at: RTMI'12: 4th Seminar on Research Trends in Media Informatics; February 14, 2012; Ulm, Baden-Württemberg p. 39-46.
25. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014 Aug 04;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
26. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
27. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
28. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
29. Sawyer B, Rejeski D. Serious Games: Improving Public Policy through Game-Based Learning and Simulation. Washington, DC: Woodrow Wilson International Center for Scholars; 2002.
30. Alvarez J, Djaouti D. An introduction to serious game - definitions and concepts. In: Fauquet-Alekhine P, Soler L, editors. Serious Games & Simulation for Risks Management. Paris: LARSEN Science; 2011:11-15.
31. Cates JR, Fuemmeler BF, Stockton LL, Diehl SJ, Crandell JL, Coyne-Beasley T. Evaluation of a serious video game to facilitate conversations about human papillomavirus vaccination for preteens: pilot randomized controlled trial. JMIR Serious Games 2020 Dec 03;8(4):e16883 [FREE Full text] [doi: 10.2196/16883] [Medline: 33270028]
32. de Carvalho CV, Coelho A. Game-based learning, gamification in education and serious games. Computers 2022 Mar 04;11(3):36. [doi: 10.3390/computers11030036]

33. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. Front Psychiatry 2016;7:215 [FREE Full text] [doi: 10.3389/fpsyt.2016.00215] [Medline: 28119636]
34. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
35. Hargrave DR, Hargrave UA, Bouffet E. Quality of health information on the internet in pediatric neuro-oncology. Neuro Oncol 2006 Apr;8(2):175-182 [FREE Full text] [doi: 10.1215/15228517-2005-008] [Medline: 16533758]
36. Petrescu P. Google organic click-through rates in 2014. MOZ Blog. 2014. URL: https://moz.com/blog/google-organic-click-through-rates-in-2014 [accessed 2024-02-10]
37. van der Marel S, Duijvestein M, Hardwick JC, van den Brink GR, Veenendaal R, Hommes DW, et al. Quality of web-based information on inflammatory bowel diseases. Inflamm Bowel Dis 2009 Dec;15(12):1891-1896. [doi: 10.1002/ibd.20976] [Medline: 19462423]
38. Resnick P, Garrett R, Kriplean T, Munson S, Stroud N. Bursting your (filter) bubble: strategies for promoting diverse exposure. 2013 Presented at: CSCW '13: Computer Supported Cooperative Work; February 23-27, 2013; San Antonio, TX. [doi: 10.1145/2441955.2441981]
39. PRISMA for scoping reviews. PRISMA. URL: http://tinyurl.com/23j8b9aw [accessed 2024-02-10]
40. Gamification tool- enviromental scan. OSF Registries. URL: https://osf.io/nv8af [accessed 2024-02-09]
41. Pluye P, Gagnon M, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. Int J Nurs Stud 2009 Apr;46(4):529-546. [doi: 10.1016/j.ijnurstu.2009.01.009] [Medline: 19233357]
42. Montgomery AA, Peters TJ, Little P. Design, analysis and presentation of factorial randomised controlled trials. BMC Med Res Methodol 2003 Nov 24;3:26 [FREE Full text] [doi: 10.1186/1471-2288-3-26] [Medline: 14633287]
43. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
44. Betsch C, Böhm R. Detrimental effects of introducing partial compulsory vaccination: experimental evidence. Eur J Public Health 2016 Jun;26(3):378-381. [doi: 10.1093/eurpub/ckv154] [Medline: 26297722]
45. Carolan K, Verran J, Crossley M, Redfern J, Whitton N, Amos M. Impact of educational interventions on adolescent attitudes and knowledge regarding vaccination: a pilot study. PLoS One 2018;13(1):e0190984 [FREE Full text] [doi: 10.1371/journal.pone.0190984] [Medline: 29351325]
46. Dale LP, White L, Mitchell M, Faulkner G. Smartphone app uses loyalty point incentives and push notifications to encourage influenza vaccine uptake. Vaccine 2019 Jul 26;37(32):4594-4600. [doi: 10.1016/j.vaccine.2018.04.018] [Medline: 29699784]
47. Darville G, Anderson–Lewis C, Stellefson M, Lee Y, MacInnes J, Pigg R, et al. Customization of avatars in a HPV digital gaming intervention for college-age males: an experimental study. Simul Gaming 2018 Oct 17;49(5):515-537. [doi: 10.1177/1046878118799472]
48. Eley CV, Young VL, Hayes CV, Verlander NQ, McNulty CAM. Young people's knowledge of antibiotics and vaccinations and increasing this knowledge through gaming: mixed-methods study using e-Bug. JMIR Serious Games 2019 Feb 01;7(1):e10915 [FREE Full text] [doi: 10.2196/10915] [Medline: 30707096]
49. McNulty CAM, Lecky DM, Farrell D, Kostkova P, Adriaenssens N, Koprivová Herotová T, e-Bug Working Group. Overview of e-Bug: an antibiotic and hygiene educational resource for schools. J Antimicrob Chemother 2011 Jun;66(Suppl 5):v3-12 [FREE Full text] [doi: 10.1093/jac/dkr119] [Medline: 21680584]
50. Fadda M, Galimberti E, Fiordelli M, Romanò L, Zanetti A, Schulz PJ. Effectiveness of a smartphone app to increase parents' knowledge and empowerment in the MMR vaccination decision: a randomized controlled trial. Hum Vaccin Immunother 2017 Nov 02;13(11):2512-2521 [FREE Full text] [doi: 10.1080/21645515.2017.1360456] [Medline: 29125783]
51. Fadda M, Galimberti E, Fiordelli M, Schulz PJ. Evaluation of a mobile phone–based intervention to increase parents' knowledge about the measles-mumps-rubella vaccination and their psychological empowerment: mixed-method approach. JMIR Mhealth Uhealth 2018 Mar 07;6(3):e59 [FREE Full text] [doi: 10.2196/mhealth.8263] [Medline: 29514772]
52. Ibuka Y, Li M, Vietri J, Chapman GB, Galvani AP. Free-riding behavior in vaccination decisions: an experimental study. PLoS One 2014;9(1):e87164 [FREE Full text] [doi: 10.1371/journal.pone.0087164] [Medline: 24475246]
53. Kaufman, Flanagan. Lost in translation: comparing the impact of an analog and digital version of a public health game on players? perceptions, attitudes, and cognitions. Int J Gaming Comput Mediat Simul 2013 Jul;5(3):1-9 [FREE Full text] [doi: 10.4018/jgcms.2013070101]
54. Lee W, Stück D, Konty K, Rivers C, Brown CR, Zbikowski SM, et al. Large-scale influenza vaccination promotion on a mobile app platform: a randomized controlled trial. Vaccine 2020 Apr 16;38(18):3508-3514. [doi: 10.1016/j.vaccine.2019.11.053] [Medline: 31787410]
55. Mitchell G, Leonard L, Carter G, Santin O, Brown Wilson C. Evaluation of a 'serious game' on nursing student knowledge and uptake of influenza vaccination. PLoS One 2021 Jan 14;16(1):e0245389 [FREE Full text] [doi: 10.1371/journal.pone.0245389] [Medline: 33444348]

56. Laplana JP. BJN Awards 2019: innovation runner up: creating a staff flu vaccination game. Br J Nurs 2019 Aug 08;28(15):1032-1034. [doi: 10.12968/bjon.2019.28.15.1032] [Medline: 31393758]
57. Mottelson A, Vandeweerdt C, Atchapero M, Luong T, Holz C, Böhm R, et al. A self-administered virtual reality intervention increases COVID-19 vaccination intention. Vaccine 2021 Nov 05;39(46):6746-6753 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.004] [Medline: 34654579]
58. Nowak GJ, Evans NJ, Wojdynski BW, Ahn SJG, Len-Rios ME, Carera K, et al. Using immersive virtual reality to improve the beliefs and intentions of influenza vaccine avoidant 18-to-49-year-olds: considerations, effects, and lessons learned. Vaccine 2020 Jan 29;38(5):1225-1233. [doi: 10.1016/j.vaccine.2019.11.009] [Medline: 31806533]
59. Real FJ, DeBlasio D, Beck AF, Ollberding NJ, Davis D, Cruse B, et al. A virtual reality curriculum for pediatric residents decreases rates of influenza vaccine refusal. Acad Pediatr 2017;17(4):431-435. [doi: 10.1016/j.acap.2017.01.010] [Medline: 28126612]
60. Woodall WG, Zimet G, Kong A, Buller D, Reither J, Chilton L, et al. Vacteens.org: a mobile web app to improve HPV vaccine uptake. Front Digit Health 2021;3:693688 [FREE Full text] [doi: 10.3389/fdgth.2021.693688] [Medline: 34713171]
61. Vandeweerdt C, Luong T, Atchapero M, Mottelson A, Holz C, Makransky G, et al. Virtual reality reduces COVID-19 vaccine hesitancy in the wild: a randomized trial. Sci Rep 2022 Mar 17;12(1):4593 [FREE Full text] [doi: 10.1038/s41598-022-08120-4] [Medline: 35301359]
62. Amresh A, Chia-Chen A, Baron C. A game based intervention to promote HPV vaccination among adolescents. 2019 Presented at: SeGAH 2019: IEEE 7th International Conference on Serious Games and Applications for Health; August 5-7, 2019; Kyoto, Japan. [doi: 10.1109/segah.2019.8882459]
63. Bertozzi E, Krilov L, Walker D. Successful game development partnerships between academics and physicians: two case studies. IJGCMS 2013 Jul;5(3):97. [doi: 10.4018/jgcms.2013070107]
64. Carolan K, Verran J, Amos M, Crossley M, Redfern J, Whitton N, et al. SimFection: a digital resource for vaccination education. J Biol Educ 2018 May 24;53(2):225-234. [doi: 10.1080/00219266.2018.1469534]
65. Kafai Y, Fields D, Giang M, Fefferman N, Sun J, Kunka D, et al. Designing for massive engagement in a tween community: participation, prevention, and philanthropy in a virtual epidemic. 2017 Presented at: IDC '17: 2017 Conference on Interaction Design and Children; June 27-30, 2017; Stanford, CA. [doi: 10.1145/3078072.3079730]
66. de Araujo Lima ID, de Leon CGRMP, Ribeiro LM, da Silva ICR, Vilela DM, Fonseca LMM, et al. A serious game (Immunitates) about immunization: development and validation study. JMIR Serious Games 2022 Feb 18;10(1):e30738 [FREE Full text] [doi: 10.2196/30738] [Medline: 35179496]
67. Real FJ, Rosen BL, Bishop JM, McDonald S, DeBlasio D, Kreps GL, et al. Usability evaluation of the novel smartphone application, HPV Vaccine: Same Way, Same Day, among pediatric residents. Acad Pediatr 2021;21(4):742-749. [doi: 10.1016/j.acap.2020.11.023] [Medline: 33279739]
68. Streuli S, Ibrahim N, Mohamed A, Sharma M, Esmailian M, Sezan I, et al. Development of a culturally and linguistically sensitive virtual reality educational platform to improve vaccine acceptance within a refugee population: the SHIFA community engagement-public health innovation programme. BMJ Open 2021 Sep 14;11(9):e051184 [FREE Full text] [doi: 10.1136/bmjopen-2021-051184] [Medline: 34521673]
69. Davies J, Bukulatjpi S, Sharma S, Caldwell L, Johnston V, Davis JS. Development of a culturally appropriate bilingual electronic app about hepatitis B for indigenous Australians: towards shared understandings. JMIR Res Protoc 2015 Jun 10;4(2):e70 [FREE Full text] [doi: 10.2196/resprot.4216] [Medline: 26063474]
70. Ruiz-López T, Sen S, Jakobsen E, Tropé A, Castle PE, Hansen BT, et al. FightHPV: design and evaluation of a mobile game to raise awareness about human papillomavirus and nudge people to take action against cervical cancer. JMIR Serious Games 2019 Apr 08;7(2):e8540 [FREE Full text] [doi: 10.2196/games.8540] [Medline: 30958271]
71. New mobile game in collaboration with WHO raises COVID-19 vaccine awareness. Eurogamer. URL: http://tinyurl.com/4p5snk46 [accessed 2024-02-09]
72. The vaccination game. mrcwimm. URL: https://mrcwimm.itch.io/the-vaccination-game [accessed 2024-02-09]
73. Help take down COVID-zilla!: protect your friends and family. Alberta Health Services. URL: http://tinyurl.com/yp82u75j [accessed 2024-02-09]
74. Just the vax!. Vaccine Education Center. URL: https://media.chop.edu/data/files/vaccine-trivia-game/index.html [accessed 2024-02-09]
75. COVID invaders. Get One Desk. URL: https://www.getonedesk.com/covid-invaders [accessed 2024-02-09]
76. Vax pack hero. Vax Pack Hero. URL: https://vaxpackhero.com/ [accessed 2024-02-09]
77. Flu's clues: virtual flu game. Families Fighting Flu. URL: https://www.familiesfightingflu.org/virtual-flu-game/ [accessed 2024-02-09]
78. Virus fighter. Virus Fighter. URL: https://www.virusfighter.org/ [accessed 2024-02-09]
79. Immunization411: for preteens and teens' online training. Missouri Department of Health and Senior Services. URL: http://tinyurl.com/553ztd28 [accessed 2024-02-09]
80. COVID chronicles. COVID Chronicles. URL: http://tinyurl.com/yc73frrh [accessed 2024-02-09]
81. I boost. Public Health Association of BC. URL: https://iboostimmunity.com/ [accessed 2024-02-09]

82. Montagni I, Mabchour I, Tzourio C. Digital gamification to enhance vaccine knowledge and uptake: scoping review. JMIR Serious Games 2020 May 18;8(2):e16983 [FREE Full text] [doi: 10.2196/16983] [Medline: 32348271]
83. Ning Y, Jia Z, Zhu R, Ding Y, Wang Q, Han S. Effect and feasibility of gamification interventions for improving physical activity and health-related outcomes in cancer survivors: an early systematic review and meta-analysis. Support Care Cancer 2022 Dec 30;31(1):92. [doi: 10.1007/s00520-022-07550-0] [Medline: 36585478]
84. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2022 Jan 04;24(1):e26779 [FREE Full text] [doi: 10.2196/26779] [Medline: 34982715]
85. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic [accessed 2023-01-24]
86. Plechatá A, Vandeweerdt C, Atchapero M, Luong T, Holz C, Betsch C, et al. Experiencing herd immunity in virtual reality increases COVID-19 vaccination intention: evidence from a large-scale field intervention study. Comput Hum Behav 2023 Feb;139:107533 [FREE Full text] [doi: 10.1016/j.chb.2022.107533] [Medline: 36277032]
87. Tozzi AE, Gesualdo F, D'Ambrosio A, Pandolfi E, Agricola E, Lopalco P. Can digital tools be used for improving immunization programs? Front Public Health 2016;4:36 [FREE Full text] [doi: 10.3389/fpubh.2016.00036] [Medline: 27014673]
88. Ohannessian R, Yaghobian S, Verger P, Vanhems P. A systematic review of serious video games used for vaccination. Vaccine 2016 Aug 31;34(38):4478-4483. [doi: 10.1016/j.vaccine.2016.07.048] [Medline: 27498211]
89. Real FJ, Ollberding NJ, Meisman AR, DeBlasio DJ, Pero MB, Davis D, et al. Impact of a virtual reality curriculum on human papillomavirus vaccination: a pilot trial. Am J Prev Med 2022 Nov;63(5):865-873 [FREE Full text] [doi: 10.1016/j.amepre.2022.05.003] [Medline: 35778065]
90. Richman AR, Torres E, Wu Q, Eldridge D, Lawson L. The evaluation of a digital health intervention to improve human papillomavirus vaccine recommendation practices of medical students. J Cancer Educ 2023 Aug 17;38(4):1208-1214. [doi: 10.1007/s13187-022-02250-z] [Medline: 36526919]
91. Raikhel AV, Blau K, Alberty K, Cornia P, Rodriguez RA, Steinberg KP, et al. Vax the Max, a gamification intervention for COVID-19 vaccination task engagement in the inpatient setting. Am J Med Qual 2023;38(1):47-56 [FREE Full text] [doi: 10.1097/JMQ.0000000000000094] [Medline: 36472420]
92. Brewer NT, Chapman GB, Rothman AJ, Leask J, Kempe A. Increasing vaccination: putting psychological science into action. Psychol Sci Public Interest 2017 Dec;18(3):149-207. [doi: 10.1177/1529100618760521] [Medline: 29611455]
93. Vaisson G, Provencher T, Dugas M, Trottier M, Chipenda Dansokho S, Colquhoun H, et al. User involvement in the design and development of patient decision aids and other personal health tools: a systematic review. Med Decis Making 2021 Apr;41(3):261-274. [doi: 10.1177/0272989X20984134] [Medline: 33655791]
94. Hakim H, Bettinger JA, Chambers CT, Driedger SM, Dubé E, Gavaruzzi T, et al. A web application about herd immunity using personalized avatars: development study. J Med Internet Res 2020 Oct 30;22(10):e20113 [FREE Full text] [doi: 10.2196/20113] [Medline: 33124994]
95. Epstein D. Game design and behaviour change using Vaxcards: a collectable card game to increase vaccine confidence. Monash University. 2022. URL: http://tinyurl.com/bdhxbbfj [accessed 2024-02-10]
96. Noda S, Shirotsuki K, Nakao M. The effectiveness of intervention with board games: a systematic review. Biopsychosoc Med 2019;13:22 [FREE Full text] [doi: 10.1186/s13030-019-0164-1] [Medline: 31641371]
97. Bayeck R. Examining board gameplay and learning: a multidisciplinary review of recent research. Simul Gaming 2020 Apr 16;51(4):411-431. [doi: 10.1177/1046878119901286]
98. Talan T, Doğan Y, Batdı V. Efficiency of digital and non-digital educational games: a comparative meta-analysis and a meta-thematic analysis. J Res Technol Educ 2020 Apr 20;52(4):474-514. [doi: 10.1080/15391523.2020.1743798]
99. Naderi S, Moafian F. The victory of a non-digital game over a digital one in vocabulary learning. Comput Educ Open 2023 Dec;4:100135. [doi: 10.1016/j.caeo.2023.100135]

Abbreviations

HPV: human papillomavirus
MMAT: Mixed Methods Appraisal Tool
PICO: Population, Intervention, Comparison, and Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach, T Leung; submitted 14.03.23; peer-reviewed by R Starling, M Coccia, D Epstein, R Delaney; comments to author 23.07.23; revised version received 22.09.23; accepted 02.01.24; published 29.02.24.

Please cite as:
Hakim H, Driedger SM, Gagnon D, Chevrier J, Roch G, Dubé E, Witteman HO
Digital Gamification Tools to Enhance Vaccine Uptake: Scoping Review
JMIR Serious Games 2024;12:e47257
URL: https://games.jmir.org/2024/1/e47257
doi:10.2196/47257
PMID:38421688

Review

Investigating the Use of Serious Games for Cancer Control Among Children and Adolescents: Scoping Review

Sunghak Kim[1], PhD; Paije Wilson[2], MLiS; Olufunmilola Abraham[3], PhD

[1]National Cancer Survivorship Center, National Cancer Control Institute, National Cancer Center, Goyang, Republic of Korea

[2]Ebling Library, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, United States

[3]Social and Administrative Sciences Division, School of Pharmacy, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Olufunmilola Abraham, PhD
Social and Administrative Sciences Division
School of Pharmacy
University of Wisconsin-Madison
Room 2515 Rennebohm Hall
777 Highland Avenue
Madison, WI, 53705
United States
Phone: 1 6082634498
Fax: 1 6082625262
Email: olufunmilola.abraham@wisc.edu

Abstract

Background: Effective health care services that meet the diverse needs of children and adolescents with cancer are required to alleviate their physical, psychological, and social challenges and improve their quality of life. Previous studies showed that serious games help promote people's health. However, the potential for serious games to be used for successful cancer control for children and adolescents has received less attention.

Objective: This scoping review aimed to map the use of serious games in cancer prevention and cancer care for children and adolescents, and provide future directions for serious games' development and implementation within the context of cancer control for children and adolescents.

Methods: This study followed a combination of the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) and the JBI (Joanna Briggs Institute) framework for the conduct of scoping reviews. PubMed, CINAHL Plus Full Text, Scopus, Web of Science Core Collection, and American Psychological Association (APA) PsycINFO databases were used for the search.

Results: From the initial 2750 search results, 63 papers were included in the review, with 28 quantitative, 14 qualitative, and 21 mixed method studies. Most of the studies were cancer care serious game papers (55/63, 87%) and a small number of studies were cancer prevention serious game papers (8/63, 13%). The majority of the included studies were published between 2019 and 2023 (cancer prevention: 5/8, 63%; cancer care: 35/55, 64%). The majority of the studies were conducted in Europe (cancer prevention: 3/8, 38%; cancer care: 24/55, 44%) and North America (cancer prevention: 4/8, 50%; cancer care: 17/55, 31%). Adolescents were the most represented age group in the studies' participants (cancer prevention: 8/8, 100%; cancer care: 46/55, 84%). All (8/8, 100%) cancer prevention serious game papers included healthy people as participants, and 45 out of 55 (82%) cancer care serious game papers included patients with cancer. The majority of cancer prevention serious game papers addressed game preference as a target outcome (4/8, 50%). The majority of cancer care serious game papers addressed symptom management as a target outcome (28/55, 51%). Of the cancer care studies examining serious games for symptom management, the majority of the studies were conducted to treat psychological (13/55, 24%) and physical symptoms (10/55, 18%).

Conclusions: This review shows both the growth of interest in the use of serious games for cancer control among children and adolescents and the potential for bias in the relevant literature. The diverse characteristics of the included papers suggest that serious games can be used in various ways for cancer control among children and adolescents while highlighting the need to develop and implement serious games in underrepresented areas.

(JMIR Serious Games 2024;12:e58724) doi:10.2196/58724

KEYWORDS

serious games; cancer control; children; adolescents; scoping review; game; games; gaming; cancer; oncology; pediatric; pediatrics; paediatric; paediatrics; child; children; youth; adolescent; adolescents; teen; teens; teenager; teenagers; synthesis; review methods; review methodology; search; searches; searching; scoping

Introduction

Cancer Control for Children and Adolescents

The efforts of cancer control to reduce the cancer burden, including interventions in cancer prevention and care, have reduced the prevalence of cancer and ameliorated its impacts on individuals and communities [1,2]. Additionally, with advances in medical technology, the survival rate of children and adolescents who contract cancer has been increasing [3]. However, cancer remains a life-threatening disease for children and adolescents and requires intensive treatment over a long period of time [4]. Practicing cancer prevention is the most efficient approach to avoid the significant physical and psychological burdens experienced during the diagnosis and treatment of cancer [5]. Adolescence, especially, is a critical stage to develop one's cognitive ability and acquire new behavioral factors. Therefore, learning about cancer risks and cancer prevention methods during adolescence may significantly impact one's future health status [6,7]. Nevertheless, cancer prevention knowledge and educational opportunities for children and adolescents are limited [6,8]. Moreover, young patients with cancer easily experience fatigue, pain, sleep disorders, and anxiety and have difficulties in establishing their values and interpersonal relationships while undergoing prolonged and intense treatment [9-11]. Systematic and continuous care that meets patients' needs is crucial to helping them overcome such challenges and improve their quality of life [12-14]. Nonetheless, due to the smaller number of young patients with cancer compared to adult patients with cancer, their needs may not be prioritized in medical policy formulation or service provision [15-17]. More active industrial and research activities supporting the development and implementation of effective cancer control methods for children and adolescents should be undertaken.

The Use of Serious Games

In the significantly growing field of digital health care, based on the rapid development of information and communication technology and computer technology, the potential for serious games to be used as a successful means of cancer control is being recognized [18,19]. Serious games are digital or computerized games used primarily for educational and training purposes rather than entertainment and amusement [20]. With the widespread availability of electronic devices such as computers, gaming consoles, and mobile devices, many users can easily access and enjoy serious games [20]. Additionally, they can experience more interactive, immersive, and engaging game-based learning through various serious game content [21]. By playing serious games, users may not only obtain enjoyable and immersive experiences but also enhance motivation, engagement, and learning outcomes. Users can also develop their skills in critical thinking, decision-making, problem-solving, social interaction, time management, and so on by actively exploring the serious game content in a safe environment without physical constraints [22-25]. Adaptive and personalized functions and immediate feedback offered by the game system also promote users' continuous learning cycles while retaining user engagement [26,27]. Previous literature has shown the effectiveness of serious games designed with diverse objectives, such as improving retention of knowledge [28], pain relief [29], and medication adherence [20] in the context of various diseases.

While several positive outcomes have been associated with the use of serious games, there are a few side effects associated with the use of serious games as a health intervention tool. Previous studies have reported the possibility that the complex features of serious games may increase users' mental workload, which may negatively impact learning [30]. Additionally, some studies have argued that the addictive nature of video games should not be overlooked in the use of serious games [31] (interestingly, though, at least one study has argued the opposite effect, stating that serious games can be used as a solution to game addiction issues [32]). These potential negative effects can be prevented through careful consideration of the user groups, purposes, and appropriate uses of serious games during their development process [33].

Serious Games in the Context of Cancer Control in Children and Adolescents

Past research has also explored the relationship between serious games and young people in the context of cancer [34,35]. Adolescents, especially, tend to have excellent adaptability to new technologies and possess substantial knowledge and experience with games as compared to users of other age groups [36]. Given the research on adolescents' engagement with video games [37] and studies conducted in game-based learning [38] and narrative persuasion [39-41], one may anticipate the positive influence of serious games on adolescents' learning and persuasion outcomes. When it comes to cancer prevention, adolescents can learn about complex cancer concepts and relevant prevention and treatment methods by interacting with engaging characters and objects and actively performing game quests embedded in serious games [34]. Concerning cancer care, serious games can help distract teenage patients with cancer from the pain and anxiety associated with treatment, facilitating successful coping with the challenges of cancer [35]. Serious games can also provide psychological and social support or assist in promoting rehabilitation and physical activity [42]. However, more research is needed relating to the use of serious games in establishing successful strategies for pediatric cancer control. Cancer control for children and adolescents differs from that for adults, from the causes of cancer to the objectives and methods of cancer treatment [43,44]. For example, because children and adolescents are still cared for by caregivers (ie, legal guardians), not only young patients with cancer but also their caregivers should be included in the scope of cancer control

[45]. This is just one of many characteristics indicating the need for different approaches and considerations when providing cancer control services for children and adolescents as compared to adults. Despite these unique needs, there is a relative lack of academic and industrial projects specifically addressing serious games within the context of cancer control in children and adolescents. Systematic and comprehensive consideration of existing studies can inform and direct future research in the use of serious games for cancer control in children and adolescents.

Aim of This Study

This paper assesses the extent to which serious games have been used for cancer control (ie, cancer prevention and care) in children and adolescents. Specifically, this scoping review aimed to understand trends in serious games research in cancer prevention and cancer care for children and adolescents within published, original research papers, and investigate future directions for the application of serious games in successful cancer control for children and adolescents. Due to the importance of introducing cancer prevention and care in children and adolescents, and as the needs in the context of cancer prevention will differ from those of cancer care, this review also sought to compare serious games research focusing on cancer prevention with those focusing on cancer care to identify any key differences in research trends for these subjects.

This study will provide valuable insight and inform successful health intervention strategies relating to the use of serious games within the context of cancer control in children and adolescents. Considering the salience of cancer prevention education in adolescence, we also anticipate this study will encourage future research focusing on the use of serious games within this context.

Methods

Study Design

As the goal of the review was to explore and summarize the literature on our topic, which aligns with one of the primary purposes of a scoping review [46], a scoping review methodology was chosen for this study. Within the context of the population, concept, and context framework, our population was children and adolescents, our concept was serious games, and our context was cancer control. This study followed the JBI (Joanna Briggs Institute) framework for the conduct of scoping reviews, which specifically involved the following steps: "(1) defining and aligning the objectives and questions; (2) developing and aligning the inclusion criteria with the objectives and questions; (3) describing the planned approach to evidence searching selection, data extraction, and presentation of the evidence; (4) searching for the evidence; (5) selecting the evidence; (6) extracting the evidence; (7) analysis of the evidence; (8) presentation of the results; and (9) summarizing the evidence per the purpose of the review, making conclusions and noting any implications of the findings" [46]. This study was reported using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [47]. The PRISMA-ScR checklist can be found in Multimedia Appendix 1. The protocol for this review was registered in the OSF (Open Science Framework) [48].

Search Strategy

The literature search was developed by a health sciences librarian (PW) and included a combination of controlled vocabulary and keywords relating to serious games and cancer. No date, language, age, or geographical filters were applied to the search. The search was translated by PW for use in PubMed, CINAHL Plus Full Text (via EBSCOhost), Scopus, Web of Science Core Collection, and American Psychological Association (APA) PsycINFO (via EBSCOhost) databases. PW executed the search in each database on June 2, 2022, and reran the search using Bramer and Bain's [49] method on December 15, 2023, to retrieve any new results since the date of the first search. The results of the search were imported into EndNote 20 (Clarivate) for the first search run and EndNote 21 (Clarivate) for the search rerun. The results were deduplicated using Bramer et al's [50] method. The deduplicated results were then exported as a Microsoft Excel sheet (version 2402), which was used for screening. The full search strategy used for each database can be found in Multimedia Appendix 2.

Study Selection

To be included in the scoping review, studies needed to meet the inclusion criteria outlined in Textbox 1.

In total, 2 authors (SK and PW) independently screened the titles and abstracts of the records based on the above eligibility criteria (ie, the authors reviewed each of the records in duplicate). Any conflicts were discussed and resolved via consensus. After the completion of the title and abstract screening, the process was repeated for the full-text screening, with the same 2 authors reviewing all records in duplicate based on the previously mentioned eligibility criteria, and all conflicts being discussed and resolved via consensus. In cases where conflicts were challenging to resolve, a third reviewer served as the tiebreaker (OA). All screening (including both title and abstract and full-text screening) was performed in a Microsoft Excel sheet (version 2402).

Textbox 1. Eligibility criteria.

Inclusion criteria

- Language: the paper was written in English.
- Publication type: the paper was a full, original research paper (ie, primary study).
- Age:
  - The study recruited participants who were aged 19 years or younger. The maximum age was based on the World Health Organization (WHO)'s definition of adolescents [51]. We also included studies where data were indirectly collected for this age group, such as when parents were interviewed about their children.
  - If children or adolescents and adults were examined in the same study, the paper separately or predominantly reported on the findings of children or adolescent participants.
- Serious games used for cancer prevention or care:
  - The study examined serious games being used for cancer prevention or care. Within the context of this study, we defined "serious games" as digital or computerized games that were used for education, behavior modification, or therapy. As a note, this definition was inclusive of digital or computerized entertainment games that were used for therapeutic purposes (eg, examining whether playing a commercial video game, such as Frogger, distracted patients from cancer-related symptoms [52]). For cancer prevention, this included studies that used serious games to educate or modify behaviors for cancer prevention. For cancer care, this included studies that used serious games to care for patients or survivors of cancer with cancer-related symptoms, helping them to overcome cancer-related challenges or educating them about their cancer diagnoses or treatments.
  - For studies that examined using serious games in combination with other interventions for cancer prevention or care, the paper separately or predominantly reported on the impact of the serious games on cancer prevention or care.
  - For studies that examined using serious games in the context of cancer and other diseases, the paper separately or predominantly reported on the impact of the serious games on cancer prevention or care.

Exclusion criteria

- Language: the paper was not written in English.
- Publication type: the paper was a literature review, editorial, commentary, essay, white paper, or a type of gray literature.
- Age:
  - The study only recruited participants that were aged 20 years or older, and no data (direct or indirect) were collected on children or adolescents.
  - If children or adolescents and adults were examined in the same study, the study predominantly consisted of adults and reported their findings cumulatively (ie, they did not separately report the findings of children or adolescent participants).
- Serious games used for cancer prevention or care:
  - The study only examined nondigital games, such as board games.
  - The study did not examine serious games being used for cancer prevention or care.
  - For studies that examined using serious games in combination with other interventions for cancer prevention or care, the paper reported their findings cumulatively (ie, they did not separately report on their findings for the serious games).
  - For studies that examined other diseases as well as cancer prevention or care, the study was not predominantly focused on cancer prevention or care and reported their findings cumulatively (ie, they did not separately report their findings for cancer prevention or care).

Data Extraction

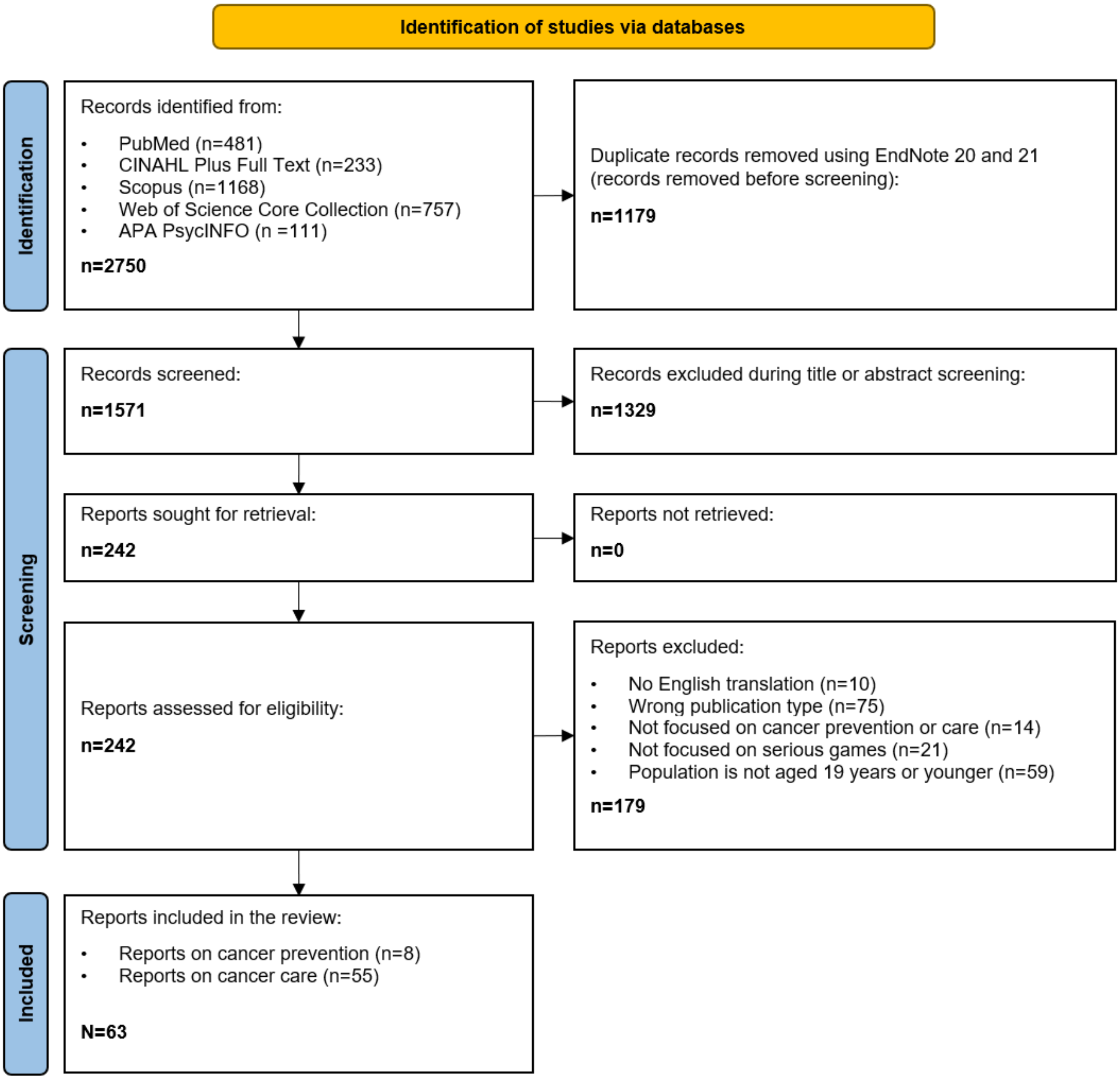
Following a full-text review, 2 authors (SK and PW) created and piloted a standardized extraction chart in Microsoft Excel (version 2402). The pilot entailed the 2 authors independently charting data from the first 10 included reports into the extraction chart, and meeting to identify areas that necessitated clarification or further standardization. After the pilot, information from all included studies was independently charted by the same authors into the extraction chart. The chart included each study's authors, title, publication year, location, participant age (ie, whether participants were children, adolescents, or adults), participant type (ie, whether the participants were patients with cancer, survivors of cancer, health professionals, caregivers, or healthy people; or whether the participant type was unspecified), serious game objective (ie, whether the serious game was being used for cancer prevention or cancer care), serious game name, target outcome (ie, what outcomes were primarily examined for the study), and, for studies that had symptom management as a target outcome, target symptom (ie, what symptoms were primarily examined for the study). Any discrepancies in the data charting were discussed by the 2 authors and resolved via consensus.

Results

Overview

The search retrieved a total of 2750 records. Of those records, 1179 were identified as duplicates and removed. Title and abstract screening was performed on the remaining 1571 records, of which 1329 were excluded. The remaining 242 records underwent full-text screening, with 179 of these records being excluded (resulting in a total of 63 included records for the review; see Figure 1 [53]).

Figure 1. PRISMA flow diagram. Adapted from Page et al. APA: American Psychological Association; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Basic Information

Of the 63 papers included in this study, cancer care was the serious game objective with the most research (55/63, 87%) [35,42,52,54-105], with cancer prevention only having 8 papers (13%) [7,34,106-111]. The majority of the included studies were published between 2019 and 2023. To be specific, 5 out of 8 (63%) cancer prevention serious game papers [7,106,109-111] and 35 out of 55 (64%) cancer care serious game papers [35,42,54,58,59,61-63,65,66,71,72,74,75,78-82,85,91-105] were published between 2019 and 2023. The studies were conducted in Asia, Europe, North America, and Oceania, with the majority of the studies being conducted in Europe and North America. Of the 8 cancer prevention serious game papers, 3 (38%) [107,109,110] and 4 (50%) [7,34,106,108] studies were conducted in Europe and North America, respectively. Of the 55 cancer care serious game papers, 24 (44%) [35,58,61-63,65-67,70,74,75,81-83,85-87,92,93,97-100,103] and 17 (31%) [52,54-57,59,60,63,68,69,73,78,79,84,89,90,102] studies were conducted in Europe and North America, respectively. Of the 8 cancer prevention serious game studies, 3 (38%) were quantitative [34,110,111], 3 (38%) were qualitative [7,107,108], and 2 (25%) were mixed methods studies [106,109]. Of the 55 cancer care serious game studies, 25 (45%) were quantitative

[35,42,52,54-56,58,64,69,71-74,76,77,80,84,87-89,91,96,97,100,105], 11 (20%) were qualitative [57,59,62,63,67,81,82,93,99,102,103], and 19 (35%) were mixed methods studies [60,61,65, 66,68,70,75,78,79,83,85,86,90,92,94,95,98,101,104]. For the distribution of these studies by publication year and location see Table 1.

Table 1. Basic information of included papers.

Characteristics	Serious game objective	
	Cancer prevention (n=8)	Cancer care (n=55)
Publication year, n (%)[a]		
1985-1998 [52,73]	0 (0)	2 (4)
1999-2003 [89]	0 (0)	1 (2)
2004-2008 [55,56,68,69]	0 (0)	4 (7)
2009-2013 [57,76,77,83,90,107]	1 (13)	5 (9)
2014-2018 [34,60,64,67,70,84,86-88,108]	2 (25)	8 (15)
2019-2023 [7,35,42,54,58,59,61-63,65,66,71,72,74,75,78-82,85,91-106,109-111]	5 (63)	35 (64)
Location, n (%)[a]		
Asia [42,64,71,72,76,77,80,88,91,94-96,101,105,111]	1 (13)	14 (26)
Europe [35,58,61-63,65-67,70,74,75,81-83,85-87,92,93,97-100,103,107,109,110]	3 (38)	24 (44)
North America [7,34,52,54-57,59,60,63,68,69,73,78,79,84,89,90,102,106,108]	4 (50)	17 (31)
Oceania [55,56,69,104]	0 (0)	4 (7)
Study type, n (%)[a]		
Quantitative [34,35,42,52,54-56,58,64,69,71-74,76,77,80,84,87-89,91,96,97,100,105,110,111]	3 (38)	25 (45)
Qualitative [7,57,59,62,63,67,81,82,93,99,102,103,107,108]	3 (38)	11 (20)
Mixed methods [60,61,65,66,68,70,75,78,79,83,85,86,90,92,94,95,98,101,104,106,109]	2 (25)	19 (35)

[a]Note that the frequency numbers may not add up to the total number of studies (N=63), as some studies included more than one category. Percentages may also not add up to 100% due to rounding.

Participant Characteristics

Children and adolescents were the main participants represented in the papers included in this review, which aligned with this study's purpose. The adolescents' age group had the greatest representation in the included studies—all (8/8, 100%) cancer prevention serious game papers [7,34,106-111] and 46 out of 55 (84%) cancer care serious game papers [35,42,52,54-62, 64-66,68-80,83,84,86-93,95-98,100,101,104,105] involved adolescents as participants. Interestingly, many studies included adult participants as well as adolescents or children. As mentioned in our inclusion criteria, studies examining adult populations were included if these also examined and isolated data for our population of interest (ie, children or adolescents) or if adults were interviewed about children or adolescents. Of our included studies, 25 of the 55 (45%) cancer care papers [52,55,56,60-62,66-69,79-82,84,85,92-95,97,99,101,102,104] and 2 of the 8 (25%) cancer prevention papers [34,109] included adults as well as adolescents or children. Further, 2 studies only included adults (being studies where parents were interviewed about their children, but where the children did not participate in the study themselves), with both of these studies being cancer care papers [63,103].

All (8/8, 100%) cancer prevention serious game papers included healthy people as participants [7,34,106-111]. Regarding cancer care serious game papers, 45 out of 55 (82%) studies included patients with cancer [35,42,52,55-57,59-62,64-66,68-79,81-84, 88-90,92-102,104,105]. Surprisingly, several papers also involved health professionals and caregivers as participants, wherein these groups were interviewed to solicit their impressions and opinions of how the serious games affected their young patients or children, respectively [60-63,66,79, 81,82,85,92-95,97,99,101-103]. In these cases, where health professionals' or caregivers' opinions or quotes were associated with our population of interest (ie, children or adolescents) and these opinions or quotes were included in the results section of the studies, our review study included those papers and counted the health professionals and caregivers as the study's participants.

Table 2 presents more details about the distribution of included studies by participant age and participant type.

Table 2. Participant characteristics of included papers.

Characteristics	Serious game objective	
	Cancer prevention (n=8)	Cancer care (n=55)
Participant age (years), n (%)[a]		
Children (0-9) [35,42,52,57-62,64-67,70,72,74,76-79,81-87,91-99,101,102,107,108]	2 (25)	38 (69)
Adolescents (10-19) [7,34,35,42,52,54-62,64-66,68-80,83,84,86-93,95-98,100,101,104-111]	8 (100)	46 (84)
Adults (≥20) [34,52,55,56,60-63,66-69,79-82,84,85,92-95,97,99,101-104,109]	2 (25)	27 (49)
Participant type, n (%)[a]		
Patients with cancer [35,42,52,55-57,59-62,64-66,68-79,81-84,88-90,92-102,104,105]	0 (0)	45 (82)
Survivors of cancer [54,57,58,80,86,87,91]	0 (0)	7 (13)
Health professionals [60,63,66,81,82,85,93,94,97]	0 (0)	9 (16)
Caregivers [60-63,66,79,81,82,92,95,97,99,101-103]	0 (0)	15 (27)
Healthy people [7,34,66,70,75,85,91,106-111]	8 (100)	5 (9)
Unspecified (not disclosed) [67]	0 (0)	1 (2)

[a]Note that the frequency numbers may not add up to the total number of studies (N=63), as some studies included more than one category. Percentages may also not add up to 100% due to rounding.

Role of Serious Games in Cancer Control for Young People

The papers included in this review examined the use of serious games for a variety of target outcomes. For cancer prevention, the target outcome most addressed by the papers was exploring participants' preferences for the content of the serious games (ie, participants' satisfaction with the serious game or likes and dislikes about features within the serious game; 4/8, 50%) [7,108,110,111], followed by educating participants about cancer (3/8, 38%) [34,106,109], and promoting healthy behaviors (1/8, 13%) [107]. For cancer care, the target outcome most addressed by the papers was using serious games for symptom management (28/55, 51%) [35,42,52,58,61,64,65,71-73,76,77,81-84,86-89,91,92,94,97,99,100,104,105], followed closely by exploring participants' preferences for the content of the serious games (24/55, 44%) [55,56,62,63,67,68,70,75,79,84-86,90,92-98,101-104]. Other target outcomes observed for cancer care included promoting healthy behaviors (10/55, 18%) [42,54,57,60,65,74,80,95,96,101], symptom reporting (7/55, 13%) [59,66,75,78,79,90,102], cancer education (6/55, 11%) [55,56,80,95,101,105], and treatment adherence (3/55, 6%) [69,95,101]. Of the cancer care studies examining serious games for symptom management, the majority of the studies were conducted to treat psychological (13/55, 24%) [61,71-73,76,77,81,83,89,92,94,97,100] and physical symptoms (10/55, 18%) [35,42,52,58,65,81,83,86,99,104]. Naturally, none of the cancer prevention serious game papers were conducted with the goal of managing cancer-related symptoms.

Table 3 presents more details about the distribution of included studies by target outcome and target symptom.

Table 3. The role of serious games in included papers.

Characteristics	Serious game objective	
	Cancer prevention (n=8)	Cancer care (n=55)
Target outcome, n (%)[a]		
Serious game preference [7,55,56,62,63,67,68,70,75,79,84-86,90,92-98,101-104,108,110,111]	4 (50)	24 (44)
Treatment adherence [69,95,101]	0 (0)	3 (6)
Cancer education [34,55,56,80,95,101,105,106,109]	3 (38)	6 (11)
Healthy behavior promotion [42,54,57,60,65,74,80,95,96,101,107]	1 (13)	10 (18)
Symptom management [35,42,52,58,61,64,65,71-73,76,77,81-84,86-89,91,92,94,97,99,100,104,105]	0 (0)	28 (51)
Symptom reporting [59,66,75,78,79,90,102]	0 (0)	7 (13)
Target symptom, n (%)[a]		
Psychological symptoms[b] [61,71-73,76,77,81,83,89,92,94,97,100]	0 (0)	13 (24)
Physical symptoms[c] [35,42,52,58,65,81,83,86,99,104]	0 (0)	10 (18)
Cognitive symptoms[d] [42,58,84,87,88,104]	0 (0)	6 (11)
General side effects[e] [73,105]	0 (0)	2 (4)
Other[f] [64,82,87,91,104,105]	0 (0)	6 (11)

[a]Note that the frequency numbers may not add up to the total number of studies (N=63), as some studies included more than one category. Percentages may also not add up to 100% due to rounding.

[b]"Psychological symptoms" includes papers that addressed "anxiety," "depression," "distress," "emotional state," "psychosocial symptoms," or "psychological symptoms" in this study.

[c]"Physical symptoms" includes papers that addressed "endurance," "fatigue," "motor performance," "nausea," "pain," or "physical activity" in this study.

[d]"Cognitive symptoms" includes papers that addressed "cognitive behavioral effects," "cognitive function," "functional capacity," or "reading deficits" in this study.

[e]"General side effects" includes papers that addressed "side effects of chemotherapy" or "symptoms" (nonspecific) in this study.

[f]"Other" includes papers that addressed "activities of daily living," "daily performance," "quality of life," or "sleep" in this study.

Discussion

Overview

This scoping review aimed to examine the extent to which serious games have been used within the context of cancer prevention and care for children and adolescents within published, original research papers, and investigate future directions for the application of serious games in successful cancer control for children and adolescents. In this review, we identified papers that explored the potential of serious games being used for successful cancer control for children and adolescents. When observing the distribution of these papers in a comprehensive map, we also identified several gaps that may introduce bias into the existing literature.

Need for Cancer Prevention Serious Games Research

The results of this study showed that there were considerably fewer cancer prevention serious game papers than cancer care serious game papers. This difference may indicate that cancer prevention serious games are perceived as being less important than cancer care serious games. The necessity of pediatric cancer prevention might also be overlooked because of the relatively insufficient medical infrastructure and interest in pediatric cancer prevention as compared to adult cancer prevention, being in part due to the smaller number of pediatric patients with cancer compared to adult patients with cancer [15-17]. Another explanation could be the complex etiology of pediatric cancer. For adult cancers, there are several studies establishing the associations between cancer occurrence and lifestyle or environmental risk factors [112,113], which may help people take preventive measures. However, it is known that lifestyle or environmental risk factors are unlikely to play a significant role in the occurrence of pediatric cancers [114]. The lack of evidence about lifestyle or environmental risk factors may hinder identifying them and reduce preventive efforts for pediatric cancers. Consequently, the awareness of pediatric cancer prevention may be low. Nevertheless, considering that cancer prevention interventions are the most effective way to reduce cancer-related risks, it is essential to develop and provide cancer prevention services for pediatric cancers. Future research should be conducted to develop serious games that educate young people about cancer so that they can have knowledge about cancer and preemptively engage in healthy behaviors.

Disparities in Cancer Control Serious Games Research Relating to the Publication Year and Location

This study's findings also confirmed that the publication year and location are concentrated in specific years and countries. Both cancer prevention serious game papers and cancer care serious game papers were largely published after the year 2019. This distribution may imply a trend that researchers are recently paying more attention to the application of serious games in digital health care [115]. The impact of serious games will be amplified by systematically delving into the conceptual and theoretical underpinnings of serious games and uncovering strategies to develop and use serious games for constructive and socially beneficial purposes. Future research needs to investigate the positive influence of serious games and accelerate their implementation in cancer prevention and care. Doing so will facilitate the successful use of serious games in pediatric cancer control and spearhead advancements in digital health care. Additionally, research on serious games in both cancer prevention and cancer care has predominantly been conducted in Europe and North America. As the positive effects of serious games in the medical field have been indicated through existing research, future research needs to explore avenues to encourage serious game studies in the medical field for researchers in other geographic regions. Such efforts would be beneficial to mitigate disparities in digital technology and reduce inequalities in access to health care services.

Embracing Various Age Groups in Cancer Control Serious Games Research

Moreover, our study revealed distinct patterns in the distribution of papers based on participant characteristics. Regarding participant age, it was observed that all papers on cancer prevention serious games focused on adolescents. Similarly, the predominant age group featured in papers concerning cancer care serious games was also adolescents. These findings are not unexpected, considering adolescents' familiarity with gaming [36] and the research objective of this review study. However, the observed distribution indirectly suggests a divergence between the child and adolescent groups, while indicating a relative neglect in research focusing on serious games for children. In addition to exploring cancer control serious games tailored for children, future studies could also compare the characteristics and uses of serious games designed for children with those designed for adolescents, and, in doing so, identify salient differences between the 2 groups. An additional finding of note was that, due to the intricate dynamics of pediatric cancer that affect not only young patients but also adults closely associated with them [45], certain studies incorporated adult participants as well as children or adolescents. More research should encompass both child and adult users when developing a serious game for effective pediatric cancer control or when evaluating user experiences, as doing so could furnish more comprehensive and nuanced findings.

Embracing Various Participant Types in Cancer Care Serious Games Research

Regarding participant type, it was observed that all papers on cancer prevention serious games exclusively targeted healthy people. Conversely, the primary participant type featured in papers focusing on cancer care serious games was patients with cancer. Notably, cancer care serious game papers included a diverse array of participant types, with some studies even including multiple participant types. This suggests the potential for serious games to cater to diverse participant types, even those characterized by different interests and attributes. Future research should strive to understand the different needs of multiple participant types to develop effective, and far-reaching cancer care serious games. A comprehensive approach to identify different user needs within the context of serious games should also be encouraged.

The diverse array of participant types also implies the potential for participants to play a variety of roles in developing serious games. Pediatric patients with cancer and survivors can share their opinions and user experiences when using serious games [55,86]. Caregivers can also share their opinions and user experiences when they or their dependents use serious games while also encouraging their dependents to play serious games [60,61]. Health professionals can provide expertise and a nuanced understanding of the needs of their patients and can therefore provide invaluable feedback for the design and content of serious games [60,82]. All these roles can facilitate the evaluation and development of serious games.

Finally, our review found that a small number of papers included survivors of cancer, caregivers, or health professionals as participants when examining the use of serious games for cancer control in children or adolescents. In particular, survivors of cancer and caregivers have difficulties in resolving their unmet needs or draw less emphasis or attention to cancer care services [116,117]. Crafting appropriate serious games for survivors of cancer and caregivers based on an advanced understanding of them may improve both the quantity and quality of cancer care services for young people. As health care providers, health professionals can be overlooked in their potential to play a role in serious game research for cancer control in children or adolescents; however, as stated previously, health professionals can provide invaluable feedback for the development and evaluation of serious games.

Enriching Cancer Control Serious Games Research by Focusing on Underrepresented Target Outcomes

The roles of the serious games differed between cancer prevention and cancer care. Serious game preference was one of the main target outcomes for both cancer prevention serious game papers and cancer care serious game papers. Cancer education emerged as another key target outcome for cancer prevention serious game papers, whereas symptom management took precedence in cancer care serious game papers. Similar to the case with participant types, cancer care serious game papers addressed a diverse range of target outcomes, with some studies addressing multiple target outcomes within a single investigation. This suggests the feasibility of developing and deploying serious games for young people's cancer care with multifaceted purposes. There were, however, some target outcomes that were underrepresented, such as healthy behavior promotion for cancer prevention studies and management of cognitive symptoms for cancer care studies. Future research should delve into effective strategies for serious game development and implementation in these underrepresented target outcomes, thus bridging existing gaps and fostering more comprehensive discussions concerning the use of serious games in cancer prevention and care.

Limitations

This study demonstrated the research state on the use of serious games for cancer control among children and adolescents and suggested the future directions of serious game development and research; however, this study did have limitations. First, only papers written in English were included in the review, so valuable data from relevant papers published in other languages may have been excluded. Second, multiple studies included in this study's data analysis recruited not only children and adolescents but also adults as participants; therefore, careful interpretation of serious games' influence on children and adolescents is necessary. Third, we opted to limit our results to a reputable and manageable selection, with the goal of the review being to provide an overview of original, primary studies on our topic. Due to this, we refrained from including paper types such as reviews, conference proceedings, etc, in our review. Fourth, we did not evaluate the statistical outcomes of the serious games within each study, as this was not within the scope of our review (with the scope of our review being to map the use of serious games within the context of cancer control). Future research, such as a systematic review, could provide valuable insight into the statistical efficacy of serious games within focused areas of cancer control. Future studies could also assess topics such as features and frameworks of serious games, providing more insight into their development.

Conclusions

This review study shows that there has been an increased interest in the use of serious games for cancer control among children and adolescents. At the same time, this study reveals that the number of papers has been skewed in terms of the purpose and context of serious games for cancer control in children and adolescents, with cancer prevention serious games in young people having received considerably less attention than those of cancer care. Additionally, the frequencies of study participants' characteristics and target outcomes differ depending on the serious game objectives. Regarding cancer prevention serious game papers, these studies primarily examined serious game preference. Regarding cancer care serious game papers, these studies primarily examined serious game preference and how serious games help alleviate cancer-related symptoms. These differences suggest that serious games can be used in multiple ways within cancer control in children and adolescents while highlighting the need to develop and implement serious games in underrepresented areas. Further studies are needed to comprehensively examine which features of serious games enhance young people's cancer control and what evidence-based and theory-driven methods are available to develop effective serious games for this purpose. Integrating these future findings into this review study's outcomes may help advance successful serious game development and implementation for young people's cancer control.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5A8081124).

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

SK contributed to the conceptualization, funding acquisition, formal analysis, investigation, methodology, project administration, writing of the original draft, and reviewing and editing of this paper. PW contributed to the data curation, formal analysis, investigation, methodology, project administration, writing of the original draft, and reviewing and editing of this paper. OA contributed to the conceptualization, resources, supervision, project administration, writing of the original draft, and reviewing and editing of this paper. All authors approved the final version of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist of this scoping review.

[PDF File (Adobe PDF File), 117 KB - games_v12i1e58724_app1.pdf]

Multimedia Appendix 2

The searching strategy of this scoping review.

[PDF File (Adobe PDF File), 546 KB - games_v12i1e58724_app2.pdf]

References

1. What is cancer control. Union for International Cancer Control. URL: https://www.uicc.org/who-we-are/about-uicc/what-cancer-control [accessed 2024-03-13]
2. Cancer control continuum. National Cancer Institute at the National Institutes of Health. URL: https://cancercontrol.cancer.gov/about-dccps/about-cc/cancer-control-continuum [accessed 2024-03-13]
3. Cancer in children and adolescents. National Cancer Institute at the National Institutes of Health. URL: https://www.cancer.gov/types/childhood-cancers/child-adolescent-cancers-fact-sheet [accessed 2023-07-10]
4. Childhood cancer. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/cancer-in-children [accessed 2023-07-10]
5. Preventing cancer. World Health Organization. URL: https://www.who.int/activities/preventing-cancer [accessed 2024-03-20]
6. Abraham O, Rosenberger CA, LeMay SM, Bittner SJ. Adolescents' perceptions about cancer and preferences for cancer education. Cancer Control 2021;28:10732748211036057 [FREE Full text] [doi: 10.1177/10732748211036057] [Medline: 34403277]
7. Abraham O, Szela L, Khan M, Geddam A. Exploring middle school students' perspectives on using serious games for cancer prevention education: focus group study. JMIR Serious Games 2022;10(1):e31172 [FREE Full text] [doi: 10.2196/31172] [Medline: 34643533]
8. Kye SY, Hwang SY, Oh KH, Jun JK. Effects of a cancer prevention education program on elementary school students' knowledge, attitude, self-efficacy, and intentions in South Korea. Epidemiol Health 2019;41:e2019027 [FREE Full text] [doi: 10.4178/epih.e2019027] [Medline: 31208191]
9. Examining the psychological impact of childhood cancer. American Psychological Association. 2009. URL: https://www.apa.org/topics/children/cancer-psychological-impact [accessed 2024-03-20]
10. Yi J, Kim MA, Tian T. Perceived long-term and physical health problems after cancer: adolescent and young adult survivors of childhood cancer in Korea. Eur J Oncol Nurs 2014;18(2):145-150. [doi: 10.1016/j.ejon.2013.11.006] [Medline: 24361038]
11. Toft T, Alfonsson S, Hovén E, Carlsson T. Feeling excluded and not having anyone to talk to: qualitative study of interpersonal relationships following a cancer diagnosis in a sibling. Eur J Oncol Nurs 2019;42:76-81. [doi: 10.1016/j.ejon.2019.07.010] [Medline: 31450042]
12. Datta SS, Saha T, Ojha A, Das A, Daruvala R, Reghu KS, et al. What do you need to learn in paediatric psycho-oncology? Ecancermedicalscience 2019;13:916 [FREE Full text] [doi: 10.3332/ecancer.2019.916] [Medline: 31123499]
13. Kim M, Choi K, You J. Current psychosocial care for children and adolescents with cancer and their families: perspectives of service providers in hospitals and community welfare agencies. Health Soc Welfare Review 2021;41(3):130-159. [doi: 10.15709/hswr.2021.41.3.130]
14. Zhang J, Wang C, Huang L, Zhang J. Continuous care needs in patients with cancer receiving chemotherapy during the recent omicron wave of COVID-19 in Shanghai: a qualitative study. Front Psychol 2022;13:1067238 [FREE Full text] [doi: 10.3389/fpsyg.2022.1067238] [Medline: 36687977]
15. Gupta S, Rivera-Luna R, Ribeiro RC, Howard SC. Pediatric oncology as the next global child health priority: the need for national childhood cancer strategies in low- and middle-income countries. PLoS Med 2014;11(6):e1001656 [FREE Full text] [doi: 10.1371/journal.pmed.1001656] [Medline: 24936984]
16. Ndlovu SR, Kuupiel D, Ginindza TG. Mapping evidence on the distribution of paediatric cancers in sub-Saharan Africa: a scoping review protocol. Syst Rev 2019;8(1):262 [FREE Full text] [doi: 10.1186/s13643-019-1184-4] [Medline: 31699138]
17. Denburg A, Wilson MG, Johnson SI, Kutluk T, Torode J, Gupta S. Advancing the development of national childhood cancer care strategies in Latin America. J Cancer Policy 2017;12:7-15. [doi: 10.1016/j.jcpo.2016.12.005]
18. Kondylakis H, Bucur A, Crico C, Dong F, Graf N, Hoffman S, et al. Patient empowerment for cancer patients through a novel ICT infrastructure. J Biomed Inform 2020;101:103342 [FREE Full text] [doi: 10.1016/j.jbi.2019.103342] [Medline: 31816400]
19. Ghazisaeidi M, Safdari R, Goodini A, Mirzaiee M, Farzi J. Digital games as an effective approach for cancer management: opportunities and challenges. J Educ Health Promot 2017;6:30 [FREE Full text] [doi: 10.4103/jehp.jehp_146_14] [Medline: 28584830]
20. Abraham O, LeMay S, Bittner S, Thakur T, Stafford H, Brown R. Investigating serious games that incorporate medication use for patients: systematic literature review. JMIR Serious Games 2020;8(2):e16096 [FREE Full text] [doi: 10.2196/16096] [Medline: 32347811]
21. Xu M, Luo Y, Zhang Y, Xia R, Qian H, Zou X. Game-based learning in medical education. Front Public Health 2023;11:1113682 [FREE Full text] [doi: 10.3389/fpubh.2023.1113682] [Medline: 36935696]
22. Yang YC. Building virtual cities, inspiring intelligent citizens: digital games for developing students’ problem solving and learning motivation. Comput Educ 2012;59(2):365-377. [doi: 10.1016/j.compedu.2012.01.012]
23. Altomari L, Altomari N, Iazzolino G. Gamification and soft skills assessment in the development of a serious game: design and feasibility pilot study. JMIR Serious Games 2023;11:e45436 [FREE Full text] [doi: 10.2196/45436] [Medline: 37494078]
24. Schrader C. Serious games and game-based learning. In: Handbook of Open, Distance and Digital Education. Singapore: Springer; 2022:1-14. [doi: 10.1007/978-981-19-0351-9_74-1]

25. Checa D, Bustillo A. A review of immersive virtual reality serious games to enhance learning and training. Multimed Tools Appl 2020;79(9-10):5501-5527 [FREE Full text] [doi: 10.1007/s11042-019-08348-9]
26. Sajjadi P, Ewais A, De Troyer O. Individualization in serious games: a systematic review of the literature on the aspects of the players to adapt to. Entertainment Comput 2022;41:100468. [doi: 10.1016/j.entcom.2021.100468]
27. Bontchev BP, Terzieva V, Paunova-Hubenova E. Personalization of serious games for learning. Interact technol smart educ 2021;18(1):50-68. [doi: 10.1108/ITSE-05-2020-0069]
28. Hu H, Xiao Y, Li H. The effectiveness of a serious game versus online lectures for improving medical students' coronavirus disease 2019 knowledge. Games Health J 2021;10(2):139-144. [doi: 10.1089/g4h.2020.0140] [Medline: 33434099]
29. Saragih ID, Suarilah I, Saragih IS, Lin YK, Lin CJ. Efficacy of serious games for chronic pain management in older adults: A systematic review and meta-analysis. J Clin Nurs 2024;33(3):1185-1194. [doi: 10.1111/jocn.17012] [Medline: 38291564]
30. Zhonggen Y. A meta-analysis of use of serious games in education over a decade. Int J Comput Games Technol 2019;2019(17):4797032. [doi: 10.1155/2019/4797032]
31. Rodrigo-Yanguas M, González-Tardón C, Bella-Fernández M, Blasco-Fontecilla H. Serious video games: angels or demons in patients with attention-deficit hyperactivity disorder? a quasi-systematic review. Front Psychiatry 2022;13:798480 [FREE Full text] [doi: 10.3389/fpsyt.2022.798480] [Medline: 35573357]
32. Ben C, Martino F, Charlene J, Martin R, Niklas R. Learning when serious: psychophysiological evaluation of a technology-enhanced learning game. J Educ Technol Soc 2014;17(1):3-16.
33. Laamarti F, Eid M, El Saddik A. An overview of serious games. Int J Comput Games Technol 2014;2014:358152. [doi: 10.1155/2014/358152]
34. Khalil GE, Beale IL, Chen M, Prokhorov AV. A video game promoting cancer risk perception and information seeking behavior among young-adult college students: a randomized controlled trial. JMIR Serious Games 2016;4(2):e13 [FREE Full text] [doi: 10.2196/games.5793] [Medline: 27470927]
35. Alonso Puig M, Alonso-Prieto M, Miró J, Torres-Luna R, Plaza López de Sabando D, Reinoso-Barbero F. The association between pain relief using video games and an increase in vagal tone in children with cancer: analytic observational study with a quasi-experimental pre/posttest methodology. J Med Internet Res 2020;22(3):e16013 [FREE Full text] [doi: 10.2196/16013] [Medline: 32224482]
36. Giovanelli A, Ozer EM, Dahl RE. Leveraging technology to improve health in adolescence: a developmental science perspective. J Adolesc Health 2020;67(2S):S7-S13 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.02.020] [Medline: 32718517]
37. Ream GL, Elliott LC, Dunlap E. Trends in video game play through childhood, adolescence, and emerging adulthood. Psychiatry J 2013;2013:301460 [FREE Full text] [doi: 10.1155/2013/301460] [Medline: 24236277]
38. Abdul Jabbar AI, Felicia P. Gameplay engagement and learning in game-based learning: a systematic review. Rev Educ Res 2015;85(4):740-779. [doi: 10.3102/0034654315577210]
39. Green MC, Jenkins KM. Interactive narratives: processes and outcomes in user-directed stories. J Commun 2014;64(3):479-500. [doi: 10.1111/jcom.12093]
40. Kim J, Song H, Merrill K, Jung Y, Kwon RJ. Using serious games for antismoking health campaigns: experimental study. JMIR Serious Games 2020;8(4):e18528 [FREE Full text] [doi: 10.2196/18528] [Medline: 33263549]
41. De Oro CM, Molinares DJ, Coronado AME, Carrillo R. From gamification to serious games: reinventing learning processes. In: Pedagogy-Challenges, Recent Advances, New Perspectives, and Applications. UNITED KINGDOM: IntechOpen Limited; 2021. [doi: 10.5772/intechopen.101324]
42. Masoud AE, Shaheen AAM, Algabbani MF, AlEisa E, AlKofide A. Effectiveness of exergaming in reducing cancer-related fatigue among children with acute lymphoblastic leukemia: a randomized controlled trial. Ann Med 2023 Dec;55(1):2224048 [FREE Full text] [doi: 10.1080/07853890.2023.2224048] [Medline: 37318119]
43. What are the differences between cancer in adults and children? American Cancer Society. 2019. URL: https://www.cancer.org/cancer/types/cancer-in-children/differences-adults-children.html [accessed 2024-03-20]
44. Kattner P, Strobel H, Khoshnevis N, Grunert M, Bartholomae S, Pruss M, et al. Compare and contrast: pediatric cancer versus adult malignancies. Cancer Metastasis Rev 2019;38(4):673-682. [doi: 10.1007/s10555-019-09836-y] [Medline: 31832830]
45. Lewandowska A, Rudzki G, Lewandowski T, Rudzki S. The problems and needs of patients diagnosed with cancer and their caregivers. Int J Environ Res Public Health 2020 Dec 24;18(1):87 [FREE Full text] [doi: 10.3390/ijerph18010087] [Medline: 33374440]
46. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Chapter 11: scoping reviews (2020 version). JBI Manual for Evidence Synthesis. URL: https://edisciplinas.usp.br/pluginfile.php/7315963/mod_resource/content/1/manual_capitulo_revisao_escopo_JBIMES_2021April.pdf [accessed 2024-06-07]
47. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
48. Kim S, Wilson P, Abraham O. The use of serious games for cancer control in children and adolescents: a scoping review protocol. OSF Preprint posted online on November 22, 2023. [doi: 10.17605/OSF.IO/TZQHD]

49. Bramer W, Bain P. Updating search strategies for systematic reviews using EndNote. J Med Libr Assoc 2017;105(3):285-289 [FREE Full text] [doi: 10.5195/jmla.2017.183] [Medline: 28670219]
50. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc 2016;104(3):240-243 [FREE Full text] [doi: 10.3163/1536-5050.104.3.014] [Medline: 27366130]
51. Adolescent health. World Health Organization. URL: https://www.who.int/health-topics/adolescent-health#tab=tab_1 [accessed 2024-06-06]
52. Redd WH, Jacobsen PB, Die-Trill M, Dermatis H, McEvoy M, Holland JC. Cognitive/attentional distraction in the control of conditioned nausea in pediatric cancer patients receiving chemotherapy. J Consult Clin Psychol 1987;55(3):391-395. [doi: 10.1037//0022-006x.55.3.391] [Medline: 3597954]
53. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372(11-12):n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
54. Fuemmeler BF, Holzwarth E, Sheng Y, Do EK, Miller CA, Blatt J, et al. Mila Blooms: a mobile phone application and behavioral intervention for promoting physical activity and a healthy diet among adolescent survivors of childhood cancer. Games Health J 2020;9(4):279-289 [FREE Full text] [doi: 10.1089/g4h.2019.0060] [Medline: 32391734]
55. Beale IL, Kato PM, Marin-Bowling VM, Guthrie N, Cole SW. Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. J Adolesc Health 2007;41(3):263-270. [doi: 10.1016/j.jadohealth.2007.04.006] [Medline: 17707296]
56. Beale IL, Marin-Bowling VM, Guthrie N, Kato PM. Young cancer patients' perceptions of a video game used to promote self care. Int Electron J Health Educ 2006;9:202-212 [FREE Full text]
57. Beltran A, Li R, Ater J, Baranowski J, Buday R, Thompson D, et al. Adapting a videogame to the needs of pediatric cancer patients and survivors. Games Health J 2013 Aug;2(4):213-221. [doi: 10.1089/g4h.2013.0018] [Medline: 26192225]
58. Benzing V, Spitzhüttl J, Siegwart V, Schmid J, Grotzer M, Heinks T, et al. Effects of cognitive training and exergaming in pediatric cancer survivors—a randomized clinical trial. Med Sci Sports Exerc 2020;52(11):2293-2302 [FREE Full text] [doi: 10.1249/MSS.0000000000002386] [Medline: 33064404]
59. Bernier Carney KM, Jung SH, Iacob E, Lewis M, Linder LA. Communication of pain by school-age children with cancer using a game-based symptom assessment app: a secondary analysis. Eur J Oncol Nurs 2021;52(11-12):101949 [FREE Full text] [doi: 10.1016/j.ejon.2021.101949] [Medline: 33813185]
60. Bruggers CS, Baranowski S, Beseris M, Leonard R, Long D, Schulte E, et al. A prototype exercise-empowerment mobile video game for children with cancer, and its usability assessment: developing digital empowerment interventions for pediatric diseases. Front Pediatr 2018;6:69 [FREE Full text] [doi: 10.3389/fped.2018.00069] [Medline: 29686977]
61. Carrion-Plaza A, Jaen J, Montoya-Castilla I. HabitApp: new play technologies in pediatric cancer to improve the psychosocial state of patients and caregivers. Front Psychol 2020;11(11-12):157 [FREE Full text] [doi: 10.3389/fpsyg.2020.00157] [Medline: 32116953]
62. Cederved C, Back J, Ångström-Brännström C, Ljungman G, Engvall G. Co-creation of a serious game about radiotherapy: participatory action research study with children treated for cancer. JMIR Hum Factors 2022;9(2):e34476. [doi: 10.2196/34476]
63. De la Hera T, Sanz CS. The role of spontaneous digital play during young patients' cancer treatment. Media Commun 2021;9(1):39-48. [doi: 10.17645/mac.v9i1.3216]
64. Fazelniya Z, Najafi M, Moafi A, Talakoub S. The impact of an interactive computer game on the quality of life of children undergoing chemotherapy. Iranian J Nursing Midwifery Res 2017;22(6):431-435. [Medline: 29184580]
65. Hamari L, Järvelä LS, Lähteenmäki PM, Arola M, Axelin A, Vahlberg T, et al. The effect of an active video game intervention on physical activity, motor performance, and fatigue in children with cancer: a randomized controlled trial. BMC Res Notes 2019;12(1):784. [doi: 10.1186/s13104-019-4821-z]
66. Hoffmann S, Schraut R, Kröll T, Scholz W, Belova T, Erhardt J, et al. AquaScouts: ePROs implemented as a serious game for children with cancer to support palliative care. Front Digit Health 2021;3:730948 [FREE Full text] [doi: 10.3389/fdgth.2021.730948] [Medline: 34957461]
67. Hoffmann S, Wilson S. The role of serious games in the iManageCancer project. Ecancermedicalscience 2018;12:850 [FREE Full text] [doi: 10.3332/ecancer.2018.850] [Medline: 30079112]
68. Kato PM, Beale IL. Factors affecting acceptability to young cancer patients of a psychoeducational video game about cancer. J Pediatr Oncol Nurs 2006;23(5):269-275. [doi: 10.1177/1043454206289780] [Medline: 16902082]
69. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. Pediatrics 2008;122(2):e305-e317. [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
70. Kayali F, Silbernagl M, Peters K, Mateus-Berr R, Reithofer A, Martinek D, et al. Design considerations for a serious game for children after hematopoietic stem cell transplantation. Entertainment Comput 2016;15:57-73. [doi: 10.1016/j.entcom.2016.04.002]

71. Khan S, Abbasi AZ, Kazmi SF, Hooi TD, Rehman U, Hlavacs H, et al. Serious video games and psychological support: a depression intervention among young cancer patients. Entertainment Comput 2022;41:100479. [doi: 10.1016/j.entcom.2022.100479]
72. Khan S, Kazmi F, Maroof RY. Therapeutic outcomes of 3D-GIT game among depressive cancer patients. FWU J Soc Sci 2019;13(3):57-64.
73. Kolko DJ, Rickard-Figueroa JL. Effects of video games on the adverse corollaries of chemotherapy in pediatric oncology patients: a single-case analysis. J Consult Clin Psychol 1985 Apr;53(2):223-228. [doi: 10.1037//0022-006x.53.2.223] [Medline: 3858296]
74. Kowaluk A, Woźniewski M. Interactive video games as a method to increase physical activity levels in children treated for leukemia. Healthcare (Basel) 2022 Apr 06;10(4):692 [FREE Full text] [doi: 10.3390/healthcare10040692] [Medline: 35455869]
75. Lawitschka A, Buehrer S, Bauer D, Peters K, Silbernagl M, Zubarovskaya N, et al. A web-based mobile app (INTERACCT App) for adolescents undergoing cancer and hematopoietic stem cell transplantation aftercare to improve the quality of medical information for clinicians: observational study. JMIR mHealth uHealth 2020;8(6):e18781 [FREE Full text] [doi: 10.2196/18781] [Medline: 32602847]
76. Li WHC, Chung JOK, Ho EKY. The effectiveness of therapeutic play, using virtual reality computer games, in promoting the psychological well-being of children hospitalised with cancer. J Clin Nurs 2011;20(15-16):2135-2143. [doi: 10.1111/j.1365-2702.2011.03733.x] [Medline: 21651633]
77. Li WHC. Effectiveness of therapeutic play intervention in promoting the psychosocial well-being of children hospitalized with cancer. Pediatr Cancer 2012:31-54.
78. Linder LA, Newman A, Bernier Carney KM, Wawrzynski S, Stegenga K, Chiu YS, et al. Symptoms and daily experiences reported by children with cancer using a game-based app. J Pediatr Nurs 2022;65(11-12):33-43 [FREE Full text] [doi: 10.1016/j.pedn.2022.04.013] [Medline: 35490550]
79. Linder LA, Newman AR, Stegenga K, Chiu Y, Wawrzynski SE, Kramer H, et al. Feasibility and acceptability of a game-based symptom-reporting app for children with cancer: perspectives of children and parents. Support Care Cancer 2021;29(1):301-310. [doi: 10.1007/s00520-020-05495-w]
80. Masumoto D, Nakagami-Yamaguchi E, Nambu M, Maeda M, Uryu H, Hayakawa A, et al. A pilot study of game-based learning programs for childhood cancer survivors. BMC Cancer 2022;22(1):340 [FREE Full text] [doi: 10.1186/s12885-022-09359-w] [Medline: 35351044]
81. Nani S, Matsouka O, Theodorakis Y, Antoniou P. Perceived benefits of a therapeutic exercise program through digital interactive games among children with cancer. Hell J Nurs 2019;58(1):64-70.
82. Nani S, Matsouka O, Theodorakis Y, Antoniou P. Exergames and implications on quality of life in pediatric oncology patients: a preliminary qualitative study. J Phys Educ Sport 2019;19(Suppl 1):262-267. [doi: 10.7752/jpes.2019.s1039]
83. Nilsson S, Finnström B, Kokinsky E, Enskär K. The use of virtual reality for needle-related procedural pain and distress in children and adolescents in a paediatric oncology unit. Eur J Oncol Nurs 2009;13(2):102-109. [doi: 10.1016/j.ejon.2009.01.003] [Medline: 19230769]
84. Palmer SL, Leigh L, Ellison SC, Onar-Thomas A, Wu S, Qaddoumi I, et al. Feasibility and efficacy of a computer-based intervention aimed at preventing reading decoding deficits among children undergoing active treatment for medulloblastoma: results of a randomized trial. J Pediatr Psychol 2014;39(4):450-458 [FREE Full text] [doi: 10.1093/jpepsy/jst095] [Medline: 24369366]
85. Piil K, Gyldenvang HH, Møller JK, Kjoelsen T, Juul J, Pappot H. Electronic games for facilitating social interaction between parents with cancer and their children during hospitalization: interdisciplinary game development. JMIR Serious Games 2021;9(1):e16029 [FREE Full text] [doi: 10.2196/16029] [Medline: 33475523]
86. Sabel M, Sjölund A, Broeren J, Arvidsson D, Saury JM, Blomgren K, et al. Active video gaming improves body coordination in survivors of childhood brain tumours. Disabil Rehabil 2016;38(21):2073-2084. [doi: 10.3109/09638288.2015.1116619] [Medline: 26728453]
87. Sabel M, Sjölund A, Broeren J, Arvidsson D, Saury J, Gillenstrand J, et al. Effects of physically active video gaming on cognition and activities of daily living in childhood brain tumor survivors: a randomized pilot study. Neurooncol Pract 2017;4(2):98-110 [FREE Full text] [doi: 10.1093/nop/npw020] [Medline: 31385977]
88. Sajjad S, Abdullah A, Sharif M, Mohsin S. Psychotherapy through video game to target illness related problematic behaviors of children with brain tumor. Curr Med Imaging Rev 2014;10(1):62-72. [doi: 10.2174/1573405610666140313004302]
89. Schneider SM, Workman ML. Effects of virtual reality on symptom distress in children receiving chemotherapy. Cyberpsychol Behav 1999;2(2):125-134. [doi: 10.1089/cpb.1999.2.125] [Medline: 19178248]
90. Stinson JN, Jibb LA, Nguyen C, Nathan PC, Maloney AM, Dupuis LL, et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. J Med Internet Res 2013;15(3):e51 [FREE Full text] [doi: 10.2196/jmir.2350] [Medline: 23475457]
91. Tanriverdi M, Cakir E, Akkoyunlu ME, Cakir FB. Effect of virtual reality-based exercise intervention on sleep quality in children with acute lymphoblastic leukemia and healthy siblings: a randomized controlled trial. Pall Supp Care 2022;20(4):455-461. [doi: 10.1017/s1478951522000268]

92. Tark R, Metelitsa M, Akkermann K, Saks K, Mikkel S, Haljas K. Usability, acceptability, feasibility, and effectiveness of a gamified mobile health intervention (Triumf) for pediatric patients: qualitative study. JMIR Serious Games 2019;7(3):e13776. [doi: 10.2196/13776]
93. Schenck L, Bäumer C, Ross B, Schäfer G, Stember N, Thomas H, et al. Development of a smartphone virtual reality game to support the radiation therapy of children and adolescents in proton centers. Front Pediatr 2023;11:1-13. [doi: 10.3389/fped.2023.1163022]
94. Yang D, Lu M, Chen C, Liu P, Hou I. Development of a therapeutic video game with the MDA framework to decrease anxiety in preschool-aged children with acute lymphoblastic leukemia: mixed methods approach. JMIR Serious Games 2022;10(3):e37079. [doi: 10.2196/37079]
95. Chai CWE, Lau BT, Tee MKT, Al Mahmud A. Evaluating a serious game to improve childhood cancer patients' treatment adherence. Digit Health 2022;8:20552076221134457 [FREE Full text] [doi: 10.1177/20552076221134457] [Medline: 36339903]
96. Shaheen AAM, Masoud AE, Algabbani MF. Examination of the applicability and intensity of Nintendo Wiitm exergaming for children with acute lymphoblastic leukemia: a preliminary study. Int J Physiother 2023;10(2):30-36. [doi: 10.15621/ijphy/2023/v10i2/1320]
97. Russo L, Tozzi AE, Mastronuzzi A, Croci I, Gesualdo F, Campagna I, et al. Feasibility of a VR intervention to decrease anxiety in children with tumors undergoing CVC dressing. Int J Environ Res Public Health 2022;19(19):11953. [doi: 10.3390/ijerph191911953]
98. Oliveira H, Patraquim N, Lima H. The HOPE project—a case study on the development of a serious game to increase pediatric cancer patients' motivation. Act Learn Digital Transform Healthcare Educ, Train Res 2023:49-63. [doi: 10.1016/b978-0-443-15248-1.00001-1]
99. Nani S, Matsouka O, Theodorakis Y, Antoniou P. The impact of exergames on cancer related fatigue among pediatric oncology patients: a qualitative approach. Nosileftiki 2022;61(3):339-346 [FREE Full text]
100. Erdős S, Horváth K. The impact of virtual reality (VR) on psychological and physiological variables in children receiving chemotherapy: a pilot cross-over study. Integr Cancer Ther 2023;22:15347354231168984. [doi: 10.1177/15347354231168984]
101. Er CW, Theng L, Tsun M, Mahmud A. Increasing motivation for cancer treatment adherence in children through a mobile educational game: a pilot study. EAI Endorsed Trans Pervasive Health Technol 2022;8(30):173453. [doi: 10.4108/eai.15-2-2022.173453]
102. Bernier Carney KM, Stegenga K, Linder LA. Informing parents as caregivers with a symptom assessment app developed for children with cancer. J Pediatr Hematol Oncol Nurs 2022;39(4):264-272 [FREE Full text] [doi: 10.1177/27527530211073307] [Medline: 35791850]
103. Cederved C, Ångström-Brännström C, Ljungman G, Engvall G. Parents' experiences of having their children take part in participatory action research creating a serious game about radiotherapy. Radiography (Lond) 2023;29(1):95-100 [FREE Full text] [doi: 10.1016/j.radi.2022.10.005] [Medline: 36327521]
104. Atkinson M, Tully A, Maher CA, Innes-Wong C, Russo RN, Osborn MP. Safety, feasibility and efficacy of Lokomat® and Armeo®Spring training in deconditioned paediatric, adolescent and young adult cancer patients. Cancers (Basel) 2023;15(4):1250 [FREE Full text] [doi: 10.3390/cancers15041250] [Medline: 36831593]
105. Kerimoglu Yildiz G, Yildiz S, Yoruk MA, Sevgen S. A tablet game or training booklet? Two methods for evaluating symptom management and quality-of-life of children receiving chemotherapy. Eur J Oncol Nurs 2022;61(11-12):102190. [doi: 10.1016/j.ejon.2022.102190] [Medline: 36215817]
106. Cates JR, Fuemmeler BF, Stockton LL, Diehl SJ, Crandell JL, Coyne-Beasley T. Evaluation of a serious video game to facilitate conversations about human papillomavirus vaccination for preteens: pilot randomized controlled trial. JMIR Serious Games 2020;8(4):e16883. [doi: 10.2196/16883]
107. Inglés-Camats G, Presno-Rivas M, Antonijoan M, Garcia-Panella O, Forrest T. Yummy tricks: a serious game for learning healthy eating habits. In: Studies in Health Technology and Informatics. Amsterdam: IOS Press Ebooks; 2012. [doi: 10.3233/978-1-61499-088-8-185]
108. Lu AS, Buday R, Thompson D, Baranowski T. What type of narrative do children prefer in active video games? An exploratory study of cognitive and emotional responses. San Diego: Academic Press; 2016:137-155. [doi: 10.1016/b978-0-12-801738-8.00007-5]
109. Ruiz-López T, Sen S, Jakobsen E, Tropé A, Castle PE, Hansen BT, et al. FightHPV: design and evaluation of a mobile game to raise awareness about human papillomavirus and nudge people to take action against cervical cancer. JMIR Serious Games 2019;7(2):e8540. [doi: 10.2196/games.8540]
110. Mallafré-Larrosa M, Papi G, Trilla A, Ritchie D. Development and promotion of an mHealth app for adolescents based on the European code against cancer: retrospective cohort study. JMIR Cancer 2023;9(11-12):e48040 [FREE Full text] [doi: 10.2196/48040] [Medline: 38015612]
111. Hasni NHM, Azalian AAM, Kadir TAA, Hisham Mohd Adib MA. Serious "human papillomavirus vaccine" game for Malaysian adolescents: development and preliminary study. Int J Public Health Sci 2023;12(1):82-90. [doi: 10.11591/ijphs.v12i1.22258]

112. Anand P, Kunnumakkara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, et al. Cancer is a preventable disease that requires major lifestyle changes. Pharm Res 2008;25(9):2097-2116 [FREE Full text] [doi: 10.1007/s11095-008-9661-9] [Medline: 18626751]
113. Danaei G, Hoorn SV, Lopez AD, Murray C, Ezzati M. Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. Lancet 2005;366(9499):1784-1793. [doi: 10.1016/s0140-6736(05)67725-2] [Medline: 16298215]
114. The American Cancer Society medical and editorial content team. Can childhood cancers be prevented? American Cancer Society. 2024. URL: https://www.cancer.org/cancer/types/cancer-in-children/preventing-childhood-cancers.html [accessed 2024-06-07]
115. Wang Y, Wang Z, Liu G, Wang Z, Wang Q, Yan Y, et al. Application of serious games in health care: scoping review and bibliometric analysis. Front Public Health 2022;10(11-12):896974 [FREE Full text] [doi: 10.3389/fpubh.2022.896974] [Medline: 35757619]
116. Kim JY, Chun M, Kim SW, Song J. Identification of Korean cancer survivors' unmet needs and desired psychosocial assistance: a focus group study. PLoS One 2020;15(1):e0228054 [FREE Full text] [doi: 10.1371/journal.pone.0228054] [Medline: 31945143]
117. Onyeneho CA, Ilesanmi RE. Burden of care and perceived psycho-social outcomes among family caregivers of patients living with cancer. Asia-Pac J Oncol Nurs 2021;8(3):330-336. [doi: 10.4103/2347-5625.308678]

Abbreviations

APA: American Psychological Association
JBI: Joanna Briggs Institute
OSF: Open Science Framework
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by G Eysenbach; submitted 22.03.24; peer-reviewed by J Han, A Poliani; comments to author 10.04.24; revised version received 02.05.24; accepted 29.05.24; published 10.07.24.

Please cite as:
Kim S, Wilson P, Abraham O
Investigating the Use of Serious Games for Cancer Control Among Children and Adolescents: Scoping Review
JMIR Serious Games 2024;12:e58724
URL: https://games.jmir.org/2024/1/e58724
doi:10.2196/58724
PMID:38985502

Review

Gamification in Mobile Apps for Children With Disabilities: Scoping Review

Ebrahim Mahmoudi[1,2], BSc; Paul Yejong Yoo[3], PhD; Ananya Chandra[1,2], BA; Roberta Cardoso[4], PhD; Carlos Denner Dos Santos[5], PhD; Annette Majnemer[1,2,4], PhD; Keiko Shikako[1,2,4], PhD

[1]School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada

[2]Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR), Montreal, QC, Canada

[3]Division of Neurosciences and Mental Health, The Hospital for Sick Children, Toronto, ON, Canada

[4]McGill University Health Center Research Institute, Montreal, QC, Canada

[5]Département de systèmes d'information et méthodes quantitatives de l'École de gestion de l'Université de Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Keiko Shikako, PhD
School of Physical and Occupational Therapy
McGill University
3654 Prom Sir-William-Osler
Montreal, QC, H3G 1Y5
Canada
Phone: 1 514 488 5552 ext 1141
Email: keiko.thomas@mcgill.ca

Abstract

Background: Children with disabilities face numerous challenges in accessing health services. Mobile health is an emerging field that could significantly reduce health inequities by providing more accessible services. Many mobile apps incorporate gamification elements such as feedback, points, and stories to increase engagement and motivation; however, little is known about how gamification has been incorporated in mobile apps for children with disabilities.

Objective: This scoping review aims to identify and synthesize the existing research evidence on the use of gamification in mobile apps for children with disabilities. Specifically, the objectives were to (1) identify the categories of these mobile apps (eg, treatment and educational) (2), describe the health-related outcomes they target, (3) assess the types and levels of gamification elements used within these apps, and (4) determine the reasons for incorporating gamification elements into mobile apps.

Methods: We searched MEDLINE, PsycINFO, CINAHL, Embase, the ACM Digital Library, and IEEE Xplore databases to identify papers published between 2008 and 2023. Original empirical research studies reporting on gamified mobile apps for children with disabilities that implemented at least 1 gamification strategy or tactic were included. Studies investigating serious games or full-fledged games were excluded.

Results: A total of 38 studies reporting on 32 unique gamified mobile apps were included. Findings showed that gamified apps focus on communication skills and oral health in children with autism spectrum disorder while also addressing self-management and academic skills for other disability groups. Gamified mobile apps have demonstrated potential benefits across different populations and conditions; however, there were mixed results regarding their impact. The gamification strategies included fun and playfulness (23/32, 72%), feedback on performance (17/32, 53%), and reinforcement (17/32, 53%) in more than half of apps, whereas social connectivity was used as a gamification strategy in only 4 (12%) mobile apps. There were 2 main reasons for integrating gamification elements into mobile apps described in 16 (42%) studies: increasing user engagement and motivation and enhancing intervention effects.

Conclusions: This scoping review offers researchers a comprehensive review of the gamification elements currently used in mobile apps for the purposes of treatment, education, symptom management, and assessment for children with disabilities. In addition, it indicates that studies on certain disability groups and examinations of health-related outcomes have been neglected, highlighting the need for further investigations in these areas. Furthermore, research is needed to investigate the effectiveness of mobile-based gamification elements on health and health behavior outcomes, as well as the healthy development of children with disabilities.

(JMIR Serious Games 2024;12:e49029) doi:10.2196/49029

KEYWORDS

mobile health; mHealth; gamification; children with disabilities; mobile phone

Introduction

Background

Worldwide, approximately 93 million children have a moderate to severe disability, and 13 million children have a severe disability [1]. Children with disabilities face numerous barriers to accessing health services and health-promoting activities [2]. Despite the abundance of research investigating different interventions to improve the lives of children with disabilities, the interventions have not been successfully implemented, limiting the impact of research on public health outcomes [3]. To address this issue, innovative technological advances could significantly improve the health and well-being of marginalized groups [4,5], such as children with disabilities, their families, and the systems of care surrounding them.

Mobile apps, as examples of innovative technological tools, are becoming important in improving access to therapeutic interventions and diagnoses for underserved groups [6]. Mobile health (mHealth), a young but rapidly evolving field, enables the delivery of planned interventions and practices via mobile devices and apps, downloaded and installed on mobile devices to perform a particular task [7]. Constant availability, broader access, fairness of service offerings, personalized content, lower cost, and increased service capacity and efficiency are some advantages of mHealth [8]. Therefore, mHealth can be a tool to create more accessible services for children and adolescents with disabilities and their families as applied to various health-related situations.

There is a growing interest in incorporating game-like elements, called "gamification," in mobile apps to promote greater engagement with the technology and motivation to achieve specific personalized goals [9,10]. Gamification is the application of various game strategies and tactics in nongame contexts [11-13]. Gamification aims to change individual behavior through a combination of game elements (often delivered within games but also through mobile apps) [14,15] in contrast to "serious games" that is "any form of interactive computer-based game software for one or multiple players to be used on any platform and that has been developed with the intention to be more than entertainment" [16]. Although gamification is a promising concept [14], the empirical research regarding its applications is still in its early stages.

Gamification is increasingly being applied in mHealth to promote healthy behaviors using a wide range of game elements, including challenges, goal setting, feedback, progress bars, points, and levels [9]. There is an increasing trend toward incorporating gamification in different health domains, such as pediatric rehabilitation, physical activity, and chronic health conditions [5,17,18]. Gamification in rehabilitation can enhance therapeutic adherence and can be used to complement traditional interventions for children with disabilities and promote physical activity and other healthy behaviors [19]. The implementation of gamification has the potential to enhance individuals' adherence to medical protocols and successfully manage their health conditions [20-22]. The inclusion of social support as a gamification component has been recognized as encouraging for enhancing one's social abilities [23-25]. Previous research has also indicated that the use of gamification has the potential to trigger desirable emotional experiences and enhance users' levels of satisfaction and self-esteem [26-30]. Furthermore, game design components have become more accessible, cost-effective, and enticing as video games have gained popularity [10].

Ryan et al [31] introduced an integrative process model that incorporates the fundamental elements of self-determination theory, which is a motivational theory. They argued that actions can support or thwart the satisfaction of basic psychological needs, namely autonomy, competence, and relatedness and consequently influence the quality of motivation. Depending on whether the individual's needs are supported or not, it may then influence mental health outcomes (eg, depression and anxiety) and physical health outcomes (eg, exercise and weight control) [31]. Research has demonstrated that gamification can both facilitate and diminish intrinsic motivation [32]. Therefore, the integration of gamification features in mobile apps entails certain nuances.

Gamification elements such as rewards have the potential to enhance motivation toward continued performance and consequently healthier behaviors; however, numerous research studies indicate that the use of extrinsic motivators or the provision of controlling feedback can significantly diminish intrinsic motivation by impeding individuals' sense of autonomy [33,34]. The presence of increased levels of extrinsic motivation in the context of gamification is not sufficient as the only criterion for evaluating its advantages [35]. Cheating may also escalate as individuals get involved in attempts to attain greater levels of achievement, primarily driven by the rewards [23] Furthermore, there is a prevailing prediction that a significant proportion of gamification implementations will be doomed to failure because of inadequate understanding regarding the effective design principles of gamification [36]. The development of gamified health solutions frequently lacks collaboration with health professionals, potentially compromising their efficacy and diminishing their credibility [23,28]. When gamification to promote health fails to prioritize the user-centered approach and neglects to consider the unique attributes and demographic factors of potential users, their effectiveness may be undermined [23,28,37,38].

As a result, tailoring the gamification features based on the users' profiles is crucial to enhance their engagement [39]. Given the diverse needs of children with disabilities, it is imperative for researchers and mobile developers to possess a comprehensive understanding of gamification principles and strategies. This knowledge will enable them to effectively customize gamification features to cater to the specific requirements of this target population. Although there is a growing interest in using gamification elements in mobile apps, there is still a lack of comprehensive understanding in the field of childhood disability. Currently, there is no literature review investigating gamification in mobile apps designed for children with disabilities.

Our scoping review aimed to bridge the following knowledge gaps:

First, there is a deficiency in the systematic identification and categorization of gamified mobile apps (eg, treatment and educational) that are specifically designed for children with disabilities. Understanding the existing evidence in this niche field is crucial to evaluating the scope and diversity of available gamified mobile apps.

Second, there is a lack of comprehensive documentation on the specific health-related outcomes these apps target. Understanding the existing evidence will help us recognize which health-related outcomes have been targeted in this population and help identify disabilities and health-related outcomes that could have been neglected.

Third, there is a lack of comprehensive documentation on the characteristics of gamification strategies and tactics used in these mobile apps. A comprehensive review of gamification types and levels is required to understand how the game elements address the unique needs of children with disabilities.

Fourth, the underlying justification for incorporating gamification elements into mobile apps for this specific group remains unclear.

Objective

Addressing the abovementioned knowledge gaps is imperative to advance the use of gamification in mobile apps for children with disabilities. Therefore, this review aimed to explore the current use of gamification strategies and tactics in mobile apps for children with disabilities. The four specific objectives are as follows: (1) to identify gamified mobile apps designed for children with disabilities, (2) to identify health-related outcomes that these mobile apps aim to target, (3) to identify the different types and levels of gamification strategies and tactics implemented in these mobile apps (4) to determine the reasons for incorporating gamification elements into mobile apps

Methods

Overview

Scoping reviews help identify the types of current literature in a specific field and key characteristics related to a particular context, and analyze the knowledge gaps, while systematic reviews investigate the conflicting results and address any variation in current practices, or compare new interventions against gold standard, established interventions [40]. As there was no current review exploring the current evidence in gamification for children with disabilities and identifying the types of gamification elements in these mobile apps, we sought a scoping review to answer our research objectives. The methodological frameworks proposed by Arksey and O'Malley [41] and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [42] (Multimedia Appendix 1) were used to guide this scoping review.

Search Strategy

The database searches were performed in November 2023 in the following web-based databases: MEDLINE, PsycINFO, CINAHL, Embase, the ACM Digital Library, and IEEE Xplore. The selection of databases, keywords, and relevant indexing (eg, Medical Subject Headings and other database-specific search techniques) were finalized in collaboration with the experienced librarian. The full search strategy is presented in Multimedia Appendix 2. In summary, we had 2 main themes: children with disabilities (population) and gamification in mobile apps (exposure). Regarding the full search strategy used on MEDLINE for the first theme, we combined different key disability terms (lines 1-51) with pediatric population terms (lines 53-55) and parent-related terms (lines 57-62). For the second theme, we combined key terms for mobile apps (lines 64-67) and gamification (line 68). The combination of these 2 themes helped us find any papers that studied mobile apps for children with disabilities. The inclusion criteria for the selection of the papers are discussed in Inclusion Criteria section. To ensure the comprehensiveness of the search, the primary author (EM) manually searched the reference lists of the relevant studies and existing reviews. Furthermore, EM searched the JMIR homepage [43], where no new studies were found. All the research results found in the databases were imported to the Rayyan reference manager website [44], where duplicates were identified and removed.

Inclusion Criteria

We included a peer-reviewed research article if the conditions presented in Textbox 1 were met.

Textbox 1. Inclusion criteria for included peer-reviewed research articles.

Inclusion criteria

- Publication language: there were no limits imposed on the language of the studies.
- Type of publication: peer-reviewed journal articles and conference proceedings
- Type of study: qualitative, quantitative, or mixed methods
- Time: published between January 2008 and November 2023. The reason for the selected start date was that the App Store and Google Play were launched in 2008, and almost all mobile apps were developed after 2008. Furthermore, another reason for the start date was that the concept of gamification was first introduced by Deterding et al [13] in 2008
- Population: children (aged up to 18 years) with any of the following disabilities: autism spectrum disorder, developmental delays, cerebral palsy, attention-deficit/hyperactivity disorder, dyslexia, intellectual disabilities, Turner syndrome, deglutition disorders, child behavior disorders, speech disorders, sensory disorder, motor disability, brain injuries (eg, traumatic brain injury), or any other brain-based disabilities
- Exposure: mobile apps on any device (smartphone, tablet, or iPad) and platform (Android or iOS) designed for children with disabilities. The mobile apps were included if they incorporated at least 1 gamification element (gamification strategy or tactics)
- Outcome: any health-related outcome that relates to the child's developmental functioning and general health status

We did not include theses, dissertations, protocols, abstracts, and letters to the editor; however, their references were screened for relevant studies. Nongamified mobile apps were excluded. Furthermore, apps labeled as "serious games" were excluded as they are complete games and fall outside of the scope of this review. Given the unique characteristics of mental health conditions and other disorders such as obesity and cancer in children and adolescents, we excluded these disorders; however, if there was any health-related comorbidity among children with disabilities (eg, if the study was on children with disabilities who are obese), we included them.

Study Selection

First, we tested the selection criteria, with 2 reviewers (EM and PYY) screening titles and abstracts independently until we reached an interrater agreement of 90%. The same process was followed for the full-text review of potentially relevant studies but with 2 dyads of reviewers (EM and PYY and EM and AC). Upon full-text screening, 1 reviewer (EM) manually searched the target journal and the reference lists of the included articles, abstracts, protocols, etc, and no relevant articles were found. Any disagreements were resolved through discussion to reach a consensus on a final decision, or a third adjudicator was implicated (KS and RC).

Data Abstraction and Charting

The data extraction form was developed and calibrated among each dyad of reviewers (EM and PYY and EM and AC) with 3 random articles. As the percent agreement was greater than 90% in each dyad, the data abstraction of the remaining articles began, and the conflicts were resolved through discussion. For each study, we extracted data on the study's first author, country, study design, population (eg, autism), sample characteristics (eg, size and age), mobile app name, device (smartphone, tablet, or iPad), platform (Android or iOS), app purpose, type of gamification strategy and tactics, health-related outcomes, and any reasons for implementing gamification in the mobile apps.

Data Synthesis and Analysis

Both quantitative and qualitative analyses were performed. A frequency analysis was conducted to illustrate the distribution of studies by publication year, country of origin, disability type in studies, gamification strategies and tactics, and the gamification level incorporated by mobile apps.

The gamification framework proposed by Cugelman [12] was used to assess the gamification elements present in these mobile apps. This framework consists of two sections: (1) gamification *strategies*, which are the persuasive principles of gamification and (2) gamification *tactics*, which are the on-screen features of gamification that app users interact with. The concepts of gamification tactics and strategies proposed by Cugelman [12] were used to operationalize gamification in this review. This framework consists of 7 gamification strategies and 10 gamification tactics.

Descriptive statistics were calculated to examine the level of gamification incorporation into mobile apps, as we wanted to understand if the number of gamification features might influence the outcomes. As there is no previous research exploring the level of mobile-based gamification for children with disabilities, we used arbitrary cutoff points to estimate the gamification level used in previous research in a different field [45]. The level of gamification strategies was labeled as none (no gamification strategies), low (1-2 gamification strategies), medium (3-5 gamification strategies), and high (6-7 gamification strategies). Similarly, the level of gamification tactics was classified as none (no gamification tactics), low (1-3 gamification tactics), medium (4-7 gamification tactics), and high (8-10 gamification tactics).

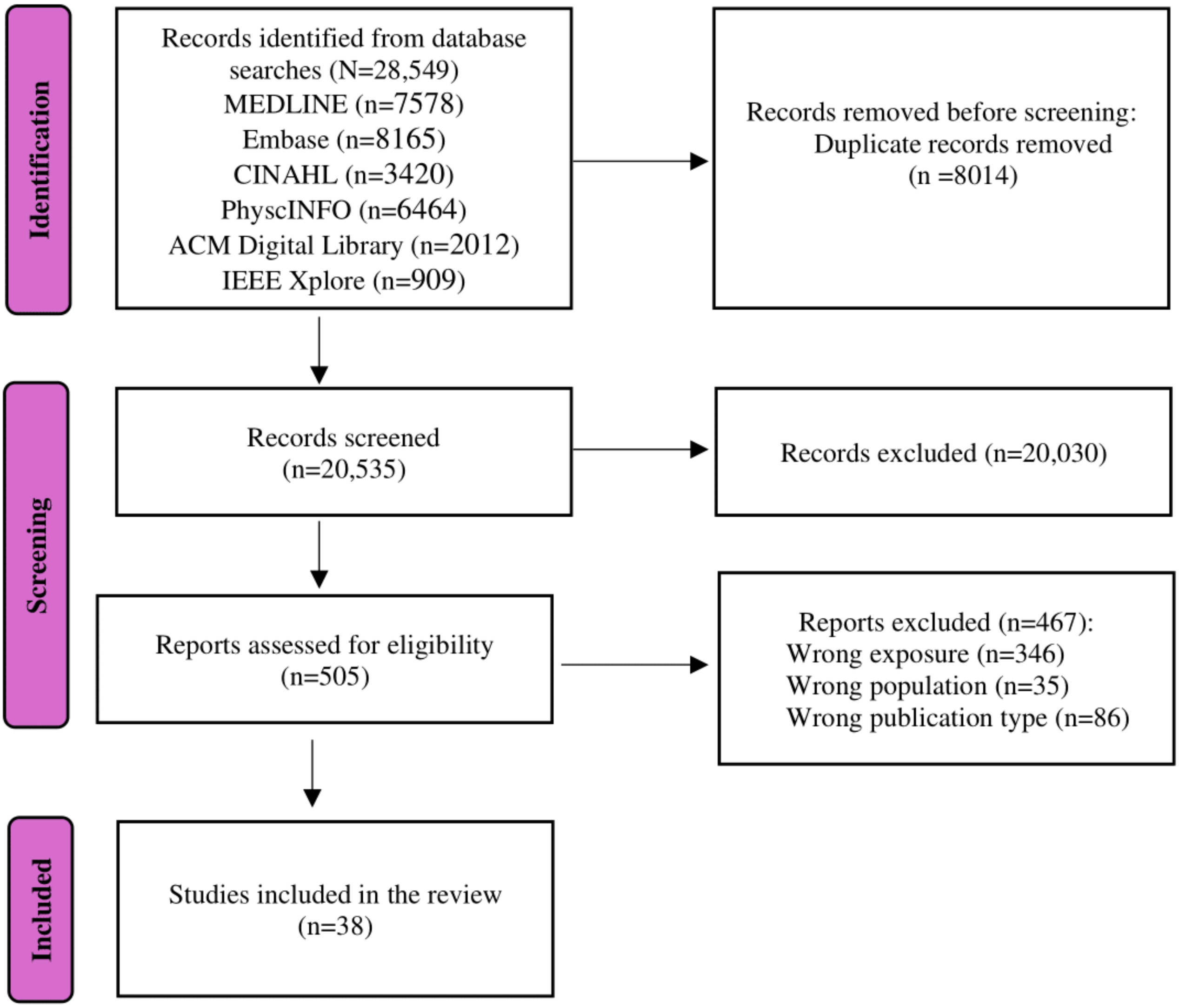
The primary author (EM) performed the content analysis to identify the health-related outcomes targeted by these mobile apps and the rationale for applying gamification in the apps. Further verification was done through discussion and collaboration with another author (RC) with expertise in conducting reviews and data synthesis.

Results

Study Selection

The flowchart of the search strategy and study selection is depicted in Figure 1. The initial database search yielded 28,549 citations; after the removal of duplicates, 20,535 (71.93%) citations remained for the title and abstract screening. The first screening phase led to 505 (2.46%) included and 20,030 (97.54%) excluded documents. The studies were excluded because they did not fit our inclusion criteria (eg, wrong population and wrong exposure). The second screening phase consisted of a full-text review of the 505 (2.46%) included documents, resulting in 38 (7.5%) included studies for this scoping review.

Figure 1. Flowchart of the review process.



Study Characteristics

Table 1 presents an overview of the study characteristics of all 38 articles included in the scoping review. Although we selected 2008 as the beginning year, all studies were published after 2013, demonstrating that gamification is a recently evolving field. The studies were implemented worldwide, with 37% (14/38) of the studies from Asian countries and 24% (9/38) of the studies from the United States. Most studies (29/38, 76%) used a quantitative research approach and were primarily quasi-experimental (9/29, 31%) or randomized clinical trials (RCTs; 6/29, 21%) studies. All 38 articles included in this review were written in English. Autism spectrum disorder (ASD) was the most common condition reported in 18 (47%) of the 38 included studies, followed by vision impairment (4/38, 11%) and dyslexia (4/38, 11%). Multimedia Appendix 3 [46-83] demonstrates details regarding the study and participant characteristics.

Table 1. Characteristics of included studies (n=38).

Characteristics	Value, n (%)
Year of publication	
2018-2023	23 (61)
2013-2017	15 (39)
2008-2012	0 (0)
Country	
United States	9 (24)
Australia	5 (13)
Malaysia	4 (11)
United Kingdom	3 (8)
Canada	2 (5)
Hungary	2 (5)
Singapore	2 (5)
Indonesia	2 (5)
Spain	1 (3)
Japan	1 (3)
Türkiye	1 (3)
Saudi Arabia	1 (3)
United Arab Emirates	1 (3)
India	1 (3)
Myanmar	1 (3)
South Africa	1 (3)
Romania	1 (3)
Study design	
Quantitative	26 (68)
Mixed methods	8 (21)
Qualitative	4 (11)
Disability	
Autism spectrum disorder	18 (47)
Vision impairments	4 (11)
Dyslexia	4 (11)
Attention-deficit/hyperactivity disorder	2 (5)
Mild traumatic brain injury	2 (5)
Neurodevelopmental disabilities	2 (5)
Physical disabilities	1 (3)
Pervasive developmental disorder	1 (3)
Mild intellectual disabilities	1 (3)
Complex needs (physical disabilities, learning, and communication difficulties)	1 (3)
Hearing impairments	1 (3)
Concussion	1 (3)

Mobile App Characteristics

Table 2 [46-83] demonstrates results regarding the general characteristics of mobile apps and gamification. A total of 32 unique gamified mobile apps were identified in the 38 included studies. Approximately 44% (14/32) of the identified mobile apps fell into the treatment category, designed to help children with disabilities improve their skill competencies, such as story creation and story sharing, and social communication. Educational (n=13, 41%) and assessment (n=3, 9%) apps were the second and third most observed categories. One app was specifically designed for managing symptoms.

Table 2. Summary descriptions of studies included in the scoping review.

Study, year	App name	App category	Health-related outcome	Gamification strategies	Gamification tactics
Kucirkova et al [46], 2014	Our story	Treatment	Social communication and story-telling abilities	Goal setting, capacity to overcome challenges, fun, and playfulness	Provides clear goals and offers a challenge
Moore et al [47], 2015	TOBY	Treatment	Different rehabilitation goals, such as sensory awareness, imitation, and social interaction	Goal setting, feedback on performance, reinforcement, comparing progress, fun, and playfulness	Offers a challenge, levels, points, shows progress, feedback, and gives reward
Parsons et al [48], 2020	TOBY	Treatment	Different rehabilitation goals, such as receptive language, social skills, and pragmatic language	Goal setting, feedback on performance, reinforcement, comparing progress, fun, and playfulness	Offers a challenge, levels, points, shows progress, feedback, gives reward
Parsons et al [49], 2019	TOBY	Treatment	Social communication	Goal setting, feedback on performance, reinforcement, comparing progress, fun, and playfulness	Offers a challenge, levels, points, shows progress, feedback, and gives reward
Penev et al [50], 2021	Guess what	Treatment	Social communication	Goal setting, the capacity to overcome challenges, feedback on performance, reinforcement, fun, and playfulness	Provides clear goals, offers a challenge, points, feedback, gives rewards, and provides badges
Saputra [51], 2016	LexiPal	Educational	Enjoyment and motivation in learning	Goal setting, feedback on performance, reinforcement, fun, and playfulness	Provides clear goals, levels, points, feedback, gives rewards, provides badges, and story or theme
Schmidt et al [52], 2020	SMART	Treatment	Self-management and relaxation	Goal setting, feedback on performance, reinforcement, fun, and playfulness	Provides clear goals, feedback, gives rewards, provides badges, and story or theme
Thida et al [53], 2020	VOIS	Treatment	Language	The capacity to overcome challenge, feedback on performance, reinforcement, fun, and playfulness	Offers a challenge, shows progress, feedback, and give rewards
Urakami [54], 2021	GROWJECTOR	Treatment	Medication adherence	Goal setting, reinforcement, and social connectivity	Provides clear goals, points, and gives rewards
Ying et al [55], 2016	NR[a]	Educational	Learning road safety	Fun and playfulness	Story or theme
Chua et al [56], 2017	NR	Educational	Emotional learning	Feedback on performance, fun, and playfulness	Feedback and story or theme
Doenyas et al [57], 2014	NR	Educational	Sequencing skill	Goal setting, the capacity to overcome challenges, feedback on performance, and reinforcement	Offers a challenge, feedback, and points
Holmes et al [58], 2016	NR	Treatment	Visual acuity	Reinforcement and compare progress	Levels, points, and shows progress
Kelly et al [59], 2016	Dig Rush	Treatment	Visual acuity	Goal setting, fun, and playfulness	Levels and points
Aburukba et al [60], 2017	AutiAid	Treatment or symptoms management	Memory and management of symptoms	Goal setting and social connectivity	Provides clear goals, levels, and points
Alnaghaimshi et al [61], 2020	Autism-World	Assessment and educational	User's literacy on autism	Social connectivity	__[b]
Barta et al [62], 2017	AutiSoft	Symptoms management	Manage daily routines	Feedback on performance and reinforcement	Feedback and gives rewards
Birtwell et al [63], 2019	SideKicks!	Treatment	Social communication	Goal setting, fun, and playfulness	Provides clear goals and story or theme

Study, year	App name	App category	Health-related outcome	Gamification strategies	Gamification tactics
Borhan et al [64], 2018	Mr. Read	Educational	Reading skills	Feedback on performance, reinforcement, fun, and playfulness	Points, feedback, shows progress, and story or theme
Brkic et al [65], 2022	FarmApp	Assessment	No specific behavior	Feedback on performance, reinforcement, fun, and playfulness	Offers a challenge, feedback, points, and story or theme
Daud and Abas [66], 2013	Dyslexia Baca	Educational	Letter recognition	Feedback on performance, reinforcement, fun, and playfulness	Feedback, gives reward, and story or theme
Dehkordi and Rias [67], 2014	GO-Go	Educational	Multiple cues responding	Reinforcement, fun, and playfulness	Offers a challenge, gives rewards, and story or theme
Gómez and Carro [68], 2014	AdaptAD-HD	Treatment	Concentration and impulse control abilities	Goal setting	Provides clear goals and levels
Guzsvinecz et al [69], 2017	Sliders	Educational	Logical thinking and deductive reasoning	Feedback on performance and reinforcement	Points, feedback, and shows game leaders
Hu et al [70], 2019	NeuroCare	Symptoms management	Self-management of pediatric concussion	Goal setting and feedback on performance	Provides clear goals and feedback
Irwin et al [71], 2015	Listening to Faces (L2F)	Educational	Audiovisual speech perception	Feedback on performance, compare progress, fun, and playfulness	Feedback and shows progress
Kalantarian et al [72], 2019	Guess what	Treatment	Social communication	Goal setting, the capacity to overcome challenges, feedback on performance, reinforcement, fun, and playfulness	Provides clear goals, offers a challenge, points, feedback, gives rewards, and provides badges
Macdonald et al [73], 2022	NR	Educational	Reading skills	Feedback on performance, reinforcement, fun, and playfulness	Feedback, gives rewards, and story or theme
Manh et al [74], 2018	NR	Treatment	Visual acuity	Capacity to overcome challenges and reinforcement	Levels and points
Mwamba et al [75], 2019	Paediatric Attention-Deficit/Hyperactivity Disorder Application Software (PANDAS)	Assessment	No specific behavior	Capacity to overcome challenges, reinforcement, fun, and playfulness	Offers a challenge, points, and story or theme
Tang et al [76], 2021	ColourSpot	Assessment	No specific behavior	Feedback on performance, reinforcement, Fun and playfulness	Feedback, gives rewards, and story or theme
Cahyono [77], 2022	LexiPal	Educational	Enjoyment and motivation in learning	Goal setting, feedback on performance, reinforcement, fun, and playfulness	Provides clear goals, levels, points, feedback, gives rewards, provides badges, and story or theme
Chistol et al [78], 2023	Autism Assistant	Treatment	Multiple behavioral skills	Fun and playfulness, goal setting, compare progress, and capacity to overcome challenges	Story or theme, provides clear goals, shows progress, and offers a challenge
Tan et al [79], 2023	NUS care	Educational	Oral health	Fun and playfulness, reinforcement, and social connectivity	Story or theme, points, and rewards
Schmidt et al [80], 2022	Self-Monitoring Activity Regulation and Relaxation Treatment (SMART)	Treatment	Self-management and relaxation	Goal setting, feedback on performance, reinforcement, fun, and playfulness	Provides clear goals, feedback, gives rewards, provides badges, and story or theme
Johnson et al [81], 2022	Zingo	Treatment	Therapy adherence	Fun and playfulness, feedback on performance, reinforcement, goal setting, and compares progress	Feedback, points, story or theme, gives rewards, provides clear goals, and shows progress

Study, year	App name	App category	Health-related outcome	Gamification strategies	Gamification tactics
Johnson et al [82], 2023	Zingo	Treatment	Therapy adherence	Fun and playfulness, feedback on performance, reinforcement, goal setting, and compares progress	Feedback, points, story or theme, gives rewards, provides clear goals, and shows progress
Krishnan et al [83], 2021	Brush Up	Educational	Oral health	Fun and playfulness, reinforcement, feedback on performance	Story or theme, feedback, and gives rewards

[a]NR: not reported.

[b]Not applicable.

A variety of health-related outcomes were identified, including a wide range of developmental, therapeutic, and educational skill competencies. Social communication (5/32, 16%), self-management (4/32, 11%), visual acuity (3/32, 9%), and oral health (2/32, 6%) were the most observed outcomes in the review. Only 2 (6%) studies investigating the LexiPal app targeted psychological outcomes such as motivation and engagement.

Multimedia Appendix 3 demonstrates other details regarding the participant and mobile app characteristics [46-83]. Regarding the 32 platforms, 14 (44%) apps were exclusively designed for iOS and 10 (31%) apps for Android, with 3 (9%) apps available on both platforms. In addition, 3 (9%) studies did not report the platform used. Of the 32 apps, 12 (38%) gamified apps were delivered on smartphones, 10 (31%) exclusively on iPads, 5 (16%) on both smartphones and tablets, 3 (9%) on both smartphones and iPads, and 2 (6%) exclusively on tablets. Most studies (28/32, 88%) did not report on costs related to app development.

Gamification Characteristics

Table 3 outlines the number and percentage of each gamification strategy and tactic adopted by 32 gamified mobile apps in this review. The most popular gamification strategy among 32 mobile apps was fun and playfulness (n=23, 72%), resulting in a higher number of gamification tactics (on-screen features) such as story or theme, avatars, a graphic representation of story characters, fun videos, and sound effects.

Table 3. Number of gamification strategies and tactics (n=32).

	Value, n (%)	Reference
Gamification strategies		
Fun and playfulness	23 (72)	[46-53,55,56,59,63-67,71,73,75,76,78,79,81,82]
Feedback on performance	17 (53)	[47-53,56,57,62,64-66,69-73,76,81,82,83]
Reinforcement	17 (53)	[47-54,57,58,62,64-67,69,72-77,79-82,83]
Goal setting	14 (44)	[46-52,54,57,59,60,63,68,70,72,78,81,82]
The capacity to overcome challenges	7 (22)	[46,50,53,57,72,74,75,78]
Compares progress	5 (16)	[47-49,58,71,78,81,82]
Social connectivity	4 (12)	[54,60,61,79]
Gamification tactics		
Feedback	18 (56)	[47-53,56,57,62,64-66,69-73,76,81,82,83]
Points	14 (44)	[47-51,54,57-60,64,65,69,74,79,81,82]
Story or theme	14 (44)	[51,52,55,56,63-67,73,75,76,78,79,81,82,83]
Gives rewards	13 (41)	[47-54,62,66,67,72,73,76,79,81,82,83]
Provides clear goals	11 (34)	[46,50-52,54,60,63,68,70,72,78,81,82]
Offers a challenge	8 (25)	[46-50,53,57,65,67,72,75,78]
Levels	7 (22)	[47-49,51,58-60,68,74]
Shows progress	7 (22)	[47-49,53,58,64,71,78,81,82]
Provides badges for achievements	3 (9)	[50-52,72]
Shows game leaders	0 (0)	—[a]

[a]Not applicable.

Furthermore, more than half (17/32, 53%) of mobile apps adopted feedback on performance and reinforcement (17/32, 53%). This finding is consistent with the high presence of on-screen features such as visual and verbal feedback, providing points and stars, and giving monetary and nonmonetary rewards upon completing a specified task. Finally, social connectivity was the least common gamification strategy observed in the apps (4/32, 12%), resulting from a low presence of on-screen social connectivity features. Only 12% (4/32) apps provided a social connection like a chat room for users where they can send messages [54,60,61,79]. The most common tactics, meanwhile, were feedback, points, and rewards. None of the apps displayed who the game leaders were.

Table 4 demonstrates the levels of gamification strategies and tactics adopted by 32 gamified mobile apps. Only 1 (3%) app did not adopt any gamification tactics [61]. Although more than half (18/32, 56%) of the mobile apps had adopted a medium level of gamification strategies, only 25% (8/32) of the mobile apps had incorporated a medium level of gamification tactics, known as on-screen features.

Table 4. Level of gamification incorporated in mobile apps (N=32).

	Value, n (%)	Reference
Number of gamification strategies adopted		
0 (none)	0 (0)	—[a]
1-2 (low)	13 (41)	[55,56,58-63,67-70,74]
3-5 (medium)	18 (56)	[46-49,51-54,57,64-66,71-73,75,76,78,79,81,82,83]
6-7 (high)	1 (3)	[50,72]
Number of gamification tactics adopted		
0 (none)	1 (3)	[61]
1-3 (low)	23 (72)	[46,54-60,62,63,65-71,73-76,79,83]
4-7 (medium)	8 (25)	[47-53,64,72,78,81,82]
8-10 (high)	0 (0)	—

[a]Not applicable.

We identified the rationale for using gamification in apps in 16 (42%) of the 38 included studies. Multimedia Appendix 4 [50-52,56,57,60,62-66,76,77,80-82] presents the complete results. The two most cited reasons were (1) to promote user engagement and motivation and (2) to increase the intervention effects. Some of the underlying reasons for the first theme are as follows: encourage use [52,65] and increase engagement with the intervention [50,81]. For example, the SMART app and FarmApp provided feedback on performance and offered different reinforcement features to keep users more involved with the app content [52,65].

Regarding the second theme, the researchers applied gamification to enhance learning [51,56,63,64] and increase intervention efficacy [66,76,81,82]. As an example, the Zingo app incorporated a digital pet where users, by adhering to the prescribed therapies in the app, receive stars to make changes to avatars. In addition, children and clinicians can monitor their weekly progress in the app.

Discussion

Principal Findings

This scoping review aimed to offer an overview of existing research using gamification in mobile apps for children with disabilities. A total of 38 studies and 32 unique mobile apps were identified, and most incorporated a limited number of gamification strategies and tactics. Of the 32 apps, 18 (56%) were specifically designed for children with ASD, while 14 (44%) were for children with other types of disabilities.

Social communication impairments are a clinical indicator of ASD [84]. Notably, mobile apps designed for children with ASD identified in this scoping review were predominantly focused on enhancing communication and social skills [46,49,63,72]. Our review found that 2 educational apps were designed specifically for children with ASD to acquire knowledge and skills related to oral health [79,83].

This review identified 4 (12%) of the 32 mobile apps that specifically target the self-management of children with disabilities. Previous research indicates that effective self-management behaviors could enhance health-related outcomes in children with complex needs [85]. Self-management interventions for individuals with intellectual disabilities primarily target self-management in the workplace, self-management of medical conditions, and self-management of daily activities [86]. Similar to the previous literature, the 4 (12%) mobile apps identified in this review focused on the management of symptoms in children with ASD [60], traumatic brain injury [52,80], and concussion [70] and on managing the daily routines of children with ASD [62].

Children with developmental delays, including those with learning disabilities, ASD, and attention-deficit/hyperactivity disorder, generally have lower academic achievements than those without developmental delays. The most impacted domains are cognitive, attention and memory, visual-motor skills, and behavioral functioning [87]. Similarly, several mobile apps identified in this review targeted memory [60], reading skills [73], letter recognition [66], multiple cues for responding [67], concentration and impulse control abilities [68], audiovisual speech perception [71], and behavioral skills [78] to help children with disabilities attain a higher level of academic achievement.

The major objective of the 9% (3/32) assessment mobile apps in this review was to assess users' cognitive control and memory [65], visual acuity [76], and screen children with attention-deficit/hyperactivity disorder [75]. These apps did not focus on any specific changes in behavioral, developmental, or other health-related outcomes.

The Gamified Apps and Their Impact

The findings of our scoping review shed light on the relationship between gamified mobile apps and health-related outcomes in children with different disabilities. Multimedia Appendix 5 [46,48-50,54,57-59,64,72-74,77,83] provides a comprehensive overview of 15 studies that have reported the impact of gamified mobile apps as interventions across different populations and conditions. Quantitative studies show mixed results, with significant improvements in targeted health-related outcomes such as social responsiveness, language skills, and visual acuity. However, studies with comparison groups often reveal that traditional methods (eg, patching for amblyopia) may still be more effective in some cases. For example, in 3 studies on children with amblyopia (a type of visual disorder where usually one eye gets poor vision), the 3 mobile apps had treatment goals that aimed to improve the visual acuity of these children. The results of 2 RCT studies showed an improvement in visual acuity using the mobile app [58,59], whereas in 1 RCT, there was no difference in visual acuity level using the mobile app compared to traditional intervention (patching) and even was less effective [74].

Furthermore, gamified systems can intrinsically motivate individuals to start and maintain the execution of healthy behaviors [88]. A meta-analysis by Bai et al [89] showed that gamification can improve student learning outcomes by fostering motivation among learners. In addition, 2 studies using the LexiPal mobile app among children with dyslexia focused on psychological outcomes such as motivation and engagement [51,77]. The study by Cahyono [77] showed that LexiPal has the potential to increase extrinsic motivation through a reward system and intrinsic motivation through activity levels and fun features; however, a longer intervention is necessary to assess the impact of gamification on long-term motivation and engagement of learners.

Moreover, the importance of user preferences and the need for more personalized gamification is highlighted. For example, the variability in the effectiveness of mobile apps in children with different levels of disability in 2 (5%) of the 38 identified studies suggests that personalization is crucial [46,49]. Personalized (or adaptive) gamification is a method for enhancing the design of game-based systems by tailoring tasks, game rules, and features to match each user's preferences or skill level [90]. Personalized gamification can be implemented through (1) customization, where users can select the elements they wish to use, and (2) automatic adaptation, in which the system selects the game design elements for each user, potentially with some user input. Therefore, developers and researchers are encouraged to consider integrating personalized gamification elements into their apps to improve user engagement and, consequently, the effectiveness of mobile apps on health-related outcomes.

Finally, it is noteworthy that most (11/15, 73%) identified studies with an intervention did not have comparison groups, which makes it challenging to draw definitive conclusions regarding the effectiveness of various mobile apps.

Gamification Strategies and Tactics

Following the gamification framework proposed by Cugelman [12], the most common gamification strategies were fun and playfulness, feedback on performance, reinforcement, and goal setting, whereas social connectivity was the least commonly used strategy, followed by comparing progress. Moreover, feedback, points, story or theme, and rewards were the most common on-screen features, while showing game leaders (leaderboards), badges, and showing progress were the least common elements applied to mobile apps.

The fun and playfulness strategy (23/32, 72%) was the gamification principle most applied to mobile apps. There was frequent use of on-screen tactics such as stories, themes, avatars, graphic representations of information, fun videos, and audio effects. Incorporating these features requires a significant amount of computational resources, time, and knowledge [91]. Most (23/32, 72%) of the mobile apps in this review had comparable playfulness elements to enhance children's experience of fun and motivate them to use the app on a regular basis. Previous studies showed that playful and fun experiences in mobile apps will increase positive attitudes toward mobile apps when the users can engage in pleasurable experiences [92,93].

Feedback on performance and reinforcement were among the most frequently used gamification strategies in mobile apps. This finding aligns with the results of previous research identifying successful behavior-change techniques in gamified mobile apps [56]. Previous reviews on gamification in other populations found that 94% and 81% of health apps had incorporated feedback on performance and reinforcement, respectively, and achievement- and progress-oriented elements such as in-app rewards [12,94]. In our scoping review, the most prevalent types of reinforcement were points (14/32, 44%) and tangible rewards (13/32, 41%), aligning with previous reviews, and indicating a positive direction toward promoting health behavior change through these strategies. This finding also emphasizes that easy-to-implement game features, such as points and feedback through messaging are the most widely used gamification features to promote engagement and motivation [94]. Nevertheless, the outcomes were frequently measured only through in-app behavior (eg, completing tasks in the app for rewards).

We found that achievement- and progress-oriented rewards were given to users as a result of their change in specific behaviors, such as completing cognitive assessments [65] or participating in daily language test challenges [66]. For instance, when children with ASD used the TOBY iPad app, an early intervention tool, they had to choose a specific picture from a set of pictures. Upon completing the task, they would gain tokens (points), which could be used to choose a reward [47-49].

Another app, LexiPal, an educational app for children with dyslexia, used various game elements, such as points, feedback, and rewards. Upon successfully completing 1 round of tasks, a pop-up window would appear to illustrate the score and reward. If the child gets a score of 4 to 5, they earn a golden cup reward and receive text and audio feedback [51]. Users were rarely rewarded for behavior changes external to the app. For instance, in the Urakami app, users who completed outpatient therapy sessions could collect points, which could be exchanged to purchase in-app avatars [54].

A growing body of literature has criticized the paucity of use of incentives through points, badges, and leaderboard elements in digital health solutions [95]. Points were among the most frequently applied game elements in our review, whereas badges and leaderboards were incorporated by a minority of apps. Many studies have investigated the effectiveness of a combination of points, leaderboards, and badges to increase outcomes such as engagement in physical activity. Several studies have shown that the application of these game features could significantly enhance individuals' physical activity [20]. In contrast, several studies found that points and leaderboards did not significantly impact walking compared to a nongamified version of the same app [96,97]. In addition, Maher et al [98] found that a combination of rewards and leaderboards led to a short-term increase in physical activity but there was no long-term positive impact on health behavior. Further research is needed to investigate the impact of using game features for short- and long-term impacts in the childhood disability field.

Goal setting is a known intervention strategy for successful health behavior change [99]. In our review, 44% (14/32) of the mobile apps used goal setting to promote user engagement. Previous research has outlined that combining goal setting with showing progress, feedback, and rewards can significantly enhance intrinsic motivation toward behavior [100]. Although feedback and rewards have been extensively applied to the identified apps in our review, comparing progress elements has been underused. There is a vast amount of literature on the possible benefits of rewards and feedback, yet each element's effectiveness still needs to be determined [101]. Another concern is that these features may enhance extrinsic motivation rather than intrinsic motivation, which leads to the weak maintenance impact of gamified apps [102]. Therefore, we recommend further investigating the independent effects of individual mobile-based gamification elements on children with disabilities.

Despite the potential advantages of social connectivity on young people's well-being [103], only 4 (12%) of the 32 apps implemented a social connectivity strategy. They provided the users with access to chat rooms [61], the ability to share the points with parents [60], and the ability to send messages to parents and health professionals [54,79]. The scarcity of social connectivity options found in our review contrasts with previous research indicating that social networks could positively impact health behavior change insofar as app users can interact with other users and share their points and experiences with one another. [103,104]. A recent review identified several social support features where app users could interact with others through sharing posts and sending private messages [105]. However, previous studies highlighted the potential negative aspects of social connectivity in mobile apps. For example, concerns were raised about inappropriate content sharing and messaging between children and information inaccuracy in the technology space [106,107]. Therefore, while much research on the potential effectiveness of social connectivity has been carried out, some critical issues need to be investigated in the childhood disability field. Moreover, no apps provide leaderboards. This may suggest a deliberate decision in the design, as children with disabilities are generally a particularly susceptible group who may experience increased levels of stress when comparing themselves to their peers [108,109]. A meta-analysis of qualitative studies of students shows that gamification can cause anxiety and jealousy among students [89]. For example, Johnson et al [81] did not incorporate some traditional gamification elements, such as badges and leaderboards, considering the needs of children with neurodevelopmental disabilities in their study. Therefore, researchers in this field should investigate the specific needs and potential stressors of children with disabilities when considering the incorporation of gamification elements such as leaderboards.

Level of Gamification and Reasons to Apply Gamification

This review highlights that most (18/32, 56%) identified apps implement a medium level of gamification strategies and a low level of gamification tactics, with few (1/32, 3%) adopting a high level of gamification strategies or tactics. The Guess What app used the greatest number of gamification tactics (6 of 7). All identified apps in our review used at least 1 gamification strategy, which supports previous research findings that gamification is meant to significantly improve psychological outcomes [110,111]. While gamification tactics, also known as on-screen features, are considered to be part of persuasive app design to promote engagement and motivation [12], most identified apps implemented a low (23/32, 72%) number of gamification tactics. Children with disabilities are an underserved group that faces numerous barriers to accessing health services [112]. Designing digital solutions for children with disabilities requires collaboration among childhood disability researchers, mHealth experts, and children and their families. The provision of these solutions such as gamified mobile apps for children with disabilities has the potential to reduce health inequities.

However, the level of incorporated gamification tactics should be interpreted with caution in our review. One vital theoretical issue is that multiple gamification frameworks have different definitions of gamification and categorizations of game elements. For example, Lister et al [11] used "gamification" to define levels, rewards, prizes, and competitions but not avatars; meanwhile, Johnson et al [110] used "gamification" to describe all these game elements. Similarly, multiple studies separate feedback and rewards [11,113], whereas Sardi et al [9] counted them as 1 game mechanic. Although we used the framework proposed by Cugelman [12] to define gamification in our study, it would be difficult to make a definite conclusion regarding the level of gamification, as the number of game elements varies in different frameworks. Therefore, there is a need to have a solid framework for mobile-based gamification for childhood disabilities.

Furthermore, researchers applied gamification to apps for various purposes. We found justification for applying particular game elements in 16 studies. Of the 16 studies, 9 (56%) used gamification elements to promote engagement and motivation. Gamification aims to include playful elements to transform a typically boring activity into one that is enjoyable and engaging [9]. For example, FarmApp, which is a mobile app used to assess cognitive skills among children with neurodevelopmental disabilities, incorporated interactive game-like elements to be more motivating and enjoyable for children to complete the assessments [65]. In addition, the SMART app, which is for the self-management of children with traumatic brain injury, was redesigned and implemented gamified components to encourage youth with mild traumatic brain injury to use the app daily and manage their symptoms [52]. These findings align with the purpose of gamification as a tool to increase engagement and motivation [10,12].

In contrast, gamification was also applied to increase the impact of the intervention. Many mobile apps included game elements to increase the efficacy of the intervention. For instance, gamification was used in the ColourSpot app to encourage users to complete the intervention [76]. In the Dyslexia Baca app, visual graphics were incorporated to assist children with dyslexia in understanding the intervention instructions [66]. In another example, Johnson et al [81] used various gamification tactics, such as avatars, weekly progress monitoring, and earning stars, to engage children in their therapy prescription app. These justifications align with the previous literature showing that engaging apps, such as gamified apps, can enhance the effectiveness of interventions by encouraging users to use them consistently and frequently [114]. Although the capacity of gamification to promote engagement and motivation has been extensively studied [10,12,110], more research is needed to confirm the ability of gamification to increase intervention efficacy.

Limitations and Recommendations

Although this scoping review was guided by the PRISMA-ScR framework [42], it has some limitations. First, the primary aim of our review was to summarize a record of all gamified apps for children with disabilities from 2008 to 2023; however, we did not assemble any information on the effectiveness of gamified apps on any evaluation metric. By including broad search keywords in the search strategy, we had a high volume of document titles and abstracts to screen. Nevertheless, this enhances the risk of accidental exclusion of relevant citations. To minimize this, before both abstract and full-text screening, we performed pilot testing on a random sample of documents, and any discrepancies were resolved by KS. This ensured that the title and abstract screening was appropriate before the full-text screening.

In addition, although our scoping review was inclusive (no restrictions on study design), we excluded studies of children with mental concerns (eg, anxiety and depression) or other health issues (eg, obesity and cancer). Given the unique characteristics of mental health problems in children and adolescents, we recommend an independent review of gamified mobile apps for children with mental health issues.

Furthermore, unpublished studies were not included in this scoping review. Because many mHealth apps are privately designed, development or evaluative information for these apps is not available in the public domain, which may result in a substantial knowledge gap. Although private companies have been increasingly transparent in publishing data in recent years [115], this knowledge gap cannot be addressed in this scoping review. Therefore, the results of this review are not generalizable to commercial apps for children with disabilities.

Although the consultation exercise is a vital yet optional component of the scoping review framework proposed by Arksey and O'Malley [41], it was not conducted in this review. Specifically, we are conducting a separate project to seek out stakeholder input to further inform this area of research. The primary author (EM) used the results of this scoping review to inform the interview guide of a qualitative project where different stakeholders, including children and youth with disabilities, parents or caregivers, clinicians, and representatives of community organizations, shared their perspectives about different gamification elements. The findings from this research will not only enhance the results of our scoping review but also make an important contribution to the deeper understanding of best practices in developing gamified mobile apps for children with disabilities.

Scoping reviews aim to rigorously survey the current body of literature and identify crucial concepts, types of evidence, and knowledge gaps. Typically, they are not structured to evaluate the effectiveness of the interventions. In accordance with this methodology, our review did not assess the effectiveness of the identified gamified mobile apps; however, we reported a summary of the mobile app's impact on child outcomes. Indeed, systemically evaluating the effectiveness would be challenging, given the considerable heterogeneity in the types of disabilities, mHealth strategies used, and the wide range of outcomes applied. Recognizing this limitation, there is a need for future research to evaluate the effectiveness of these gamified mobile apps on specific populations and outcomes.

Finally, limited evidence was provided in this review on the extent to which health behaviors "outside the app" were augmented in children with disabilities, and its association with the gamification features proposed. Johnson et al [110] conducted a review to assess the impact of gamified interventions on health and well-being in a broader population indicating that gamification could have a positive impact on healthy behaviors (eg, physical activity). Future research should investigate the association of gamification features in mobile apps with subsequent healthy behaviors "outside the app" for children with disabilities.

Conclusions

This review provides a summary of the current literature on mobile-based gamification used for children with disabilities reported after 2008. A total of 6 databases were comprehensively searched, and 38 studies with 32 unique apps were identified that focused predominantly on treatment goals and were in most cases used in children with ASD. This review demonstrates that gamified mobile apps for children with ASD are mainly designed to enhance communication, social skills, and oral health knowledge. In addition, several mobile apps address self-management in various conditions, academic achievements in learning disabilities, and psychological outcomes such as motivation and engagement, demonstrating their potential in improving diverse health-related outcomes in children with disabilities. The results of this study showed that gamification could provide potential benefits across different populations and conditions; however, there were mixed results regarding its impact and benefits. These results can guide other researchers in the childhood disability field in recognizing disabilities or behavioral outcomes that have been neglected, thus informing future mobile app development and research on those disabilities. Collectively, this information will enable the researchers in this field to understand how gamification can improve intervention effects on relevant outcomes and meet the specific needs of this population.

Acknowledgments

The authors would like to express their gratitude to Ms Jill Boruff, an associate librarian at McGill University, for her valuable support in developing the search strategy.

This project was supported by the CHILD-BRIGHT Strategic Patient Oriented Research Network- Canadian Institutes of Health Research and the TELUS Friendly Future Foundation.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

All authors were involved in the design and development of the study protocol, which was undertaken as part of a PhD for EM under the supervision of KS. EM and RC contributed to the database search. EM, PYY, and AC contributed to abstract and full-text screening and data extraction. EM wrote the first draft of the paper, and all authors agreed to be accountable for all aspects of the research project and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 498 KB - games_v12i1e49029_app1.pdf]

Multimedia Appendix 2
Search strategy.
[DOCX File , 21 KB - games_v12i1e49029_app2.docx]

Multimedia Appendix 3
Summary descriptions of studies included in the scoping review.
[DOCX File , 69 KB - games_v12i1e49029_app3.docx]

Multimedia Appendix 4
Reasons for applying gamification found in scoping review.
[DOCX File , 24 KB - games_v12i1e49029_app4.docx]

Multimedia Appendix 5
Included studies with results directly related to an intervention (n=15).
[DOCX File , 28 KB - games_v12i1e49029_app5.docx]

References

1. World report on disability 2011. World Health Organization. URL: https://www.who.int/teams/noncommunicable-diseases/sensory-functions-disability-and-rehabilitation/world-report-on-disability [accessed 2024-08-06]
2. Davis K, Carter S, Myers E, Rocca N. Health promotion for young people with profound and multiple learning disabilities. Nurs Child Young People 2018 Feb 07;30(1):28-34. [doi: 10.7748/ncyp.2018.e1000] [Medline: 29412537]

3. Bhattacharyya O, Reeves S, Zwarenstein M. What is implementation research?: rationale, concepts, and practices. Res Soc Work Pract 2009 May 27;19(5):491-502. [doi: 10.1177/1049731509335528]
4. Quelly SB, Norris AE, DiPietro JL. Impact of mobile apps to combat obesity in children and adolescents: a systematic literature review. J Spec Pediatr Nurs 2015 Oct 22;21(1):5-17. [doi: 10.1111/jspn.12134] [Medline: 26494019]
5. Rodríguez Mariblanca M, Cano de la Cuerda R. Mobile applications in children with cerebral palsy. Neurología (English Edition) 2021 Mar;36(2):135-148. [doi: 10.1016/j.nrleng.2019.07.002] [Medline: 29275969]
6. Sugie NF. Utilizing smartphones to study disadvantaged and hard-to-reach groups. Sociol Methods Res 2016 Jan 18;47(3):458-491. [doi: 10.1177/0049124115626176]
7. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2013 Aug 05;21(5):427-436. [doi: 10.1002/cpp.1855] [Medline: 23918764]
8. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015 May 19;6:27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
9. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
10. King D, Greaves F, Exeter C, Darzi A. 'Gamification': influencing health behaviours with games. J R Soc Med 2013 Mar;106(3):76-78 [FREE Full text] [doi: 10.1177/0141076813480996] [Medline: 23481424]
11. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014 Aug 04;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
12. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
13. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011 Presented at: MindTrek '11; September 28-30, 2011; Tampere, Finland. [doi: 10.1145/2181037.2181040]
14. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. Using game-design elements in non-gaming contexts. In: Proceedings of the CHI '11 Extended Abstracts on Human Factors in Computing Systems. 2011 Presented at: CHI EA '11; May 7-12, 2011; Vancouver, BC URL: https://dl.acm.org/doi/10.1145/1979742.1979575 [doi: 10.1145/2181037.2181040]
15. Marczewski A. What's the difference between Gamification and Serious Games? Game Developer. 2013 Mar 11. URL: https://www.gamedeveloper.com/business/what-s-the-difference-between-gamification-and-serious-games- [accessed 2024-08-07]
16. Ritterfeld U, Vorderer P, Cody M. Serious Games: Mechanisms and Effects. Milton Park, UK: Taylor & Francis; 2009.
17. Xu L, Shi H, Shen M, Ni Y, Zhang X, Pang Y, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. JMIR Mhealth Uhealth 2022 Feb 03;10(2):e27794 [FREE Full text] [doi: 10.2196/27794] [Medline: 35113034]
18. Sola D, Couturier J, Voyer B. Unlocking patient activation in chronic disease care. Br J Healthcare Manag 2015 May 11;21(5):220-225. [doi: 10.12968/bjhc.2015.21.5.220]
19. Pimentel-Ponce M, Romero-Galisteo RP, Palomo-Carrión R, Pinero-Pinto E, Merchán-Baeza JA, Ruiz-Muñoz M, et al. Gamification and neurological motor rehabilitation in children and adolescents: a systematic review. Neurologia (Engl Ed) 2021 Apr 15:S0213-4853(21)00049-9 [FREE Full text] [doi: 10.1016/j.nrl.2021.02.011] [Medline: 33867183]
20. Allam A, Kostova Z, Nakamoto K, Schulz PJ. The effect of social support features and gamification on a web-based intervention for rheumatoid arthritis patients: randomized controlled trial. J Med Internet Res 2015 Jan 09;17(1):e14 [FREE Full text] [doi: 10.2196/jmir.3510] [Medline: 25574939]
21. AlMarshedi A, Wills GB, Ranchhod A. The wheel of Sukr: a framework for gamifying diabetes self-management in Saudi Arabia. Procedia Comput Sci 2015;63:475-480. [doi: 10.1016/j.procs.2015.08.370]
22. Dennis TA, O'Toole L. Mental health on the go: effects of a gamified attention bias modification mobile application in trait anxious adults. Clin Psychol Sci 2014 Sep 01;2(5):576-590 [FREE Full text] [doi: 10.1177/2167702614522228] [Medline: 26029490]
23. Pereira P, Duarte E, Rebelo F, Noriega P. A review of gamification for health-related contexts. In: Proceedings of the Design, User Experience, and Usability: User Experience Design Practice Third International Conference. 2014 Presented at: DUXU 2014; June 22-27, 2014; Crete, Greece. [doi: 10.1007/978-3-319-07626-3_70]
24. Ahtinen A, Huuskonen P, Häkkilä J. Let's all get up and walk to the North Pole: design and evaluation of a mobile wellness application. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. 2010 Presented at: NordiCHI '10; October 16-20, 2010; Reykjavik, Iceland. [doi: 10.1145/1868914.1868920]
25. Chen Y, Pu P. HealthyTogether: exploring social incentives for mobile fitness applications. In: Proceedings of the Second International Symposium of Chinese CHI. 2014 Presented at: Chinese CHI '14; April 26-27, 2014; Toronto, ON. [doi: 10.1145/2592235.2592240]
26. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012 May 08;14(3):e70. [doi: 10.2196/jmir.2058] [Medline: 22564332]

27. Ferreira C, Guimarães V, Santos A, Sousa I. Gamification of stroke rehabilitation exercises using a smartphone. In: Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare. 2014 Presented at: PervasiveHealth '14; May 20-23, 2014; Oldenburg, Germany. [doi: https://doi.org/10.4108/icst.pervasivehealth.2014.255326]
28. Helf C, Hlavacs H. Apps for life change: critical review and solution directions. Entertain Comput 2016 May;14:17-22. [doi: 10.1016/j.entcom.2015.07.001]
29. Hu R, Fico G, Cancela J, Teresa Arredondo M. Gamification system to support family-based behavioral interventions for childhood obesity. In: Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics. 2014 Presented at: BHI 2014; Valencia, Spain; June 1-4, 2014. [doi: 10.1109/bhi.2014.6864334]
30. Zhao Z, Etemad SA, Arya A. Gamification of exercise and fitness using wearable activity trackers. In: Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS). 2015 Presented at: ISCSS 2015; Loughborough, UK; September 9-11, 2015. [doi: 10.1007/978-3-319-24560-7_30]
31. Ryan RM, Patrick H, Deci EL, Williams GC. Facilitating health behaviour change and its maintenance: interventions based on Self-Determination Theory. Eur Health Psychol 2008;10:2-5 [FREE Full text]
32. Dahlstrøm C. Impacts of gamification on intrinsic motivation. Norwegian University of Science and Technology. 2017. URL: https://www.ntnu.edu/documents/139799/1279149990/04+Article+Final_camildah_fors%C3%B8k_2017-12-06-13-53-55_TPD4505.Camilla.Dahlstr%C3%B8m.pdf [accessed 2024-08-07]
33. Deterding S. Situated motivational affordances of game elements: a conceptual model. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC.
34. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol Bull 1999;125(6):627-668. [doi: 10.1037//0033-2909.125.6.627]
35. Aparicio AF, Luis Gutiérrez Vela F, González-Sánchez JL, Isla-Montes JL. Analysis and application of gamification. In: Proceedings of the 13th International Conference on Interacción Persona-Ordenador. 2012 Presented at: INTERACCION '12; October 3-5, 2012; Elche, Spain. [doi: 10.1145/2379636.2379653]
36. Gamification, huh? What IS it good for? The Guardian. URL: https://www.theguardian.com/technology/appsblog/2012/nov/28/gamification-what-is-it-good-for [accessed 2024-08-07]
37. Koivisto J, Hamari J. Demographic differences in perceived benefits from gamification. Comput Hum Behav 2014 Jun;35:179-188. [doi: 10.1016/j.chb.2014.03.007]
38. Garde A, Umedaly A, Abulnaga SM, Robertson L, Junker A, Chanoine JP, et al. Assessment of a mobile game ("MobileKids Monster Manor") to promote physical activity among children. Games Health J 2015 Apr;4(2):149-158. [doi: 10.1089/g4h.2014.0095] [Medline: 26181809]
39. Kreuter MW, Farrell DW, Brennan LK, Olevitch LR. Tailoring Health Messages: Customizing Communication With Computer Technology. Milton Park, UK: Routledge; 1999.
40. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
41. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616] [Medline: 20854677]
42. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
43. Journal of Medical Internet Research homepage. Journal of Medical Internet Research. URL: https://www.jmir.org/ [accessed 2024-08-07]
44. Rayyan homepage. Rayyan. URL: https://www.rayyan.ai/ [accessed 2024-08-07]
45. Rajani NB, Weth D, Mastellos N, Filippidis FT. Use of gamification strategies and tactics in mobile applications for smoking cessation: a review of the UK mobile app market. BMJ Open 2019 Jun 17;9(6):e027883 [FREE Full text] [doi: 10.1136/bmjopen-2018-027883] [Medline: 31213452]
46. Kucirkova N, Messer D, Critten V, Harwood J. Story-making on the iPad when children have complex needs: two case studies. Commun Disord Q 2014 Mar 06;36(1):44-54. [doi: 10.1177/1525740114525226]
47. Moore DW, Venkatesh S, Anderson A, Greenhill S, Phung D, Duong T, et al. TOBY play-pad application to teach children with ASD - a pilot trial. Dev Neurorehabil 2015;18(4):213-217. [doi: 10.3109/17518423.2013.784817] [Medline: 23869435]
48. Parsons D, Vaz S, Lee H, Robinson C, Cordier R. A twelve-month follow-up of an information communication technology delivered intervention for children with autism spectrum disorder living in regional Australia. Res Dev Disabil 2020 Nov;106:103743. [doi: 10.1016/j.ridd.2020.103743] [Medline: 32835895]
49. Parsons D, Wilson NJ, Vaz S, Lee H, Cordier R. Appropriateness of the TOBY application, an iPad intervention for children with autism spectrum disorder: a thematic approach. J Autism Dev Disord 2019 Oct;49(10):4053-4066. [doi: 10.1007/s10803-019-04115-9] [Medline: 31289988]

50. Penev Y, Dunlap K, Husic A, Hou C, Washington P, Leblanc E, et al. A mobile game platform for improving social communication in children with autism: a feasibility study. Appl Clin Inform 2021 Oct;12(5):1030-1040 [FREE Full text] [doi: 10.1055/s-0041-1736626] [Medline: 34788890]
51. Utama Saputra MR. LexiPal: design, implementation and evaluation of gamification on learning application for dyslexia. Int J Comput Appl 2015 Dec;131(7):37-43. [doi: 10.5120/ijca2015907416]
52. Schmidt M, Fisher AP, Sensenbaugh J, Ling B, Rietta C, Babcock L, et al. User experience (re)design and evaluation of a self-guided, mobile health app for adolescents with mild Traumatic Brain Injury. J Form Des Learn 2020 Dec;4(2):51-64 [FREE Full text] [doi: 10.1007/s41686-019-00038-x] [Medline: 33860150]
53. Thida A, Han NN, Thawtar Oo S, Li S, Ding C. VOIS: the first speech therapy app specifically designed for Myanmar hearing-impaired children. In: Proceedings of the 23rd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques. 2020 Presented at: O-COCOSDA 2020; November 5-7, 2020; Yangon, Myanmar. [doi: 10.1109/o-cocosda50338.2020.9295024]
54. Urakami T. Effectiveness of a smartphone application on medication adherence in children with short stature receiving GH therapy: a multicenter prospective cohort study (GTL-App). Clin Pediatr Endocrinol 2021;30(2):85-92 [FREE Full text] [doi: 10.1297/cpe.30.85] [Medline: 33867668]
55. Ying KT, Sah SB, Abdullah MH. Personalised avatar on social stories and digital storytelling: fostering positive behavioural skills for children with autism spectrum disorder. In: Proceedings of the 4th International Conference on User Science and Engineering. 2016 Presented at: i-USEr 2016; August 23-25, 2016; Melaka, Malaysia. [doi: 10.1109/iuser.2016.7857970]
56. Chua L, Goh J, Nay ZT, Huang L, Cai Y, Seah R. ICT-enabled emotional learning for special needs education. In: Cai Y, Goei S, Trooster W, editors. Simulation and Serious Games for Education. Singapore, Singapore: Springer; 2017.
57. Doenyas C, Şimdi E, Özcan E, Çataltepe Z, Birkan B. Autism and tablet computers in Turkey: teaching picture sequencing skills via a web-based iPad application. Int J Child Comput Interact 2014 Jan;2(1):60-71. [doi: 10.1016/j.ijcci.2014.04.002]
58. Holmes JM, Manh VM, Lazar EL, Beck RW, Birch EE, Kraker RT, et al. Effect of a binocular iPad game vs part-time patching in children aged 5 to 12 years with amblyopia: a randomized clinical trial. JAMA Ophthalmol 2016 Dec 01;134(12):1391-1400 [FREE Full text] [doi: 10.1001/jamaophthalmol.2016.4262] [Medline: 27812703]
59. Kelly KR, Jost RM, Dao L, Beauchamp CL, Leffler JN, Birch EE. Binocular iPad game vs patching for treatment of amblyopia in children: a randomized clinical trial. JAMA Ophthalmol 2016 Dec 01;134(12):1402-1408 [FREE Full text] [doi: 10.1001/jamaophthalmol.2016.4224] [Medline: 27832248]
60. Aburukba R, Aloul F, Mahmoud A, Kamili K, Ajmal S. AutiAid: a learning mobile application for autistic children. In: Proceedings of the IEEE 19th International Conference on e-Health Networking, Applications and Services. 2017 Presented at: Healthcom 2017; October 12-15, 2017; Dalian, China. [doi: 10.1109/healthcom.2017.8210788]
61. Alnaghaimshi NI, Alhazmi A, Alqanwah SA, Aldablan MS, Almossa MA. Autismworld: an Arabic application for autism spectrum disorder. In: Proceedings of the 3rd International Conference on Computer Applications & Information Security. 2020 Presented at: ICCAIS 2020; March 19-21, 2020; Riyadh, Saudi Arabia. [doi: 10.1109/iccais48893.2020.9096811]
62. Barta EA, Guzsvinecz T, Sik Lanyi C, Szucs V. Android-based daily routine organizing application for elementary school students living with ASD. Stud Health Technol Inform 2017;242:283-290. [Medline: 28873812]
63. Birtwell KB, Platner AK, Nowinski LA. Exploring the use of sidekicks! For children with autism spectrum disorder (ASD). Psychol Serv 2019 May;16(2):266-270. [doi: 10.1037/ser0000301] [Medline: 30421945]
64. Borhan NH, Shiang CW, Chiu PC, Sharbini H, Ping TP, Othman RM, et al. An enhancement of dyslexic mobile application using sight word reading strategy: results and findings. J Comput Sci 2018;14(7):919-929. [doi: 10.3844/jcssp.2018.919.929]
65. Brkić D, Ng-Cordell E, O'Brien S, Martin J, Scerif G, Astle D, et al. [Formula: see text]FarmApp: a new assessment of cognitive control and memory for children and young people with neurodevelopmental difficulties. Child Neuropsychol 2022 Nov;28(8):1097-1115. [doi: 10.1080/09297049.2022.2054968] [Medline: 35332845]
66. Daud SM, Abas H. 'Dyslexia Baca' mobile app -- the learning ecosystem for dyslexic children. In: Proceedings of the International Conference on Advanced Computer Science Applications and Technologies. 2013 Presented at: ACSAT 2013; December 23-24, 2013; Kuching, Malaysia. [doi: 10.1109/acsat.2013.87]
67. Dehkordi SR, Rias RM. Using mobile game application to teach children with Autism Spectrum Disorder (ASD) multiple cues responding: a pilot study. In: Proceedings of the 3rd International Conference on User Science and Engineering. 2014 Presented at: i-USEr 2014; September 2-5, 2014; Shah Alam, Malaysia. [doi: 10.1109/iuser.2014.7002705]
68. Gómez L, Carro RM. Adaptive training of children with attention deficit hyperactivity disorder through multi-touch surfaces. In: Proceedings of the IEEE 14th International Conference on Advanced Learning Technologies. 2014 Presented at: ICALT 2014; July 7-10, 2014; Athens, Greece. [doi: 10.1109/icalt.2014.164]
69. Guzsvinecz T, Koszegi-Vigh D, Szucs V, Sik Lanyi C. "Sliders" android game - improving logical skills of people with disabilities. Stud Health Technol Inform 2017;242:279-282. [Medline: 28873811]
70. Sandhu H, Wilson K, Reed N, Mihailidis A. A mobile phone app for the self-management of pediatric concussion: development and usability testing. JMIR Hum Factors 2019 May 31;6(2):e12135 [FREE Full text] [doi: 10.2196/12135] [Medline: 31152527]

71. Irwin J, Preston J, Brancazio L, D'angelo M, Turcios J. Development of an audiovisual speech perception app for children with autism spectrum disorders. Clin Linguist Phon 2015 Jan;29(1):76-83 [FREE Full text] [doi: 10.3109/02699206.2014.966395] [Medline: 25313714]
72. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess what?: towards understanding autism from structured video using facial affect. J Healthc Inform Res 2019;3(1):43-66 [FREE Full text] [doi: 10.1007/s41666-018-0034-9] [Medline: 33313475]
73. Macdonald D, Luk G, Quintin EM. Early reading comprehension intervention for preschoolers with autism spectrum disorder and hyperlexia. J Autism Dev Disord 2022 Apr;52(4):1652-1672. [doi: 10.1007/s10803-021-05057-x] [Medline: 34046778]
74. Manh VM, Holmes JM, Lazar EL, Kraker RT, Wallace DK, Kulp MT, et al. A randomized trial of a binocular iPad game versus part-time patching in children aged 13 to 16 years with amblyopia. Am J Ophthalmol 2018 Feb;186:104-115 [FREE Full text] [doi: 10.1016/j.ajo.2017.11.017] [Medline: 29196184]
75. Mwamba HM, Fourie PR, van den Heever D. PANDAS: paediatric attention-deficit/hyperactivity disorder application software. Appl Sci 2019 Apr 20;9(8):1645. [doi: 10.3390/app9081645] [Medline: 31946165]
76. Tang T, Álvaro L, Alvarez J, Maule J, Skelton A, Franklin A, et al. ColourSpot, a novel gamified tablet-based test for accurate diagnosis of color vision deficiency in young children. Behav Res Methods 2022 Jun;54(3):1148-1160 [FREE Full text] [doi: 10.3758/s13428-021-01622-5] [Medline: 34463952]
77. Cahyono D. Gamification for education: using LexiPal to foster intrinsic and extrinsic learning motivation of students with dyslexia. In: Proceedings of the 14th International Conference on Education Technology and Computers. 2022 Presented at: ICETC '22; October 28-30, 2022; Barcelona, Spain. [doi: 10.1145/3572549.3572555]
78. Chistol M, Turcu C, Danubianu M. Autism assistant: a platform for autism home-based therapeutic intervention. IEEE Access 2023 Aug 30;11:94188-94204. [doi: 10.1109/access.2023.3310397]
79. Tan BL, Nair R, Duggal MS, Aishworiya R, Tong HJ. Development of oral health resources and a mobile app for caregivers and autistic children through consensus building. Autism 2024 Apr;28(4):959-974. [doi: 10.1177/13623613231188768] [Medline: 37537880]
80. Schmidt M, Babcock L, Kurowski BG, Cassedy A, Sidol C, Wade SL. Usage patterns of an mHealth symptom monitoring app among adolescents with acute mild traumatic brain injuries. J Head Trauma Rehabil 2022;37(3):134-143 [FREE Full text] [doi: 10.1097/HTR.0000000000000768] [Medline: 35125434]
81. Johnson RW, White BK, Gucciardi DF, Gibson N, Williams SA. Intervention mapping of a gamified therapy prescription app for children with disabilities: user-centered design approach. JMIR Pediatr Parent 2022 Aug 09;5(3):e34588. [doi: 10.2196/34588] [Medline: 35943782]
82. Johnson RW, White BK, Gibson N, Gucciardi DF, Williams SA. A mixed-methods feasibility study of a gamified therapy prescription app for children with neurodisability. Phys Occup Ther Pediatr 2024;44(4):586-603. [doi: 10.1080/01942638.2023.2263576] [Medline: 37814984]
83. Krishnan L, Iyer K, Kumar PD. Effectiveness of two sensory-based health education methods on oral hygiene of adolescent with autism spectrum disorders: an interventional study. Spec Care Dentist 2021 Sep;41(5):626-633. [doi: 10.1111/scd.12606] [Medline: 34050975]
84. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013.
85. Lindsay S, Kingsnorth S, Mcdougall C, Keating H. A systematic review of self-management interventions for children and youth with physical disabilities. Disabil Rehabil 2014;36(4):276-288 [FREE Full text] [doi: 10.3109/09638288.2013.785605] [Medline: 23614359]
86. Sandjojo J, Eltringham EG, Gebhardt WA, Zedlitz AM, Embregts PJ, Evers AW. Self-management interventions for people with intellectual disabilities: a systematic review. Patient Educ Couns 2020 Oct;103(10):1983-1996. [doi: 10.1016/j.pec.2020.06.009] [Medline: 32576421]
87. Ares EM, Iglesias P, Gómez SL, Torres RM. Neurodevelopmental difficulties as a comprehensive construct of learning disabilities in children with developmental delay: a systematic review. Ann Psychol 2020 Apr;36(2):271-282. [doi: 10.6018/analesps.36.2.347741]
88. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Hum Comput Interact 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471]
89. Bai S, Hew KF, Huang B. Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. Educ Res Rev 2020 Jun;30:100322. [doi: 10.1016/j.edurev.2020.100322]
90. Tondello GF, Nacke LE. Validation of user preferences and effects of personalized gamification on task performance. Front Comput Sci 2020 Aug 6;2. [doi: 10.3389/fcomp.2020.00029]
91. Quinn AJ, Bederson BB, Bonsignore EM, Druin A. StoryKit: designing a mobile application for story creation by children and older adults. Human Computer Interaction Lab, University of Maryland. 2009. URL: https://www.researchgate.net/publication/252148516_StoryKit_Designing_a_Mobile_Application_for_Story_Creation_By_Children_And_Older_Adults [accessed 2024-08-16]

92. Chou CH, Chiu CH, Ho CY, Lee JC. Understanding mobile apps continuance usage behavior and habit: an expectance-confirmation theory. In: Proceedings of the Pacific Asia Conference on Information Systems 2013. 2013 Presented at: PACIS 2013; June 18-22, 2013; Jeju Island, Korea.
93. Maghnati F, Ling K. Exploring the relationship between experiential value and usage attitude towards mobile apps among the smartphone users. Int J Bus Manage 2013;8(4). [doi: 10.5539/ijbm.v8n4p1]
94. Steiner B, Elgert L, Saalfeld B, Wolf KH. Gamification in rehabilitation of patients with musculoskeletal diseases of the shoulder: scoping review. JMIR Serious Games 2020 Aug 25;8(3):e19914 [FREE Full text] [doi: 10.2196/19914] [Medline: 32840488]
95. Hunter D, Werbach K. For the Win: How Game Thinking Can Revolutionize Your Business. Philadelphia, PA: Wharton Digital Press; 2012.
96. Ahn SJ, Johnsen K, Ball C. Points-based reward systems in gamification impact children's physical activity strategies and psychological needs. Health Educ Behav 2019 Jun;46(3):417-425 [FREE Full text] [doi: 10.1177/1090198118818241] [Medline: 30678507]
97. Zuckerman O, Gal-Oz A. Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. Pers Ubiquit Comput 2014;18:1705-1719. [doi: 10.1007/s00779-014-0783-2]
98. Maher C, Ferguson M, Vandelanotte C, Plotnikoff R, De Bourdeaudhuij I, Thomas S, et al. A web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: randomized controlled trial. J Med Internet Res 2015 Jul 13;17(7):e174 [FREE Full text] [doi: 10.2196/jmir.4086] [Medline: 26169067]
99. Epton T, Currie S, Armitage CJ. Unique effects of setting goals on behavior change: systematic review and meta-analysis. J Consult Clin Psychol 2017 Dec;85(12):1182-1198. [doi: 10.1037/ccp0000260] [Medline: 29189034]
100. Hicks K, Gerling K, Richardson G, Pike T, Burman O, Dickinson P. Understanding the effects of gamification and juiciness on players. In: Proceedings of the IEEE Conference on Games (CoG). 2019 Presented at: CoG 2019; August 20-23, 2019; London, UK. [doi: 10.1109/cig.2019.8848105]
101. Lewis ZH, Swartz MC, Lyons EJ. What's the point?: a review of reward systems implemented in gamification interventions. Games Health J 2016 Apr;5(2):93-99. [doi: 10.1089/g4h.2015.0078] [Medline: 26812253]
102. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. Contemp Educ Psychol 2020 Apr;61:101860. [doi: 10.1016/j.cedpsych.2020.101860]
103. Martin AJ, Dowson M. Interpersonal relationships, motivation, engagement, and achievement: yields for theory, current issues, and educational practice. Rev Educ Res 2009 Mar 01;79(1):327-365. [doi: 10.3102/0034654308325583]
104. van Achterberg T, Huisman-de Waal GG, Ketelaar NA, Oostendorp RA, Jacobs JE, Wollersheim HC. How to promote healthy behaviours in patients? An overview of evidence for behaviour change techniques. Health Promot Int 2011 Jun;26(2):148-162 [FREE Full text] [doi: 10.1093/heapro/daq050] [Medline: 20739325]
105. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
106. Bedell GM, Wade SL, Turkstra LS, Haarbauer-Krupa J, King JA. Informing design of an app-based coaching intervention to promote social participation of teenagers with traumatic brain injury. Dev Neurorehabil 2017 Oct 28;20(7):408-417. [doi: 10.1080/17518423.2016.1237584] [Medline: 27792407]
107. Bui TA, Pohl M, Rosenfelt C, Ogourtsova T, Yousef M, Whitlock K, et al. Identifying potential gamification elements for a new chatbot for families with neurodevelopmental disorders: user-centered design approach. JMIR Hum Factors 2022 Aug 19;9(3):e31991 [FREE Full text] [doi: 10.2196/31991] [Medline: 35984679]
108. Brunelle K, Abdulle S, Gorey KM. Anxiety and depression among socioeconomically vulnerable students with learning disabilities: exploratory meta-analysis. Child Adolesc Soc Work J 2019 Aug 2;37:359-367. [doi: 10.1007/s10560-019-00631-w]
109. White SW, Oswald D, Ollendick T, Scahill L. Anxiety in children and adolescents with autism spectrum disorders. Clin Psychol Rev 2009 Apr;29(3):216-229 [FREE Full text] [doi: 10.1016/j.cpr.2009.01.003] [Medline: 19223098]
110. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
111. Banfield J, Wilkerson B. Increasing student intrinsic motivation and self-efficacy through gamification pedagogy. Contemp Issues Educ Res 2014 Sep 16;7(4):291-298. [doi: 10.19030/cier.v7i4.8843]
112. Yu SW, Hill C, Ricks ML, Bennet J, Oriol NE. The scope and impact of mobile health clinics in the United States: a literature review. Int J Equity Health 2017 Oct 05;16(1):178 [FREE Full text] [doi: 10.1186/s12939-017-0671-2] [Medline: 28982362]
113. Hoffmann A, Christmann CA, Bleser G. Gamification in stress management apps: a critical app review. JMIR Serious Games 2017 Jun 07;5(2):e13 [FREE Full text] [doi: 10.2196/games.7216] [Medline: 28592397]

114. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
115. Joshi M, Bhardwaj P. Impact of data transparency: scientific publications. Perspect Clin Res 2018;9(1):31-36. [doi: 10.4103/picr.picr_104_17]

Abbreviations

ASD: autism spectrum disorder
mHealth: mobile health
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RCT: randomized clinical trial

Edited by T Leung, S Ma; submitted 15.05.23; peer-reviewed by P Gorp, Y Yang; comments to author 23.11.23; revised version received 31.01.24; accepted 27.06.24; published 06.09.24.

Please cite as:
Mahmoudi E, Yejong Yoo P, Chandra A, Cardoso R, Denner Dos Santos C, Majnemer A, Shikako K
Gamification in Mobile Apps for Children With Disabilities: Scoping Review
JMIR Serious Games 2024;12:e49029
URL: https://games.jmir.org/2024/1/e49029
doi:10.2196/49029
PMID:39240675

Review

Effectiveness of Technological Interventions for Older Adults With Parkinson Disease: Systematic Review

Roberta Bevilacqua[1], MSPsyc; Marco Benadduci[1], MSc; Federico Barbarossa[1], MEng; Giulio Amabili[1], MEng; Valentina Di Donna[1], MD; Clotilda Martella[1], MD; Giuseppe Pelliccioni[1], MD; Giovanni Renato Riccardi[1], MD; Elvira Maranesi[1], PhD

IRCCS INRCA, Ancona, Italy

Corresponding Author:
Federico Barbarossa, MEng
IRCCS INRCA
Via Santa Margherita 5
Ancona, 60124
Italy
Phone: 39 0718004767
Email: f.barbarossa@inrca.it

Abstract

Background: Among the older population, Parkinson disease (PD) stands out as a leading contributor to disability. Clinically, the foremost objectives in managing PD involve proactively delaying and preventing disability. Understanding the pivotal role of gait and balance in daily functionality holds substantial clinical significance, signaling imminent disability and prompting a reevaluation of management approaches. A key priority lies in identifying novel and effective interventions for symptoms that substantially contribute to disability.

Objective: This paper presents a systematic review that critically examines the existing body of literature on the use of technology in the rehabilitation of older patients with PD. By synthesizing current evidence, we aim to provide insights into the state of the field, identify gaps in knowledge, and offer recommendations for future research and clinical practice.

Methods: A systematic review of the literature was conducted in September 2023 analyzing manuscripts and papers of the last 5 years from the PubMed, Scopus, Embase, Web of Science, and CINAHL databases following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. A total of 14 papers were included. The inclusion criteria are as follows: (1) randomized controlled trial, (2) PD in people aged 65 years and older, and (3) use of technology in the rehabilitation training in the older population.

Results: A large portion of effective interventions relies on the incorporation of technology, particularly through virtual reality exergames. This technology appears to have effects not only on the cognitive aspect but also on the physical domain. The analysis of the results clearly indicates that, in terms of gait and balance performance, the technological intervention outperforms the traditional approach, irrespective of the specific technology employed.

Conclusions: This systematic review seeks to shed light on the evolving landscape of technology-assisted rehabilitation for older individuals with PD. As we delve into the available evidence, we will assess the extent to which technology can serve as a valuable adjunct to conventional therapy, offering new avenues for optimized care and improved outcomes in this growing patient demographic. As we sift through the existing evidence, our goal is to evaluate the potential of technology as a valuable supplement to traditional therapy, presenting fresh opportunities for enhanced care and better outcomes in this expanding patient demographic.

(JMIR Serious Games 2024;12:e53431) doi:10.2196/53431

KEYWORDS

technological intervention; Parkinson disease; randomized controlled trail; older adults; efficacy

Introduction

Neurological disorders stand as the primary cause of global disability [1]. Within this category, Parkinson disease (PD) is exhibiting the most rapid increase in disability, fatalities, and prevalence, with an anticipated doubling within the next 2 decades [2]. In Europe, the current PD population exceeds 1.2 million, contributing to an annual economic burden of €13.9

billion (US $15.5 billion) [3]. The annual cost per individual rises with the severity of the condition, while nonmotor symptoms significantly contribute to hospitalizations and institutionalizations. Beyond the challenging aspects of PD itself, individuals with PD grapple with notable comorbidities and a high frequency of falls [4]. Aging is correlated with an accelerated motor progression of the disease; reduced responsiveness to levodopa; heightened severity in gait, posture, and cognitive impairment; and an increased likelihood of developing dementia [5].

With the aging of the global population, the prevalence of PD is on the rise, making it an increasingly significant public health concern. Effective management and rehabilitation strategies are essential to mitigate the debilitating effects of this condition and to enhance the functional independence of older individuals living with PD.

From a clinical perspective, the top priorities in PD management include the delay and prevention of disability [6]. Despite symptomatic relief through medical, surgical, and rehabilitative interventions, older individuals with PD experience a persistent deterioration in disability, marked by a decline in quality of life, diminished functional mobility, decreased performance in daily activities, and worsening neurological impairments. Guidelines suggest early initiation of physical therapy at the onset of the disease for addressing functional decline [7]. However, the evidence regarding its effectiveness in delaying symptom onset or reducing severity is not robust. Recognizing the crucial role of gait and balance in daily function has significant clinical implications, indicating impending disability and necessitating a reassessment of management strategies. The identification of new and effective interventions for symptoms contributing to significant disability is a key priority [8]. In particular, there is a growing body of evidence indicating that exercise and physical activity interventions can slow the rate of functional mobility decline, ultimately improving quality of life, as tertiary prevention solutions [8].

Over the years, advances in technology have revolutionized the field of health care, offering innovative tools and interventions for the diagnosis, monitoring, and treatment of various medical conditions. In recent decades, there has been a growing interest in harnessing the potential of technology to aid in the rehabilitation of patients with PD, particularly in the older population. This shift reflects a broader trend toward incorporating digital solutions into health care, driven by the desire to optimize therapeutic outcomes and improve the overall well-being of patients.

For this reason, recent research [9-11] indicates that balance training delivered through technology leads to performance enhancements that align with noticeable neurobiological changes in the cerebral cortex [9,10]. This underscores the encouraging potential of technological interventions in aiding individuals with PD in managing balance and other motor disorders.

This paper presents a systematic review that critically examines the existing body of literature on the use of technology in the rehabilitation of older patients with PD. By synthesizing current evidence, we aim to provide insights into the state of the field, identify gaps in knowledge, and offer recommendations for future research and clinical practice.

In this comprehensive review, we will explore a wide range of technological applications, including wearable devices, telerehabilitation platforms, virtual reality systems, and robotics, among others. We will evaluate the effectiveness, usability, and safety of these technologies in improving motor function, balance, mobility, and overall quality of life in older patients with PD. Furthermore, we will consider the potential challenges and ethical implications associated with the integration of technology into rehabilitation protocols for this vulnerable population.

Methods

Overview

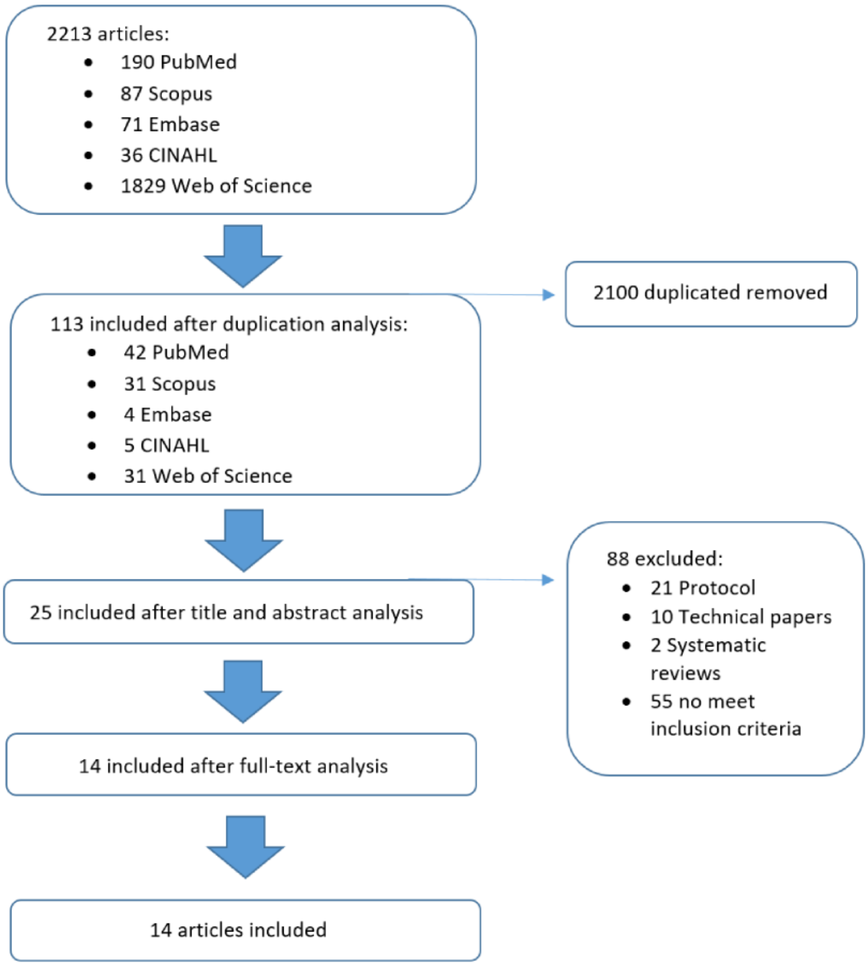
The methodology of this systematic review was based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Multimedia Appendix 1 [12]) [13], with the main aim of analyzing the use of technology in the rehabilitation of older patients with PD.

We used the PICO (Population, Intervention, Comparator, Outcome) framework as follows:

- P: Older patients with PD
- I: Rehabilitation program with technology
- C: Control group that receives a traditional rehabilitation program
- O: The efficacy of the treatment

A systematic review of the literature was conducted in September 2023. The data were collected from PubMed, Embase, Scopus, CINAHL, and Web of Sciences, analyzing manuscripts and papers of the last 5 years (from September 2018 to September 2023) in order to obtain the latest evidence in the field. The inclusion criteria were as follows: (1) randomized controlled trial, (2) PD in people aged 65 years and older, and (3) use of technology in the rehabilitation training in the older population. Systematic and narrative reviews were excluded. Based on consultation with the multidisciplinary research team, technological intervention studies were searched using the following search terms and the combination thereof: Parkinson, rehabilitation, technology, old*, elder*, effectiveness, randomized control trial. The detailed search strategy used in each database is reported in Multimedia Appendix 2. After the preliminary search, 190 papers resulted from PubMed, 71 from Embase, 87 from Scopus, 36 from CINAHL, and 1829 from Web of Sciences. Since the CINAHL database was consulted only after the first round of searches and after the duplicate analysis, a different filtering of the results was chosen. In fact, in order to narrow down the number of results, we looked for papers containing the term “randomized control trial” in the title. The findings were analyzed and screened by 4 experts of the team: a bioengineer, a clinical neurologist, a psychologist, and a physiotherapist. Rayyan software has been used in order to remove duplicates. Identified studies were independently reviewed for eligibility in a 2-step process; a first screening was performed by 4 independent authors (EM, GP, RB, and MB) based on the title and abstract. Then, full texts were retrieved for a second screening by the same 4 authors. At both stages, disagreements among reviewers were resolved by consensus and discussed with the other 2 authors (GA and VDD). After these steps, 25 papers resulted from selected databases. An additional researcher, with a background in biomedical engineering, confirmed the accuracy of the papers' selection and screened for any possible omission. After the screening based on the inclusion or exclusion criteria, conducted on the full-text papers, 14 studies were selected. Figure 1 shows the flowchart search strategy applied. Data extraction, performed by 2 independent raters (EM and MB) included full reference details, sample size, study population details, including gender and mean age, type of technology used, outcomes, and results.

Figure 1. Descriptive analysis of the included clinical studies.



Quality appraisal of included studies was carried out by applying the PEDro scale suggested for evidence-based reviews [14]. The final score was settled when 3 authors reached agreement after repeated review and analysis.

Data Synthesis

A descriptive synthesis of the data was performed. The data were grouped and analyzed based on the types of technological interventions used and their reported outcomes. The main areas of focus included cognitive and physical functioning, specifically gait and balance performance. The synthesis aimed to identify common findings and trends across the studies, highlighting the efficacy of technological interventions compared with traditional rehabilitation methods. No meta-analysis was conducted due to the heterogeneity of the study designs, interventions, and outcome measures.

Results

Study Quality Evaluation

A total of 14 papers were included [15-28]. Findings reported in this section are organized under macroconcept areas of interest.

Table 1 reports the quality evaluation of 14 population-based studies. In particular, the PEDro score ranged from 6 to a maximum of 8. The descriptive analysis of the technological intervention and devices used in the studies selected is shown in Multimedia Appendix 3.

Table 1. Scores of methodological quality assessment of the included studies.

PEDro	Eligibility	Randomized allocation	Concealed allocation	Baseline comparability	Blinded subject	Blinded therapists	Blinded raters	Key outcomes	Intention to treat	Comparison between groups	Precision and variability	Score
Del Din et al [15]	Y[a]	Y	Y	Y	N[b]	N	N	Y	Y	Y	Y	8/11
Jäggi et al [16]	Y	Y	Y	Y	N	N	N	Y	N	Y	N	6/11
Lau et al [17]	Y	Y	Y	Y	N	N	N	Y	N	Y	N	6/11
Kim et al [18]	Y	Y	Y	Y	N	N	N	Y	Y	Y	N	7/11
Maranesi et al [19]	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	7/11
Gryfe et al [20]	Y	Y	Y	Y	N	N	N	Y	Y	Y	N	7/11
Feng et al [21]	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	7/11
Cikajlo and Peterlin Potisk [22]	Y	Y	Y	Y	N	N	N	Y	N	Y	N	6/11
Capecci et al [23]	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	7/11
Spina et al [24]	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	7/11
Pazzaglia et al [25]	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	7/11
Yuan et al [26]	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	7/11
Calabrò et al [27]	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	8/11
Alagumoorthi et al [28]	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	8/11

[a]Y: yes.
[b]N: No.

General Characteristics of the Study Population

All the studies were focused on older people with PD with a mean age of 69.8 (SD 14.4) years for the technological intervention group and mean age of 70.3 (SD 7.3) years in the control group. The number of participants involved in all the studies is 578, ranging from a minimum sample of 18 patients to a maximum of 128. There were 297 males and 281 females.

Descriptive Analysis and Outcome Measures

Table 2 shows the characteristics of the included studies. All studies evaluated the impact of technological interventions on physical and cognitive domains. Regarding evaluated domains, cognitive functioning was assessed by 4 studies [16,17,20,22] and physical functioning was assessed by 13 studies [15,17-28], including gait assessment [15,17-20,23-25,27,28] in terms of gait speed or gait performance, falls and fear of falling evaluation [15,18,19,28], and balance estimation [16-18,20,21,24,28]. Examining the technology employed, there is a notable preference for using treadmills coupled with virtual reality and exergames [15,17,18,27]. In addition, numerous studies [16,19,23,24,28] incorporate sensitive pressure platforms, both static and dynamic, often in conjunction with nonimmersive virtual exercises. However, for the specific pathology under consideration, the use of exoskeletons [20] or gait robots [23] appears to be more restricted. It is worth noting that nearly all studies consistently incorporate virtual reality—whether immersive with visors or caves, or nonimmersive through exergames. This technology appears to have effects not only on the cognitive aspect but also on the physical domain. The analysis of the results clearly indicates that, in terms of gait and balance performance, the technological intervention outperforms the traditional approach, irrespective of the specific technology employed. However, from a cognitive perspective, the improvement is limited.

Table 2. Descriptive analysis of the included clinical studies.

	Population		Technology	Outcomes	Results
	Participants in EG[a]	Participants in CG[b]			
Del Din et al [15]	N=66 (older fallers with PD[c]); 81 male/47 female, aged 78.03 (6.21) years (older fallers with PD). HY[d] stage: 2-3	N=62 (older fallers with PD)	Treadmill plus virtual reality	*Primary*: FRA[e] index. *Secondary*: walking activity (total walking time per day, percentage of walking time per day, number of bouts, and steps per day)	*Between*: FRA was higher in PD compared with older fallers and MCI[f] (*P*=.043). Walker activity was lower in MCI and patients with PD compared with older fallers (*P*<.012) *Within*: Walking activity did not change. FRA significantly decreased for all groups following both interventions (treadmill reduced FRA by 26%; treadmill + VR[g] by 39%)
Jäggi et al [16]	N=19 (older people with PD); 12 male/7 female; aged 71.89 (9.09) years. HY stage: 1-4	N=21 (older people with PD); 15 male/6 female; aged 72.86 (10.14) years. HY stage: 1-4	Cognitive-motor training on the exergame device. Dividat Senso, a pressure-sensitive platform	*Primary*: feasibility (adherence rate; attrition rate; occurrence of adverse events; SUS[h], NASA TLX[i] score) *Secondary*: cognitive (Go/no go test; RTT[j]; D-KEFS[k]; TMT[l]); motor (gait speed; SPPB[m]; TUG[n]; 5×StS[o])	*Primary*: Overall adherence rate was 96.5%. EG had an adherence rate of >70%. *Secondary*: no differences between EG and CG. Significant time-group interaction effects for 5×StS, SPPB, RTT, go/no go test, and D-KEFS
Lau et al [17]	N=9 (older people with PD); 6 male/3 female; aged 64 (9) years. HY stage: 1-3	N=9 (older people with PD); 6 male/3 female; aged 71 (5) years. HY stage: 1-3	Treadmill combined with a first-person immersive video game targeting visuospatial skills and working memory.	*Primary*: motor outcomes (6MWT[p], TUG, TUG cognitive) *Secondary*: cognitive outcomes: (MoCA[q], verbal fluency, SDMT[r])	*Primary*: EG improves gait speed and walking distance during 6MWT, TUG cognitive (*P*=.05). EG improves TUG cognitive more that CG (between group: *P*=.04) *Secondary*: EG improves MoCA (*P*=.007) and SDMT (*P*=.01)
Kim et al [18]	N=22 (older people with PD); 6 male/16 female; aged 68.7 (6.9) years. HY stage: 2-3	N=22 (older people with PD); 7 male/15 female; aged 67.5 (9.3) years. HY stage: 2-3	Robot-assisted gait training (treadmill-based exoskeleton robot)	*Primary*: gait speed on the 10mWT[s] *Secondary*: dual task interference on gait speed on the 10mWT, balance (TUG), disability score (BBS[t]), fear of falling (MDS-UPDRS[u]), freezing of gait (KFES-I[v]), brain functional connectivity changes (NFOGQ[w])	*Primary*: no significant difference in gait speed on the 10mWT *Secondary*: EG shows significant improvements in BBS score and in MDS-UPDRS score. *Between*: EG shows significant difference in MDS-UPDRS
Maranesi et al [19]	N=16 (older people with PD); 6 male/10 female; aged 72.7 (6.3) years. HY stage: 2-3	N=14 (older people with PD); 5 male/9 female; aged 75.5 (5.4) years. HY stage: 2-3	Tymo system (wireless platform with nonimmersive virtual reality exergame)	*Primary*: POMA[x] balance, POMA gait, and POMA total *Secondary*: gait speed, fear of falling (FES-I), autonomy in daily living activities (Barthel Index), and physical and psychological state of the patients (12-item Short Form Survey)	*Primary: Between*: POMA balance and POMA gait (EG vs CG) *Secondary*: *Between*: Barthel Index, 12-item Short Form Survey (EG vs CG)
Gryfe et al [20]	N=13 (older people with PD); 4 male/9 female; aged 67.6 (5.9) years. HY stage: 2-3	N=14 (older people with PD); 7 male/7 female; aged 70.7 (7.3) years. HY stage: 2-3	Gait robot (exoskeleton)	*Primary*: cognitive function (SCOPA-COG[y]) and mood *Secondary*: gait speed, 6MWT, freezing of gait, balance, and quality of life	*Primary*: significant improvement in SCOPA-COG in EG than in CG *Secondary*: significant improvement in 6MWT in EG than in CG

	Population		Technology	Outcomes	Results
	Participants in EG[a]	Participants in CG[b]			
Feng et al [21]	N=14 (older people with PD); 7 male/7 female; aged 67.5 (4.8) years. HY stage: 2-4	N=14 (older people with PD); 8 male/6 female; aged 66.9 (4.6) years. HY stage: 2-4	Virtual reality technology	Motor ability (BBS, TUG, UPDRS-III, and FGA[z])	*Within*: BBS, TUG, and FGA improve significantly in EG and CG (*P*<.05). *Between groups*: BBS, TUG, UPDRS-III, and FGA in EG are better than in EG (*P*<.05).
Cikajlo et al [22]	N=10 (older people with PD), 4 male/6 female, aged 71.3 (8.4) years	N=10 (older people with PD), 5 male/5 female, aged 67.6 (7.6) years	Immersive VR using 3D Oculus Rift CV1	*Primary*: effectiveness of treatment (BBT, UPDRS) *Secondary*: motivation effect (Motivation Inventory)	*Primary*: time of manipulation (group × time, *P*=.009), number of successfully placed cubes (group × time, *P*=.028), average tremor (group × time, *P*=.002), and UPDRS for upper limb (U3=0.35). The LCD and 3D groups substantially improved their BBT score with training (U3=0.7, U3=0.6, respectively).
Capecci et al [23]	N=48 (older people with PD), 19 male/29 female, aged 68.1 (9.8) years. HY stage: ≥2	N=48 (older people with PD), 24 male/24 female, aged 67.0 (7.6) years. HY stage: ≥2	GE-O system	*Primary*: motor function (6MWT, TUG, 10MWT, minimal clinically important difference) *Secondary*: Freezing of Gait Questionnaire, UPDRS	Both groups showed significant improvement in all outcomes. As compared with baseline, with robot-assisted gait training and treadmill training, endurance and gait capacity were enhanced by 18% and 12%, respectively, and motor symptoms and quality of life were improved by 17% and 15%, respectively. The maximum advantage was observed with the Freezing of Gait Questionnaire score, which decreased by 20% after either treatment.
Spina et al [24]	N=11 (older people with PD), 5 male/6 female, aged 68 (6.9) years. HY stage: 1-2	N=11 (older people with PD), 4 male/7 female, aged 67.27 (4.85) years. HY stage: 1-2	Hunova (robotic platform)	*Primary*: Quantified balance impairments (mini-BESTest, BBS) *Secondary*: Motor ability (10MWT, 5×StS, PDQ-39[aa], TUG)	*Between*: no significant differences *Within*: primary outcomes improved in EG and CG.
Pazzaglia et al [25]	N=25 (older people with PD), 18 male/7 female, aged 72 (7) years	N=26 (older people with PD), 17 male/9 female, aged 70 (10) years	Nirvana (VR system)	*Primary*: changes in functional standing balance (BBS) *Secondary*: ability to adapt gait to complex walking tasks (DGI[ab]); physical function of the upper limb (DASH[ac]); and physical and emotional scores (SF-36[ad])	*Between*: in EG increases BBS score (*P*=.003), DGI score (*P*=.003), and SF-36 mental composite score (*P*=.037), and a decrease in DASH scale score (*P*=.009). In CG DASH scale score decreases (*P*=.007).
Yuan et al [26]	N=12 (older people with PD), 2 male/10 female, aged 67.8 (5.5) years. HY stage: 1-3	N=12 (older people with PD), 9 male/3 female, aged 66.5 (8.8) years. HY stage: 1-3	Video game–based treatment	*Primary*: Quantified balance impairments (BBS score) *Secondary*: Quality of life and motor/balance ability (SF-36, MFES[ae], MDRT[af], and Maximum Step Length test)	*Between*: changes in MFES and MDRT to the right and left sides were significantly different in the first 6-week period. Changes in BBS, MFES, and MDRT to the right and left sides were significantly different in the second 6-week period.

	Population		Technology	Outcomes	Results
	Participants in EG[a]	Participants in CG[b]			
Calabrò et al [27]	N=25 (older people with PD), 11 male/9 female, aged 70 (8) years. HY stage: 2-3	N=25 (older people with PD), 9 male/3 female, aged 73 (8) years. HY stage: 2-3	Treadmill plus music	*Primary*: gait performance (FGA) *Secondary*: brain oscillation changes related to gait cycle; gait and motor performance: (UPDRS, BBS, FES, 10MWT, TUG, and GQI[ag])	*Between*: improvement in FGA (*P*<.001), FES (*P*<.001), UPDRS (*P*=.001), and overall GQI (*P*<.001) in EG
Alagumoorthi et al [28]	N=96 (with PD), 51 male/45 female, aged 69.7 (10) years. HY stage 2-3-4	N=96 (with PD), 63 male/33 female, aged 68.5 (9.8) years. HY stage: 2-3-4	Nintendo Wii Console	*Primary*: number of fallers *Secondary*: fall rate; risk of falling and quality of life (BBS, TUG, and PDQ-39)	End of treatment: *Between*: number of fallers (*P*=.77), BBS (*P*=.658), TUG (*P*=.967), and PDQ-39 (*P*=.402). Follow-up 36th week: *Between*: number of fallers (*P*=.039), BBS (*P*=.867), TUG (*P*=.959), and PDQ-39 (*P*=.405).

[a]EG: experimental group.
[b]CG: control group.
[c]PD: Parkinson disease.
[d]HY: Hoehn & Yahr.
[e]FRA: Falls Rate and Activity Exposure Index.
[f]MCI: mild cognitive impairment.
[g]VR: virtual reality.
[h]SUS: System Usability Scale.
[i]NASA TLX: NASA Task Load Index.
[j]RTT: reaction time test.
[k]D-KEFS: Delis-Kaplan Executive Function System; color-word interference test.
[l]TMT: Trail Making Test A and B.
[m]SPPB: Short Physical Performance Battery.
[n]TUG: time up and go.
[o]5×StS: 5 times sit to stand.
[p]6MWT: 6-minute walking test.
[q]MoCA: Montreal Cognitive Assessment.
[r]SDMT: Symbol Digit Modality Test.
[s]10mWT: 10-minute walking test.
[t]BBS: Berg Balance Scale.
[u]MDS-UPDRS: Movement Disorder Society-sponsored version of the Unified Parkinson’s Disease Rating Scale.
[v]KFES-I: Korean version of Falls Efficacy Scale-International.
[w]NFOGQ: New Freezing of Gait Questionnaire.
[x]POMA: Tinetti Performance Oriented Mobility Assessment.
[y]SCOPA-COG: Scales for Outcomes in Parkinson’s Disease-Cognition.
[z]FGA: Functional Gait Assessment.
[aa]PDQ-39: 39-item Parkinson’s Disease Questionnaire.
[ab]DGI: Dynamic Gait Index.
[ac]DASH: Disabilities of the Arm, Shoulder and Hand.
[ad]SF-36: 36-item Short Form.
[ae]MFES: Modified Falls Efficacy Scale.
[af]MDRT: Multi-Directional Reach Test.
[ag]GQI: Gait Quality Index.

Discussion

Principal Findings

This systematic review seeks to shed light on the evolving landscape of technology-assisted rehabilitation for older people with PD. Delving into the available evidence, we assess how technology can serve as a valuable complement to conventional therapy, offering new ways to optimize care and improve outcomes in this growing patient demographic. In particular, the included studies evaluated the impact of technological interventions on physical and cognitive domains. Cognitive functioning was assessed in 4 studies, while physical functioning, including gait, falls, and balance, was assessed in 13 studies. Technologies used included treadmills with virtual reality, exergames, and sensitive pressure platforms. Exoskeletons and gait robots were less commonly used. Most studies incorporated virtual reality, either immersive or nonimmersive. The results indicate that technological interventions significantly improve gait and balance performance compared with traditional methods, though cognitive improvements were limited.

PD poses a serious burden on patients, carers, families, health care providers, and health authorities globally. Within the relevant unmet need in the management of PD [10], identifying preventive strategies, particularly tertiary, to mitigate severe progression and ensure an adequate quality of life for older patients is one of the most felt needs, with the final aim of slowing disease progression, enhancing the long-term quality of life, and reducing the costs associated with PD, as the majority of expenses occur in advanced stages.

From our analysis, it seems that including technology into a physical intervention can constitute a driver for ensuring the adherence and compliance to rehabilitation trainings. In particular, our analysis reveals an overall improvement in the physical status and outcomes related to disease severity, even if a not clear improvement in cognitive status and quality of life was observed. This result may suggest the need of conducting longer trials to collect evidence on the impact of rehabilitative interventions with technology also in the long term. Moreover, even if the adherence to rehabilitation training is high in hospital facilities, this should be complemented and sustained with training at home to stimulate compliance to physical exercise, considering that patients with PD should remain active as long as possible to counteract the disease progression. Additionally, digital solutions can be a viable option to support the older patients at home by coaching them in an interactive and engaging way, as already demonstrated by numerous evidence in the field [29].

In fact, although in a different context, our analysis has shown that the most engaging interventions involved virtual reality technology, which appears to be the most promising tool for positively impacting cognitive domains due to its highly interactive features. This is in line with recent findings that highlight that virtual reality technology has a particularly promising tool for investigating and rehabilitating gait and balance impairments in people with PD by enabling users to engage in an enriched and highly personalized complex environment [30].

The inclusion of technology and virtual reality into rehabilitative training imposes to stress the attention to assessing the usability and accessibility of such technologies by including the patients-in-the-loop since the design of the solutions in order to be suitable and to not compromise the appropriate use. At this purpose, even if there is a plethora of studies and disciplines that underline the needs of involving older adults and vulnerable user groups in the design process, also with a specific focus on age-related conditions such as dementia [29], there has been limited exploration of how individuals with PD can actively participate in the design process [31], despite the fact that the PD population represents an intriguing user group for design investigation due to the intricate and individual nature of the condition that includes physical and cognitive symptoms, with fluctuations in severity, while dealing with emotional challenges related to social stigma and embarrassment associated with their condition [32].

However, this systematic review has its limitations. The low number of studies and their heterogeneity do not favor valid and generalizable conclusions, but this very limitation shows that, despite the good promises, there is still a need to experiment with technological interventions and, in particular, with virtual reality, in the rehabilitation of people with PD. Moreover, all the studies reported show more or less significantly positive results: this aspect is typical of scientific publications, in which there is a tendency to publish this kind of study to the detriment of those with negative results. For the purposes of research into the implementation of new technologies for the treatment of patients with PD, it is very important to know and take into account even those studies that did not produce the desired results, so that lessons should not be learned again.

The integration of technology into PD rehabilitation holds the promise of enhancing the efficacy and accessibility of interventions, addressing the unique challenges faced by older individuals, and facilitating personalized, patient-centered care. Future challenges of sociotechnical interventions should be focused on unmet needs of the older population with PD by combining a multidisciplinary, multicomponent, and personalized approach and addressing challenges related to health literacy, availability of physical and psychological services, and the stigma associated with PD [10].

Conclusions

In summary, this systematic review seeks to shed light on the evolving landscape of technology-assisted rehabilitation for older individuals with PD. The integration of technology into PD rehabilitation holds the promise of enhancing the efficacy and accessibility of interventions, addressing the unique challenges faced by older individuals, and facilitating personalized patient-centered care.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses).
[DOC File , 64 KB - games_v12i1e53431_app1.doc]

Multimedia Appendix 2
Search strategy used in each database.
[DOCX File , 17 KB - games_v12i1e53431_app2.docx]

Multimedia Appendix 3
Descriptive analysis of the technology used in the interventions analyzed.
[DOCX File , 18 KB - games_v12i1e53431_app3.docx]

References

1. Kasner SE, Swaminathan B, Lavados P, Sharma M, Muir K, Veltkamp R, et al. Rivaroxaban or aspirin for patent foramen ovale and embolic stroke of undetermined source: a prespecified subgroup analysis from the NAVIGATE ESUS trial. Lancet Neurol 2018 Dec;17(12):1053-1060 [FREE Full text] [doi: 10.1016/S1474-4422(18)30319-3] [Medline: 30274772]
2. Tinelli M, Kanavos P, Grimaccia F. The value of early diagnosis in treatment in Parkinson's disease - a literature review of the potential clinical and socioeconomic impact of targeting unmet needs in Parkinson's disease. LSE 2016:1-41.
3. Olesen J, Gustavsson A, Svensson M, Wittchen H, Jönsson B, CDBE2010 study group, European Brain Council. The economic cost of brain disorders in Europe. Eur J Neurol 2012;19(1):155-162. [doi: 10.1111/j.1468-1331.2011.03590.x] [Medline: 22175760]
4. Canning CG, Paul SS, Nieuwboer A. Prevention of falls in Parkinson's disease: a review of fall risk factors and the role of physical interventions. Neurodegener Dis Manag 2014;4(3):203-221. [doi: 10.2217/nmt.14.22] [Medline: 25095816]
5. Yilmaz R, Hopfner F, van Eimeren T, Berg D. Biomarkers of Parkinson's disease: 20 years later. J Neural Transm (Vienna) 2019;126(7):803-813. [doi: 10.1007/s00702-019-02001-3] [Medline: 30949837]
6. Shulman LM. Understanding disability in Parkinson's disease. Mov Disord 2010;25 Suppl 1:S131-S135. [doi: 10.1002/mds.22789] [Medline: 20187231]
7. KNGF Guidelines for Physical Therapy in Patients with PD. UK: Dutch Physioth; 2004.
8. Dibble LE, Cavanaugh JT, Earhart GM, Ellis TD, Ford MP, Foreman KB. Charting the progression of disability in Parkinson disease: study protocol for a prospective longitudinal cohort study. BMC Neurol 2010;10:110 [FREE Full text] [doi: 10.1186/1471-2377-10-110] [Medline: 21047426]
9. Sehm B, Taubert M, Conde V, Weise D, Classen J, Dukart J, et al. Structural brain plasticity in Parkinson's disease induced by balance training. Neurobiol Aging 2014;35(1):232-239. [doi: 10.1016/j.neurobiolaging.2013.06.021] [Medline: 23916062]
10. Robinson AG, Dennett AM, Snowdon DA. Treadmill training may be an effective form of task-specific training for improving mobility in people with Parkinson's disease and multiple sclerosis: a systematic review and meta-analysis. Physiotherapy 2019;105(2):174-186. [doi: 10.1016/j.physio.2018.11.007] [Medline: 30876717]
11. Gaßner H, Steib S, Klamroth S, Pasluosta CF, Adler W, Eskofier BM, et al. Perturbation treadmill training improves clinical characteristics of gait and balance in Parkinson's disease. J Parkinsons Dis 2019;9(2):413-426. [doi: 10.3233/JPD-181534] [Medline: 30958316]
12. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
13. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
14. Maher C, Sherrington C, Herbert RMA, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther 2003;83:713-721. [doi: 10.1093/ptj/83.8.713]
15. Del Din S, Galna B, Lord S, Nieuwboer A, Bekkers E, Pelosin E, et al. Falls risk in relation to activity exposure in high-risk older adults. J Gerontol A Biol Sci Med Sci 2020 May 22;75(6):1198-1205 [FREE Full text] [doi: 10.1093/gerona/glaa007] [Medline: 31942969]
16. Jäggi S, Wachter A, Adcock M, de Bruin ED, Möller JC, Marks D, et al. Feasibility and effects of cognitive-motor exergames on fall risk factors in typical and atypical Parkinson's inpatients: a randomized controlled pilot study. Eur J Med Res 2023;28(1):30 [FREE Full text] [doi: 10.1186/s40001-022-00963-x] [Medline: 36647177]

17. Lau J, Regis C, Burke C, Kaleda M, McKenna R, Muratori LM. Immersive technology for cognitive-motor training in Parkinson's disease. Front Hum Neurosci 2022;16:863930 [FREE Full text] [doi: 10.3389/fnhum.2022.863930] [Medline: 35615742]
18. Kim H, Kim E, Yun SJ, Kang M, Shin HI, Oh B, et al. Robot-assisted gait training with auditory and visual cues in Parkinson's disease: a randomized controlled trial. Ann Phys Rehabil Med 2022;65(3):101620 [FREE Full text] [doi: 10.1016/j.rehab.2021.101620] [Medline: 34896605]
19. Maranesi E, Casoni E, Baldoni R, Barboni I, Rinaldi N, Tramontana B, et al. The effect of non-immersive virtual reality exergames versus traditional physiotherapy in Parkinson's disease older patients: preliminary results from a randomized-controlled trial. Int J Environ Res Public Health 2022;19(22):14818 [FREE Full text] [doi: 10.3390/ijerph192214818] [Medline: 36429537]
20. Gryfe P, Sexton A, McGibbon CA. Using gait robotics to improve symptoms of Parkinson's disease: an open-label, pilot randomized controlled trial. Eur J Phys Rehabil Med 2022;58(5):723-737 [FREE Full text] [doi: 10.23736/S1973-9087.22.07549-9] [Medline: 35708047]
21. Feng H, Li C, Liu J, Wang L, Ma J, Li G, et al. Virtual reality rehabilitation versus conventional physical therapy for improving balance and gait in Parkinson's disease patients: a randomized controlled trial. Med Sci Monit 2019 Jun 05;25:4186-4192 [FREE Full text] [doi: 10.12659/MSM.916455] [Medline: 31165721]
22. Cikajlo I, Peterlin Potisk K. Advantages of using 3D virtual reality based training in persons with Parkinson's disease: a parallel study. J Neuroeng Rehabil 2019;16(1):119 [FREE Full text] [doi: 10.1186/s12984-019-0601-1] [Medline: 31623622]
23. Capecci M, Pournajaf S, Galafate D, Sale P, Le Pera D, Goffredo M, et al. Clinical effects of robot-assisted gait training and treadmill training for Parkinson's disease. A randomized controlled trial. Ann Phys Rehabil Med 2019;62(5):303-312 [FREE Full text] [doi: 10.1016/j.rehab.2019.06.016] [Medline: 31377382]
24. Spina S, Facciorusso S, Cinone N, Armiento R, Picelli A, Avvantaggiato C, et al. Effectiveness of robotic balance training on postural instability in patients with mild Parkinson's disease: a pilot, single blind, randomized controlled trial. J Rehabil Med 2021 Feb 17;53(2):jrm00154. [doi: 10.2340/16501977-2793] [Medline: 33585943]
25. Pazzaglia C, Imbimbo I, Tranchita E, Minganti C, Ricciardi D, Lo Monaco R, et al. Comparison of virtual reality rehabilitation and conventional rehabilitation in Parkinson's disease: a randomised controlled trial. Physiotherapy 2020;106:36-42. [doi: 10.1016/j.physio.2019.12.007] [Medline: 32026844]
26. Yuan RY, Chen SC, Peng CW, Lin YN, Chang YT, Lai CH. Effects of interactive video-game-based exercise on balance in older adults with mild-to-moderate Parkinson's disease. J Neuroeng Rehabil 2020;17(1):91 [FREE Full text] [doi: 10.1186/s12984-020-00725-y] [Medline: 32660512]
27. Calabrò RS, Naro A, Filoni S, Pullia M, Billeri L, Tomasello P, et al. Walking to your right music: a randomized controlled trial on the novel use of treadmill plus music in Parkinson's disease. J Neuroeng Rehabil 2019;16(1):68 [FREE Full text] [doi: 10.1186/s12984-019-0533-9] [Medline: 31174570]
28. Alagumoorthi G, Beulah JD, Thirunavukarasu S, Ramachandaran V, Kumaresan A. Effectiveness of Wii sports- based strategy training in reducing risk of falling, falls and improving quality of life in adults with idiopathic Parkinson's disease- a randomized comparative trial. Clin Rehabil 2022;36(8):1097-1109. [doi: 10.1177/02692155221089030] [Medline: 35315706]
29. Bevilacqua R, Casaccia S, Cortellessa G, Astell A, Lattanzio F, Corsonello A, et al. Coaching through technology: a systematic review into efficacy and effectiveness for the ageing population. Int J Environ Res Public Health 2020;17(16):5930 [FREE Full text] [doi: 10.3390/ijerph17165930] [Medline: 32824169]
30. Canning CG, Allen NE, Nackaerts E, Paul SS, Nieuwboer A, Gilat M. Virtual reality in research and rehabilitation of gait and balance in Parkinson disease. Nat Rev Neurol 2020;16(8):409-425. [doi: 10.1038/s41582-020-0370-2] [Medline: 32591756]
31. McNaney R, Balaam M, Holden A, Schofield G, Jackson D, Webster M, et al. Designing for and with people with Parkinson's: a focus on exergaming. 2015 Presented at: CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; 2015 April 18-23; Seoul Republic of Korea p. 501-510. [doi: 10.1145/2702123.2702310]
32. McNaney R, Lindsay S, Ladha K, Ladha C, Schofield G, Plötz T. Cueing swallowing in Parkinson's disease. 2011 Presented at: CHI 2011 - 29th Annual CHI Conference on Human Factors in Computing Systems, Conference Proceedings and Extended Abstracts; 2011 May 7-12; Vancouver BC Canada p. 619-622. [doi: 10.1145/1978942.1979030]

Abbreviations

PD: Parkinson disease
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by T de Azevedo Cardoso, T Leung; submitted 06.10.23; peer-reviewed by L Wang; comments to author 30.11.23; revised version received 14.12.23; accepted 06.06.24; published 09.09.24.

Please cite as:
Bevilacqua R, Benadduci M, Barbarossa F, Amabili G, Di Donna V, Martella C, Pelliccioni G, Riccardi GR, Maranesi E
Effectiveness of Technological Interventions for Older Adults With Parkinson Disease: Systematic Review
JMIR Serious Games 2024;12:e53431
URL: https://games.jmir.org/2024/1/e53431
doi:10.2196/53431
PMID:39250193

Review

Game-Based eHealth Interventions for the Reduction of Fatigue in People With Chronic Diseases: Systematic Review and Meta-Analysis

Leonie S Warlo[1,2*], MSc; Souraya El Bardai[1,2*], MSc; Andrica de Vries[2,3*], MD, PhD; Marie-Lise van Veelen[4*], MD, PhD; Suzan Moors[5*], BSc; Edmond HHM Rings[6*], MD, PhD; Jeroen S Legerstee[1,7,8*], PhD; Bram Dierckx[1*], MD, PhD

[1]Department of Child and Adolescent Psychiatry/Psychology, Sophia Children's Hospital, Erasmus Medical Center, Rotterdam, Netherlands
[2]Department of Pediatric Oncology, Sophia Children's Hospital, Erasmus Medical Center, Rotterdam, Netherlands
[3]Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands
[4]Department of Neurosurgery, Sophia Children's Hospital, Erasmus Medical Center, Rotterdam, Netherlands
[5]Department of Physiotherapy, Sophia Children's Hospital, Erasmus Medical Center, Rotterdam, Netherlands
[6]Department of Pediatrics, Sophia Children's Hospital, Erasmus Medical Center, Rotterdam, Netherlands
[7]Levvel, Academic Center for Child and Adolescent Psychiatry and Specialized Youth Care, Amsterdam, Netherlands
[8]Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, Netherlands
[*]all authors contributed equally

Corresponding Author:
Souraya El Bardai, MSc
Department of Child and Adolescent Psychiatry/Psychology
Sophia Children's Hospital
Erasmus Medical Center
Wytemaweg 80
Rotterdam, 3015 CN
Netherlands
Phone: 31 630403394
Email: s.elbardai@erasmusmc.nl

Abstract

Background: Fatigue is a common and debilitating side effect of chronic diseases, significantly impacting patients' quality of life. While physical exercise and psychological treatments have been shown to reduce fatigue, patients often struggle with adherence to these interventions in clinical practice. Game-based eHealth interventions are believed to address adherence issues by making the intervention more accessible and engaging.

Objective: This study aims to compile empirical evidence on game-based eHealth interventions for fatigue in individuals with chronic diseases and to evaluate their effectiveness in alleviating fatigue.

Methods: A comprehensive literature search was performed across Embase, MEDLINE ALL, PsycINFO, Web of Science Core Collection, Cochrane Central Register of Controlled Trials, and Google Scholar in August 2021. Study characteristics and outcomes from the included studies were extracted, and a random-effects meta-analysis was conducted. Sensitivity and subgroup analyses were performed to identify sources of heterogeneity.

Results: Of 1742 studies identified, 17 were included in the meta-analysis. These studies covered 5 different chronic diseases: multiple sclerosis (n=10), cancer (n=3), renal disease (n=2), stroke (n=1), and Parkinson disease (n=1). All but 1 study used exergaming interventions. The meta-analysis revealed a significant moderate effect size in reducing fatigue favoring the experimental interventions (standardized mean difference [SMD] –0.65, 95% CI –1.09 to –0.21, *P*=.003) compared with control conditions consisting of conventional care and no care. However, heterogeneity was high (I2=85.87%). Subgroup analyses were conducted for the 2 most prevalent diseases. The effect size for the multiple sclerosis subgroup showed a trend in favor of eHealth interventions (SMD –0.47, 95% CI –0.95 to 0.01, *P*=.05, I2=63.10%), but was not significant for the cancer group (SMD 0.61, 95% CI –0.36 to 1.58, *P*=.22). Balance exercises appeared particularly effective in reducing fatigue (SMD –1.19, 95% CI –1.95 to –0.42, *P*=.002).

Conclusions: Game-based eHealth interventions appear effective in reducing fatigue in individuals with chronic diseases. Further research is needed to reinforce these findings and explore their impact on specific diseases. Additionally, there is a lack of investigation into interventions beyond exergaming within the field of game-based learning.

(JMIR Serious Games 2024;12:e55034) doi:10.2196/55034

KEYWORDS

fatigue; chronic disease; eHealth; serious games; exergames

Introduction

Chronic diseases are a major cause of morbidity worldwide, with their prevalence steadily increasing due to a growing and aging population, improved disease detection, and advancements in medical treatments, leading to greater longevity [1,2]. Chronic diseases are conditions that persist over a long period or recur frequently, often requiring ongoing medical attention [3,4]. The management of chronic diseases is shifting from cure to care and prevention strategies, with a particular focus on lifestyle management [5]. In addition to the impact of the chronic disease itself, several commonly associated symptoms—such as depression, anxiety, and fatigue—affect quality of life and should be included in routine care [6-8].

Fatigue is one of the most prevalent of these symptoms [8]. It is defined as an overwhelming sense of tiredness and exhaustion that arises without provocation and cannot be relieved by rest [9,10]. Connolly et al [8] found that patients often report fatigue as one of the most debilitating symptoms, significantly impacting daily functioning and quality of life. They report that fatigue occurs across a range of chronic diseases, including multiple sclerosis (MS) and cancer. A recent meta-analysis evaluated the prevalence of severe and chronic fatigue in a cohort of individuals with chronic diseases, finding that 23% experienced severe fatigue and 17% suffered from chronic fatigue [11].

Over the past decades, nonpharmacological treatments for fatigue have been increasingly developed and investigated. Meta-analyses indicate that physical exercise can reduce fatigue severity across various chronic diseases [12-17]. Other successful interventions include psychological therapies—such as cognitive-behavioral therapy, psychoeducation, or mindfulness—whether as standalone treatments or in combination with exercise, as well as relaxation therapies [18,19].

Despite these findings, patients often struggle with adhering to interventions [20,21]. Evidence indicates that the reasons for nonadherence are diverse, including barriers such as time, costs, location, comorbidities, and particularly a lack of motivation [21-25]. To be successful, interventions should be designed to address and overcome these barriers to adherence. This is the goal of game-based interventions, which aim to make treatment easily accessible and highly engaging. This is one reason why such interventions have become increasingly popular in recent years. Games are known to enhance motivation, attention, and learning, among other benefits [26]. Game-based interventions leverage these benefits by embedding therapeutic goals within a game (serious gaming). Evidence indicates that these interventions can significantly improve treatment adherence in chronic conditions compared with standard care [27]. Additionally, from a financial perspective, game-based interventions are attractive to health care providers and insurance companies due to their cost-effectiveness [28-30].

Several studies have investigated the effects of game-based eHealth interventions in individuals with chronic diseases, yielding promising results across a range of outcomes. These interventions include exergames (ie, game-based exercise programs [31]), virtual reality (VR) tools [32], and serious game applications [33]. For example, Kato et al [34] investigated the effect of a serious game designed to improve adherence and other behavioral outcomes in children with cancer, finding that it successfully enhanced medication adherence and self-efficacy in the target group. In a study by Del Corral et al [35], exergaming was found to lead to significantly greater improvements in exercise capacity, muscular strength, and quality of life in children with cystic fibrosis compared with the control group receiving conventional care.

With the accumulation of numerous studies over the past decade, evidence in this field has been synthesized in meta-analyses. Rutkowski et al [36] found that VR interventions appear to be effective in alleviating fatigue in individuals with cancer. Cugusi et al [37] reported small but significant effect sizes for improving health-related quality of life with experimental exergaming interventions in people with various chronic diseases, including Parkinson disease, Alzheimer disease, and stroke. Seiler et al [38] also found promising effects of various types of eHealth interventions in reducing fatigue in individuals with cancer.

However, to date, no meta-analysis has investigated the effects of (1) different game-based eHealth interventions on (2) the reduction of fatigue in (3) individuals with various chronic conditions.

This paper aims to fill this gap by systematically aggregating the findings from these studies to assess the effectiveness of game-based eHealth interventions in alleviating fatigue. The goal is to determine whether these interventions can serve as a suitable alternative to conventional treatments.

Methods

Selection Criteria

We included randomized and nonrandomized controlled trials that reported the effects of video game interventions on fatigue in individuals with chronic diseases. For this study, we defined a video game as a digital or electronic game where players interact with the game by manipulating images on a video screen. A "game" was defined as an engaging, amusing, and structured form of play conducted according to a set of rules with the aim of achieving a specific objective. Chronic diseases are defined as conditions that persist or recur over an extended period and require ongoing medical attention or limit activities of daily living [3,4]. We focused on pathological fatigue, defined as physical, emotional, or mental tiredness/exhaustion related to chronic disease or its treatment [39]. This type of fatigue is characterized by its prolonged, severe, progressive nature or its occurrence without provocation. For practical reasons, we included only journal articles published in English. All studies had to include a T1 measure with a measure of change from the baseline and a control group from the same disease population receiving a different or no intervention.

We excluded trials involving healthy volunteers, individuals with acute diseases, and those with fibromyalgia. The clinical population with fibromyalgia was excluded due to its high heterogeneity, unclear etiology, and purely clinical diagnosis, as there are no specific laboratory abnormalities associated with it [40]. Therefore, fibromyalgia is unsuitable for this meta-analysis due to its heterogeneous nature and unclear etiology, making it difficult to detect the group effects of an intervention. Articles focusing on different types of fatigue, such as fatigue after exertion or transient fatigue, were also excluded. Additionally, we excluded reviews, descriptive and observational studies, study protocols, case studies, uncontrolled studies, conference abstracts, trial registries, posters, and books, as well as studies that used nonstandardized measuring scales for fatigue.

Search Strategy

A medical information specialist from the Erasmus MC Medical Library conducted a comprehensive literature search on August 25, 2021. To ensure the findings were up-to-date, a second search was carried out on March 2, 2023. Both searches utilized the following databases: Embase, MEDLINE ALL, PsycINFO, Web of Science Core Collection, Cochrane Central Register of Controlled Trials, and Google Scholar. For both searches, the coverage years varied by database (Multimedia Appendix 1). Nonetheless, the majority of articles were published within the last 3 decades.

The search terms "game," "video," "fatigue," and related keywords were used. A separate search strategy was developed for each database (Multimedia Appendix 1). We did not include "chronic disease" or specific diseases as search terms, as we deemed the risk of missing relevant studies due to incomplete disease terms to be too high.

Selection Procedure

After removing duplicates, the titles and abstracts were screened by 2 of the authors (LSW and JSL). The full-text papers were then extracted and screened by the same authors along with an additional author (BD). The selection of articles was compared among the authors at all stages of the process. In cases of disagreement, the articles were discussed until a consensus was reached. The authors of the papers were contacted by email when relevant information was missing or inconsistent. Articles were excluded if the authors did not respond.

Assessment of Study Quality

For each study, the risk of bias was assessed by the author LSW using the risk of bias 2 tool (Cochrane Risk of Bias Tool for Randomized Trials) [41]. The assessment covered the following categories: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of reported results. Based on the assessment of the individual categories, each study was classified into 1 of 3 overall risk of bias levels: "low risk of bias," "some concerns," or "high risk of bias." The overall classification was determined by the lowest rating among the individual categories (eg, if 4 of the 5 categories were rated "low risk of bias" but 1 was rated "high risk of bias," the overall classification would be "high risk of bias").

Data Extraction

After the selection process, one author (LSW) performed data extraction for each article. The extracted data included diagnosis, author, year of publication, sample size, mean age, percentage of female participants, interventions in both the experimental and control groups, duration of the interventions in weeks, and key findings. All data were entered into Comprehensive Meta-Analysis (CMA) software version 2 (Biostat, Inc.).

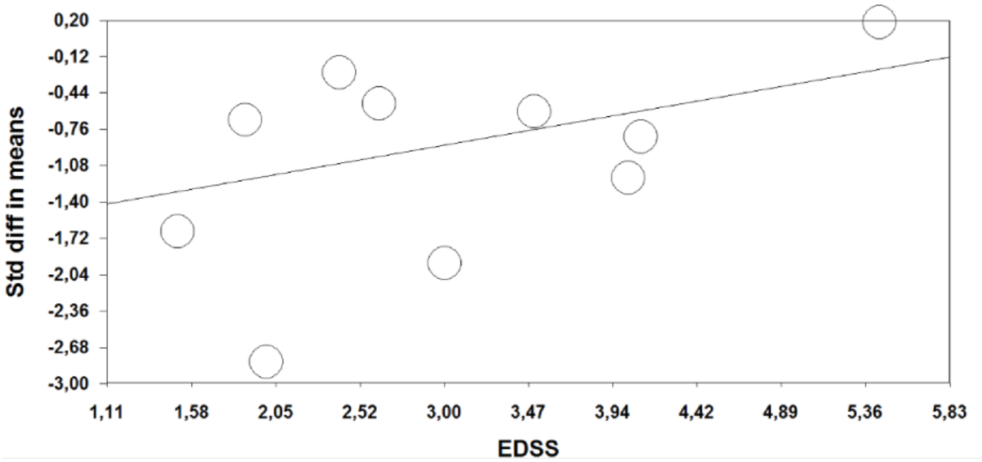
Synthesis of Results

The analyses were conducted using CMA [42]. For our outcome variable, fatigue, the mean scores and SDs for pre- and postintervention (ie, baseline and T1) were either extracted directly from the articles or calculated from median scores and IQRs using the formula described by Wan et al [43]. We followed the guidelines in the Cochrane Handbook for Systematic Reviews of Interventions to calculate the pre-post correlation [44]. It was calculated directly for studies where the SD of change from baseline to T1 was available. For other studies, we imputed the correlation by averaging the calculated pre-post correlations. Additionally, we entered data on sample size per condition, diagnosis, assessment instrument (Visual Analog Scale [VAS] vs questionnaire), mean age, percentage of female participants, modality and type of intervention (nonimmersive, immersive VR, non–VR game; balance, fitness, cognition), intervention duration in weeks, type of control intervention (no care vs conventional care), setting (hospital vs home), supervision, and, where possible, disease severity into the CMA worksheet. Two of the included studies were crossover randomized controlled trials (the remainder were parallel randomized controlled trials). For 1 of the crossover studies, we used only the T1 measure for comparison, which included data only from the period before the crossover [45]. For the other crossover study, we used data from both periods combined (ie, before and after the crossover) because only these combined data were available [46]. In studies measuring different dimensions of fatigue, the dimension reporting the average fatigue measure was used. In studies with 2 control groups—1 receiving a conventional intervention and 1 with no intervention—we chose the inactive control group. In a study comparing 2 interventions using different VR systems with a control group, the aggregated mean of the 2 experimental conditions was used [47]. For studies with 3 measurement points, data from the measures immediately before and after the intervention were utilized.

First, effect sizes were calculated as standardized mean differences (SMDs) to account for possible differences in measurement scales. We conducted a meta-analysis to determine the overall effect sizes for the experimental condition compared with the control condition using a random-effects model. The more conservative random-effects model was chosen over the fixed-effects model due to the expected heterogeneity among studies and because random-effects models are recommended for analyzing data collected in real-world settings rather than controlled laboratory environments [48]. Heterogeneity was estimated using the I^2 index, which describes the percentage of variation attributable to study heterogeneity rather than chance [49], with ≥75% indicating considerable heterogeneity. To explore sources of heterogeneity, we performed sensitivity analyses (by excluding low-quality studies and outliers), moderator analyses, and meta-regressions. Low-quality studies were defined as those with a high risk of bias, as identified by the risk of bias assessment. Outliers were defined as studies where the 95% CI did not overlap with the 95% CI of the pooled effect size. For the moderator analyses, studies were grouped by diagnosis, age, and type of experimental and control interventions, provided there was more than 1 study per group. Additionally, we conducted analyses excluding studies using VAS, exploring the impact of supervision, and distinguishing between studies conducted at home versus those conducted in a hospital setting. A random-effects meta-regression using the method of moments was conducted with gender and disease duration as predictors. For studies on MS, we additionally performed a meta-regression with disease severity, which was measured using the Expanded Disability Status Scale (EDSS) [50], as a predictor. This is shown in Figure 1.

Finally, to check for publication bias, we generated a funnel plot by plotting the SMD against the SE of all studies and assessed it for asymmetry. Additionally, we quantified potential publication bias statistically using the Egger test of the intercept [51].

Figure 1. Scatterplot showing the meta-regression of all multiple sclerosis studies with the Expanded Disability Status Scale (EDSS) as the predictor variable.

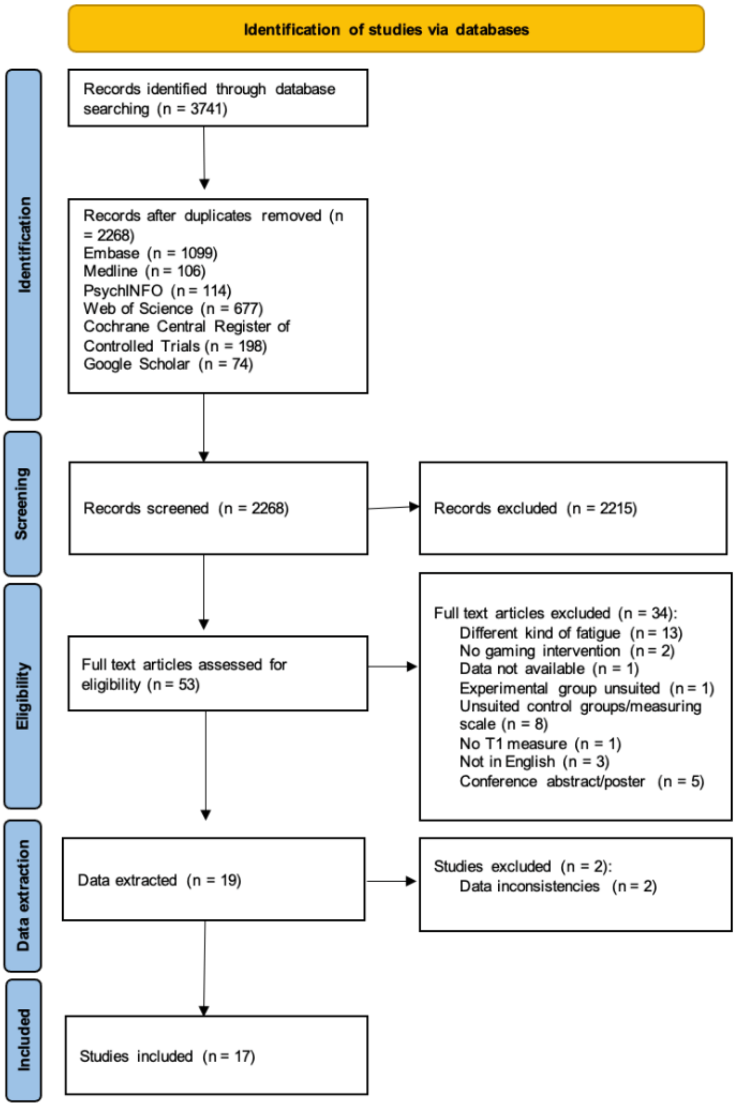


Results

Study Selection

A total of 3741 articles were identified through the literature search, of which 2268 remained after duplicate removal. An overview of the selection process is shown in Figure 2. After screening titles and abstracts, and discussing differences in opinion between the authors (for 14 articles), a total of 53 articles were selected for full-text screening. After independently screening the full text of these studies, the authors discussed discrepancies in study selection for 8 studies until a consensus was reached. Authors were contacted for missing information in 2 cases. We received a reply for 1 article, which led to its exclusion. The other study was excluded due to the lack of a response. In total, 19 articles were excluded after full-text review. Most were excluded because the fatigue measure pertained to exertion from the intervention itself. Relevant data were then extracted from the remaining 19 studies. During this process, 2 additional studies were excluded due to data inconsistencies; we contacted the authors but did not receive a reply. This left us with a total of 17 studies included in the data synthesis for the meta-analysis.

Figure 2. Preferred Reporting Items for Systematic Reviews (PRISMA) flowchart-diagram for study selection.



Characteristics of the Included Studies

Table 1 presents the characteristics of the included studies, organized by participant diagnosis: MS (n=10) [45,47,52-59], cancer (n=3) [46,60,61], renal disease (n=2) [62,63], Parkinson disease (n=1) [64], and poststroke (n=1) [65]. The studies were published between 2013 and 2023 and were all randomized controlled trials, including 2 crossover trials [45,46] and 15 parallel trials [47,52-65]. The number of participants ranged from 20 to 52. The mean age of participants ranged from 7.9 to 68.7 years, with only 1 study [60] including children. For the 16 studies that included adults, the age range was 32.3-68.7 years (median 45 years) [45-47,52-59,61-65]. The mean percentage of female participants ranged from 0% (0/42) to 90% (38/42; median 61.50%). Sixteen studies used VR exergames, which included either balance exercises (n=5) [47,52,55,57,65] or fitness exercises (n=11) [45,46,53,56,58-64]. One study used a serious Nintendo DS game designed to train cognitive functions such as working memory, spatial recognition, processing speed, and mental reasoning in healthy individuals [54]. Thus, the term "game-based eHealth interventions" is technically too broad. Throughout this paper, we used this term to include the serious gaming study as well. The VR technology varied across studies, with most using nonimmersive VR systems (n=14) [45-47,52,53,55,56,58-64]. The control interventions also exhibited some heterogeneity across studies. We categorized the control interventions into 2 groups: the conventional training group (n=7) [46,52,53,55,59,64,65], which involved traditional nonvideo game interventions targeting the same abilities as the experimental intervention, and the no-exercise control group (n=10) [45,47,54,56-58,60-63], which involved no specific intervention beyond the normal level of activity. The duration of the interventions varied significantly between studies, ranging from 2 to 24 weeks (median 8 weeks). Fatigue was measured using various questionnaires, with 1 study using a VAS to assess fatigue severity [62].

Table 1. Study characteristics and key findings of the included studies that reported on the effect of game-based eHealth interventions on the reduction of fatigue in people with chronic diseases. Studies are sorted according to diagnosis.

Diagnosis and study	Participants			Intervention			Key findings
	N	Age (years), mean	Female, n/N (%)	Experimental	Control	Duration (weeks)	
Multiple sclerosis							
Brichetto et al [52]	36	42	22/36 (61)	Balance, nonimmersive VR[a]	Conventional training	12	Significant decrease in fatigue (MFIS[b]) after treatment compared with the baseline and control groups
Cuesta-Gómez et al [53]	30	46.3	18/30 (60)	Fitness, nonimmersive VR	Conventional training	10	No significant decrease in fatigue after treatment (FSS[c])
De Giglio et al [54]	35	43.8	26/35 (74)	Cognition, Gameboy	No exercise	8	No significant decrease in fatigue after treatment (MFIS)
Khalil et al [55]	32	37.4	22/32 (69)	Balance, nonimmersive VR	Conventional training	6	Significant decrease in fatigue (MFIS) after treatment compared with the control group
Ozdogar et al [56]	59	40.2	43/59 (73)	Fitness, nonimmersive VR	No exercise	8	No significant decrease in fatigue in the experimental and control groups (MFIS)
Ozdogar et al [59]	30	37.6	21/31 (68)	Fitness, nonimmersive VR	Conventional rehabilitation	6	Significant decrease in fatigue (MFIS) in the experimental group compared with baseline and the control group
Ozdogar et al [58][d]	31	40.3	20/31 (65)	Fitness, nonimmersive VR	No exercise	8	Significant increase in sleep quality in the experimental group compared with baseline and the control group
Ozdogar et al [58][e]	34	40.8	21/34 (62)	Fitness, nonimmersive VR	No exercise	8	No significant increase in sleep quality in the experimental and control groups
Ozkul et al [57][e]	39	32.3	30/39 (77)	Balance, immersive VR	No exercise	8	Significant decrease in fatigue (FSS) after treatment compared with baseline and the control group
Thomas et al [45][e]	29	49.3	27/30 (90)	Fitness, nonimmersive VR	No exercise	24	Increased fatigue after treatment in both experimental and control groups (FSI[f])
Yazgan et al [47][e]	42	43.7	38/42 (90)	Balance, nonimmersive VR	No exercise	8	Significant decrease in fatigue (FSS) after treatment compared with baseline and the control group
Cancer							

Diagnosis and study	Participants			Intervention			Key findings
	N	Age (years), mean	Female, n/N (%)	Experimental	Control	Duration (weeks)	
Hamari et al [60]	36	7.8	10/36 (28)	Fitness, nonimmersive VR	No exercise	8	Change in fatigue was similar in both groups (PedsQL[g])
Kobayashi et al [46]	22	44.8	4/22 (18)	Fitness, nonimmersive VR	Conventional training	2	Increase in fatigue after the intervention in the experimental group and a significant decrease in the control group (POMS-sf Fatigue[h])
Villumsen et al [61]	46	68.7	0/46 (0)	Fitness, nonimmersive VR	No exercise	12	No significant decrease and no significant difference in change in fatigue between groups (FACT-F[i])
Renal disease							
Cho and Sohng [62]	48	59.3	20/48 (42)	Fitness, nonimmersive VR	No exercise	8	Significant decrease in fatigue in the experimental group but not in the control group (VAS[j])
Chou et al [63]	64	59.3	28/64 (44)	Fitness, nonimmersive VR	No exercise	4	Significant decrease in fatigue in both groups (NFSHD[k]); no significant difference between groups
Parkinson disease							
Ribas et al [64]	20	61	8/20 (40)	Fitness, nonimmersive VR	Conventional training	12	Significant decrease in fatigue in the experimental group but not in the control group (FSS)
Stroke							
de Rooij et al [65]	52	63	16/52 (31)	Balance, immersive VR	Conventional training	6	No significant decrease and no significant difference in change of fatigue between groups (FSS)

[a]VR: virtual reality.

[b]MFIS: Modified Fatigue Impact Scale.

[c]FSS: Fatigue Severity Scale.

[d]With restless legs syndrome.

[e]Without restless legs syndrome.

[f]FSI: Fatigue Symptom Inventory.

[g]PedsQL: Pediatric Quality of Life Inventory.

[h]POMS-sf: Profile of Mood States—short form.

[i]FACT-F: Functional Assessment of Cancer Therapy—Fatigue.

[j]VAS: Visual Analog Scale.

[k]NFSHD: Novel Fatigue Scale for Hemodialysis.

Risk of Bias of Studies

Study quality was low or moderate in all studies (n=5 and n=12, respectively), with none rated as high quality (Figure 3). The primary reasons for low-quality ratings were issues with the randomization process and handling of missing data.

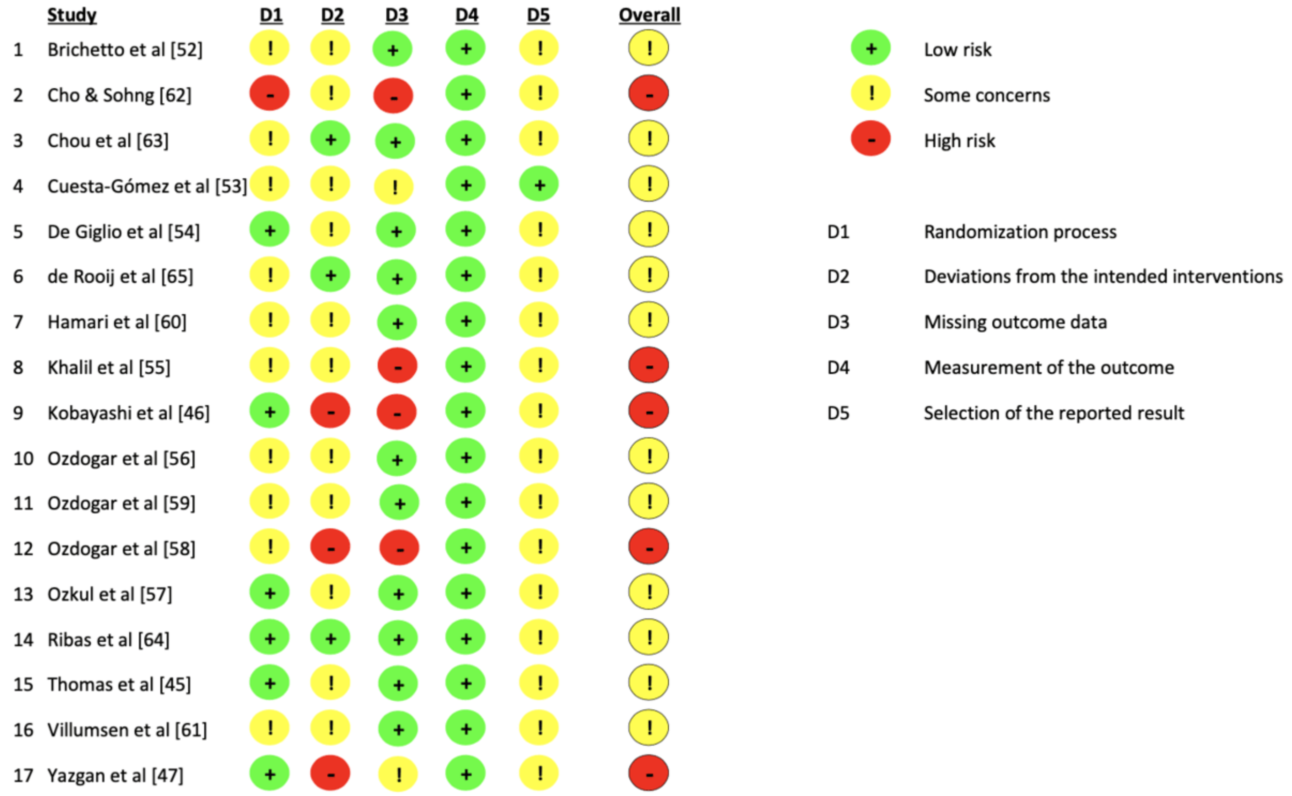
While all studies were randomized, 2 utilized a cluster-randomization procedure, with treatment allocation based on either the days participants visited the hospital [62] or the hospital wards to which they were assigned [63]. For Chou et al [63], we can assume that participant allocation was concealed from the investigator; however, this was not clear for Cho and Sohng [62]. For the other studies, the randomization process was truly random. However, there was some concern about the risk of bias in 9 of the 17 (53%) studies [52,53,55,56,59-61,63,65] due to missing or doubtful information about allocation concealment.

In 10 of the 17 (59%) studies [45-47,53-57,60,62], there was concern about the risk of bias because the authors did not implement an intention-to-treat analysis to account for missing or lost data. However, the missing data were either balanced across studies or not substantial enough to significantly impact the results. By contrast, 2 (12%) studies [46,47] exhibited a high risk of bias due to substantial issues with missing data.

In 8 of the 17 (47%) studies [46,47,53,55,57,58,60,62], it could not be ruled out that the missing data were related to the outcome itself (ie, fatigue), potentially influencing the overall results. Possible reasons related to the outcome were lack of motivation or excessive fatigue preventing participation. Issues related to participants' schedules or travel time were judged as unrelated to the outcome. In 4 of these studies [46,55,58,62], a high risk of bias was concluded as it was likely that the missing data depended on the true value of the outcome. In the other 4 studies [47,53,57,60], there were some concerns about the risk of bias, but the proportion of missing data and reasons for it were balanced across groups.

The risk of bias in the measurement of outcomes was low across all studies. However, there was some concern regarding the selection procedure of reported results in 16 of the 17 (94%) studies [45-47,52-61,63-65]; specifically, 16 (94%) studies did not indicate whether the analysis followed a prespecified plan [45-47,52,54-65]. Additionally, 4 (24%) studies began before the trial had been preregistered.

Figure 3. Risk of bias assessment for all included studies.



Main Analyses

It was possible to calculate effect sizes for fatigue reduction across all studies. Figure 4 presents the SMD and 95% CIs for each study. A negative effect size indicates that the experimental intervention was more effective in reducing fatigue compared with the control intervention. In 4 studies [45,46,53,61], the effect sizes were positive, meaning that the control intervention was more effective in reducing fatigue than the experimental intervention. Three studies were identified as outliers, 2 [46,62] of which were also rated as low quality. The third outlier was the study by Villumsen et al [61], which also reported no reduction in fatigue following the intervention. This lack of effect may be related to the study's population, which consisted entirely of males with a mean age of 68.7 years, the oldest among all the studies. The authors suggest that the lack of supervision over the exercise intensity in the home-based intervention might explain the findings. The overall effect size, calculated using a random-effects model, indicated a significant moderate effect of video game interventions on fatigue reduction compared with control interventions (SMD –0.65, 95% CI –1.09 to –0.21, P=.003). However, there was considerable heterogeneity (I^2=85.87%). To investigate the sources of this heterogeneity, several additional analyses were conducted.

First, we performed sensitivity analyses by excluding the low-quality and outlier studies. This included 4 studies [46,47,55,62] with a high risk of bias and 1 additional outlier [61], leaving us with 10 studies. Despite this more rigorous analysis, the effect size remained significant (SMD –0.42, 95% CI –0.74 to –0.10, P=.01). Although heterogeneity was reduced, it remained substantial (I^2=54.88%). When the study using a VAS was excluded (n=1) [62], the effect size was smaller but still significant (SMD –0.55, 95% CI –0.95 to –0.14, P=.009).

Second, we performed moderator analyses with diagnosis, type of intervention, type of control condition, and age as moderators for all groups where data from more than 1 study were available. Grouping studies according to diagnosis revealed a large and significant effect size for MS (SMD –0.87, 95% CI –1.34 to –0.41, P<.001, n=10). After removing low-quality and outlier studies, the effect size decreased to a trend (SMD –0.47, 95% CI –0.95 to 0.01, P=.05, I^2=63.10%, n=6). For cancer, the pooled effect size was positive, indicating that the control intervention was more effective than the experimental intervention, but this effect size was not statistically significant (SMD 0.61, 95% CI –0.36 to 1.58, P=.22, n=3). For both MS and cancer, heterogeneity was reduced but remained substantial (I^2=77.9% for MS and I^2=72.33% for cancer). For renal disease, with only 2 studies available [62,63], the pooled effect size was –1.13 (95% CI –2.32 to 0.05, P=.06), and heterogeneity was notably high (I^2=96.08%).

Interventions involving balance exercises (n=5) showed a large effect size of –1.19 (95% CI –1.95 to –0.42, P=.002). By contrast, for fitness interventions (n=10), the effect size was nonsignificant (SMD –0.44, 95% CI –1.02 to 0.13, P=.20). Heterogeneity remained substantial for both categories (I^2=62.78% for balance exercises and I^2=87.15% for fitness interventions). We could not pool effects for cognitive interventions, as only 1 study investigated this category [54]. Sensitivity analysis, which excluded low-quality and outlier studies, confirmed these results while substantially reducing heterogeneity (I^2=58.83% with n=3 for the balance group and I^2=23.60% with n=6 for the fitness group). These findings suggest that game-based balance exercises, in particular, are effective interventions for reducing fatigue in individuals with chronic diseases.

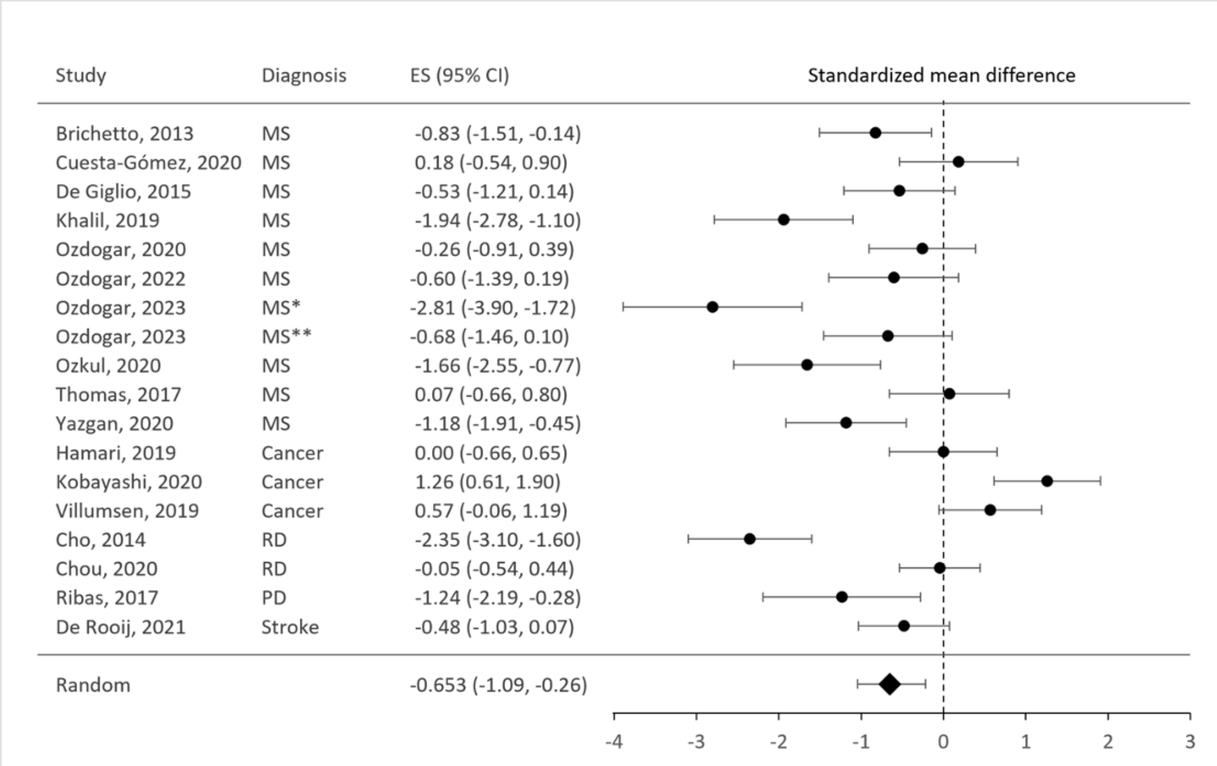
Grouping studies by the type of control group used revealed nonsignificant effect sizes for the conventional training control groups (SMD –0.49, 95% CI –1.12 to 0.15, P=.20, I^2=86.7%, n=7), but significant effect sizes for the no-exercise control groups (SMD –0.75, 95% CI –1.33 to –0.18, P=.01, I^2=85.6%, n=10).

Significant differences were found when comparing participants with a mean age below 55 years (SMD –0.65, 95% CI –1.16 to –0.13, P=.02, I^2=84%, n=12), but no significant difference was observed for those above 55 years (SMD –0.68, 95% CI –1.59 to 0.23, P=.15, I^2=90%, n=5).

Third, we conducted a meta-regression using the method of moments in a random-effects model to estimate the effect of gender and duration of intervention on the impact of game-based interventions. Neither gender nor duration significantly influenced the effect size for fatigue reduction (P=.08 and P=.86, respectively). However, the presence of supervision and the location of the studies have significantly influenced the effect size. For both factors, a meta-regression was conducted using a random-effects model. Supervision significantly affected the effect size, with supervised interventions showing a significant effect (SMD –0.86, 95% CI –1.39 to –0.33, P=.001), whereas interventions without supervision showed no significant effect (SMD 0.04, 95% CI –0.42 to 0.49, P=.88). The meta-regression also revealed a significant effect for studies conducted in a hospital setting (SMD –0.79, 95% CI –1.30 to –0.30, P=.002), contrasting with those conducted at home, which showed no significant effect (SMD 0.04, 95% CI –0.61 to 0.69, P=.90).

We performed an additional meta-regression with disease severity as a predictor for all the MS studies where it was reported (n=9; not reported in n=1) [45,47,52-59]. Increased severity was associated with a smaller effect of the intervention (SMD –0.27, 95% CI 0.06-0.48, P=.01). When high risk of bias studies were excluded, the association remained, although slightly weaker (SMD 0.25, 95% CI 0.02-0.49, P=.04, n=5) [46,47,55,58,62].

Figure 4. Random-effect meta-analysis for the effect of serious gaming on fatigue. ES: effect size; MS: multiple sclerosis; PD: Parkinson disease; RD: renal disease. *With restless leg syndrome; **without restless leg syndrome.



Publication Bias

The funnel plots for all included studies and the MS subgroup displayed asymmetry [45,47,52-59]. The Egger test of the intercept confirmed that this asymmetry was statistically significant, with *P* values of .006 and .004, respectively, indicating evidence of publication bias. This suggests an overrepresentation of studies with positive results, which should be considered when interpreting the findings. However, after excluding studies with a high risk of bias, the observed asymmetry was no longer statistically significant (*P*=.25) [46,47,55,58,62].

Discussion

Principal Findings

In recent years, the digitalization and gamification of interventions have garnered increasing attention as alternatives or complements to conventional treatment approaches. This paper aimed to evaluate the efficacy of game-based eHealth interventions in reducing fatigue among individuals with chronic diseases. We included 17 randomized controlled trials published between 2013 and 2023, encompassing 5 different types of chronic diseases. The relative recency of publications and the small number of studies illustrate that the field of (game-based) eHealth is still in its infancy. The types of interventions were fairly homogeneous, with all but 1 study focusing on exergaming interventions [54]. The remaining studies evaluated a serious game aimed at improving cognition. This trend, although on a smaller scale, mirrors the evidence base for conventional rehabilitation approaches for fatigue, where the majority of studies also focus on physical exercise interventions. However, given the positive findings for psychological interventions, particularly when combined with exercise interventions [18], future eHealth interventions should also explore these approaches.

Findings from this meta-analysis suggest that current game-based eHealth interventions may effectively reduce fatigue in people with chronic diseases. With a moderate effect size, these interventions could potentially be more effective for fatigue compared with other treatment goals, such as knowledge and self-management [66], self-efficacy [67], and health-related quality of life [37], where previous meta-analyses reported smaller effect sizes. Additionally, they appear to be as effective as many conventional (non–game-based) interventions, which typically report moderate effect sizes [18]. Some meta-analyses investigating the effect of exercise therapy on disease-related fatigue (such as in cancer and chronic obstructive pulmonary disease) reported larger effect sizes for physical exercise therapies [13,15], a finding also supported by individual studies in this meta-analysis [47,52,55,57,62,64]. The comparable effectiveness of game-based eHealth interventions is crucial for them to become a viable alternative to conventional interventions. Thus, this result represents an important first step in exploring the potential of game-based eHealth interventions.

With regard to individual chronic diseases, findings from this meta-analysis were less straightforward. Game-based interventions appear effective for MS, but not for cancer. Cancer is a heterogeneous disease with variable cancer-related fatigue. The underlying pathophysiology is relatively well investigated and is likely multifactorial, involving inflammation, disruptions in the hypothalamic–pituitary–adrenal axis, and activation of the autonomic nervous system [68,69]. However, it is influenced by several factors including the type of cancer, the stage of the disease, and the treatment—all of which varied across and within the study populations of the included cancer studies. The considerable heterogeneity of the cancer group, including a wide mean age range from 8 to 69 years, might explain the lack of a treatment effect in this group. By contrast, MS typically presents with a more homogeneous course, commonly consisting of exacerbations and stable phases [70], and all of the MS studies included here focused on patients in a stable phase. As the onset of MS typically occurs between 20 and 40 years of age [71], the study population for MS was more homogeneous in terms of age, ranging from 31 to 49 years. Besides disease-related differences, variations in results might be attributed to statistical power issues due to the limited number of studies [72]. This limitation increases the likelihood of fluctuations due to chance, particularly in the cancer group, where 1 study was an outlier [61] and another was of low quality [46]. Interestingly, 1 of the cancer studies [46] found that the experimental group experienced an increase in fatigue after the intervention. The authors suggested that this might be due to an inappropriate exercise load, as patients were unable to adjust it according to their needs in the experimental condition. Additionally, "fatigue" might have been interpreted as "exercise load," given that the Profile of Mood States—short form (POMS-sf) measuring scale used in the study assesses "general fatigue" nonspecifically.

Another striking finding of this meta-analysis was the clear difference between balance exercises and fitness exercises. The balance exercises showed a markedly larger effect size in reducing fatigue compared with fitness exercises (SMD –1.19 vs SMD –0.17). It is important to note that the balance exercise group was more homogeneous in terms of patient diagnoses, with only 1 study including patients after stroke and the rest consisting of patients with MS [65]. The fitness group included patients with 4 different diagnoses, which might contribute to the observed heterogeneity and complicate the comparison with the balance group. However, a similar observation is evident within the MS studies: all 4 balance studies favored the experimental intervention [47,52,55,57], whereas only 3 [57-59] of the 6 fitness studies did [45,53,56]. These findings contrast with a recent randomized controlled trial by Callesen et al [73], which reported conventional balance training and exercise training as equally effective in reducing fatigue among patients with MS. However, our results align with Hebert and Corboy [74], who demonstrated a significant relationship between fatigue and balance in patients with MS. Additionally, evidence from healthy participants suggests that balance exercises not only improve balance but also muscle strength [74,75]. This dual benefit might make balance exercises more effective than pure strength exercises in reducing fatigue, as they address both balance and strength—factors associated with fatigue. Additionally, balance exercises might be more enjoyable and less demanding than fitness exercises. It is also worth noting that 2 studies in the balance group incorporated additional interventions: 1 included walking alongside balance exercises [57] and another combined Pilates with balance exercises [65]. Given the small number of studies in the balance group (n=5), the effects of these 2 studies significantly influence the overall effect size for this group. Overall, the observation that balance exercises appear particularly effective in reducing fatigue is intriguing and warrants further investigation in future research. It also underscores the importance of developing tailored treatment programs for fatigue, as the underlying mechanisms may vary between different diseases [76].

In this meta-analysis, only 4 studies utilized tailored interventions specifically designed for rehabilitation [53-55,65], while the remaining 13 studies used off-the-shelf commercial games [45-47,52,56-64]. According to serious game design theory, considering the unique interests and needs of the target group leads to the best outcomes [77,78]. Nonetheless, 6 [47,52,60,62-64] out of the 11 studies using commercial games were successful in alleviating fatigue [45-47,52,54,56,60-64]. Gender did not appear to influence the effectiveness of the interventions, which contrasts with the assumption of game design theory. This suggests that the success of commercial games might stem from their broad appeal, as developers aim to meet the needs of diverse target groups to maximize their reach. Yet again, age, particularly a mean age below 55 years, had a significant effect on the effectiveness of the intervention. The literature presents mixed findings regarding the influence of age and gender on treatment outcomes in eHealth interventions. Some studies report differences attributed to these variables [79,80], while others do not [81,82]. From an economic perspective, it is important to determine whether the costly tailoring of games yields better results compared with conventional or commercial interventions. Further research is needed to address this question.

Limitations and Implications for Future Research

The current meta-analysis has several limitations that should be considered when interpreting the findings. First, a notable limitation is the lack of adherence to open science principles, particularly the absence of preregistration before conducting the research.

Second, evidence for publication bias was found among the studies included in this analysis. This suggests that the findings may not fully represent the true effects due to a potential overrepresentation of studies with positive results [83,84]. However, when studies with a high risk of bias were excluded, the asymmetry was no longer significant, indicating that publication bias was not evident in the remaining studies.

Third, the included studies exhibited substantial heterogeneity concerning the target group, interventions, software used, and intervention duration. Although we utilized a random-effects model to account for this variability, considerable heterogeneity remained in the findings. Our sensitivity and moderator analyses managed to reduce, but not entirely resolve, this heterogeneity. Potential sources of heterogeneity that were not examined are the type of software used for the interventions, whether fatigue was a primary or secondary outcome, and intervention intensity and frequency, rather than just duration. We opted to focus on duration because this information was available for all studies.

Fourth, the overall number of studies was rather limited, covering a small variety of chronic diseases and interventions. This limitation was particularly pronounced for studies involving children, which is concerning given that up to 21% of children with chronic disease experience severe fatigue [85]. The need for effective treatment in this population is as urgent as it is for adults. Additionally, no high-quality studies were available for analysis, as determined by the Cochrane risk of bias assessment tool. However, it is worth noting that this tool has been reported to have relatively low reliability [86] and is considered more conservative compared with other risk of bias assessments [87].

Fifth, adherence to the study protocol and treatment satisfaction were not systematically measured nor compared with conventional active intervention groups. This aspect is crucial for determining whether game-based interventions are indeed more motivating than their conventional counterparts and should be a focus of future studies.

Sixth, on a more technical note, different measurement scales for fatigue were used across studies. One study used the VAS to measure fatigue, which is methodologically suboptimal as it is not specifically developed or validated for fatigue assessment and does not differentiate between various aspects of fatigue. To minimize the impact of this on the results, a second analysis was conducted, excluding the study that used VAS. The result remained significant, although the effect size was smaller. This suggests that while the VAS had a substantial influence on the outcome, it was not the sole contributor, as the significant effect appears robust. Although we attempted to mitigate potential discrepancies by standardizing outcome measures using the SMD, variations in psychometric properties may have influenced the results within the studies themselves. Additionally, we had to impute SDs for 1 study [58], means and SDs for 3 studies [53,57,60], and pre-post correlation for 14 studies [45,46,52-55,58-65]. This introduces a degree of uncertainty to our findings, as the reliability of these estimates is uncertain.

Finally, the findings presented here reflect short-term outcomes. As most of the studies did not include follow-up measures, we are unable to draw any conclusions about the long-term efficacy of game-based eHealth interventions.

Overall, more studies are needed across all age groups and various chronic diseases where fatigue is a side effect, to better determine whether these interventions are suited for each disease. These studies should adhere to rigorous design and methodology, including follow-up measures, to assess long-term treatment effects and the use of an intention-to-treat analysis approach for data analysis. We recommend testing not only commercial games but also developing more tailored and personalized games that allow for the investigation of treatments beyond physical activity. In particular, a combination of psychological interventions and physical activity is warranted [18].

Conclusions

Based on the current meta-analysis, we cannot yet make clear recommendations for the use of eHealth interventions in clinical practice. However, we can cautiously conclude that eHealth interventions are effective in reducing fatigue in chronic diseases. As the number of studies in this field is steadily increasing, we hope to soon be able to back up our findings and extend them to other chronic conditions as well.

Acknowledgments

We thank the Stichting Roparun and the Stichting Vrienden van het Sophia for their funding of this research. We also express our gratitude to Sabrina Meertens-Gunput from the Medical Library of the Erasmus Medical Center Rotterdam, the Netherlands, for her invaluable assistance in conducting the systematic literature search.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature search protocol.
[DOCX File , 19 KB - games_v12i1e55034_app1.docx]

Multimedia Appendix 2
PRISMA checklist.
[DOCX File , 32 KB - games_v12i1e55034_app2.docx]

References

1. van Oostrom SH, Gijsen R, Stirbu I, Korevaar JC, Schellevis FG, Picavet HSJ, et al. Time trends in prevalence of chronic diseases and multimorbidity not only due to aging: data from general practices and health surveys. PLoS One 2016;11(8):e0160264. [doi: 10.1371/journal.pone.0160264] [Medline: 27482903]
2. World Health Organization. Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013-2020. Geneva, Switzerland: World Health Organization; 2013.

3. Bernell S, Howard SW. Use your words carefully: what is a chronic disease? Front Public Health 2016 Aug 02;4:159 [FREE Full text] [doi: 10.3389/fpubh.2016.00159] [Medline: 27532034]
4. National Center for Chronic Disease Prevention and Health Promotion. About chronic diseases. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/chronicdisease/about/index.htm [accessed 2024-09-26]
5. Budreviciute A, Damiati S, Sabir DK, Onder K, Schuller-Goetzburg P, Plakys G, et al. Management and prevention strategies for non-communicable diseases (NCDs) and their risk factors. Front Public Health 2020 Nov 26;8:574111 [FREE Full text] [doi: 10.3389/fpubh.2020.574111] [Medline: 33324597]
6. DeJean D, Giacomini M, Vanstone M, Brundisini F. Patient experiences of depression and anxiety with chronic disease: a systematic review and qualitative meta-synthesis. Ont Health Technol Assess Ser 2013;13(16):1-33 [FREE Full text] [Medline: 24228079]
7. Pitman A, Suleman S, Hyde N, Hodgkiss A. Depression and anxiety in patients with cancer. BMJ 2018 Apr 25;361:k1415 [FREE Full text] [doi: 10.1136/bmj.k1415] [Medline: 29695476]
8. Connolly D, O'Toole L, Redmond P, Smith SM. Managing fatigue in patients with chronic conditions in primary care. Fam Pract 2013 Apr 21;30(2):123-124. [doi: 10.1093/fampra/cmt005] [Medline: 23520365]
9. Ream E, Richardson A. Fatigue: a concept analysis. Int J Nurs Stud 1996 Oct;33(5):519-529. [doi: 10.1016/0020-7489(96)00004-1] [Medline: 8886902]
10. Swain MG. Fatigue in chronic disease. Clinical Science 2000;99(1):1-8. [doi: 10.1042/cs19990372]
11. Goërtz YMJ, Braamse AMJ, Spruit MA, Janssen DJA, Ebadi Z, Van Herck M, et al. Fatigue in patients with chronic disease: results from the population-based Lifelines Cohort Study. Sci Rep 2021 Oct 25;11(1):20977 [FREE Full text] [doi: 10.1038/s41598-021-00337-z] [Medline: 34697347]
12. Heine M, van de Port I, Rietberg MB, van Wegen EEH, Kwakkel G. Exercise therapy for fatigue in multiple sclerosis. Cochrane Database Syst Rev 2015 Sep 11;2015(9):CD009956 [FREE Full text] [doi: 10.1002/14651858.CD009956.pub2] [Medline: 26358158]
13. Kessels E, Husson O, Van der Feltz-Cornelis CM. The effect of exercise on cancer-related fatigue in cancer survivors: a systematic review and meta-analysis. NDT 2018 Feb;Volume 14:479-494. [doi: 10.2147/ndt.s150464]
14. Nakano J, Hashizume K, Fukushima T, Ueno K, Matsuura E, Ikio Y, et al. Effects of aerobic and resistance exercises on physical symptoms in cancer patients: a meta-analysis. Integr Cancer Ther 2018 Dec 23;17(4):1048-1058. [doi: 10.1177/1534735418807555] [Medline: 30352523]
15. Paneroni M, Vitacca M, Venturelli M, Simonelli C, Bertacchini L, Scalvini S, et al. The impact of exercise training on fatigue in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. Pulmonology 2020 Sep;26(5):304-313 [FREE Full text] [doi: 10.1016/j.pulmoe.2020.02.004] [Medline: 32184070]
16. Razazian N, Kazeminia M, Moayedi H, Daneshkhah A, Shohaimi S, Mohammadi M, et al. The impact of physical exercise on the fatigue symptoms in patients with multiple sclerosis: a systematic review and meta-analysis. BMC Neurol 2020 Mar 13;20(1):93 [FREE Full text] [doi: 10.1186/s12883-020-01654-y] [Medline: 32169035]
17. Tian L, Lu HJ, Lin L, Hu Y. Effects of aerobic exercise on cancer-related fatigue: a meta-analysis of randomized controlled trials. Support Care Cancer 2016 Feb 19;24(2):969-983. [doi: 10.1007/s00520-015-2953-9] [Medline: 26482381]
18. Hulme K, Safari R, Thomas S, Mercer T, White C, Van der Linden M, et al. Fatigue interventions in long term, physical health conditions: a scoping review of systematic reviews. PLoS One 2018 Oct 12;13(10):e0203367 [FREE Full text] [doi: 10.1371/journal.pone.0203367] [Medline: 30312325]
19. Mustian KM, Alfano CM, Heckler C, Kleckner AS, Kleckner IR, Leach CR, et al. Comparison of pharmaceutical, psychological, and exercise treatments for cancer-related fatigue: a meta-analysis. JAMA Oncol 2017 Jul 01;3(7):961-968 [FREE Full text] [doi: 10.1001/jamaoncol.2016.6914] [Medline: 28253393]
20. Bellizzi KM, Rowland JH, Jeffery DD, McNeel T. Health behaviors of cancer survivors: examining opportunities for cancer control intervention. JCO 2005 Dec 01;23(34):8884-8893. [doi: 10.1200/jco.2005.02.2343]
21. Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. Ther Clin Risk Manag 2005 Sep;1(3):189-199 [FREE Full text] [Medline: 18360559]
22. Albergoni A, Hettinga FJ, La Torre A, Bonato M, Sartor F. The role of technology in adherence to physical activity programs in patients with chronic diseases experiencing fatigue: a systematic review. Sports Med Open 2019 Sep 12;5(1):41 [FREE Full text] [doi: 10.1186/s40798-019-0214-z] [Medline: 31512075]
23. Cox CL, Montgomery M, Oeffinger KC, Leisenring W, Zeltzer L, Whitton JA, et al. Promoting physical activity in childhood cancer survivors: results from the Childhood Cancer Survivor Study. Cancer 2009 Feb 01;115(3):642-654 [FREE Full text] [doi: 10.1002/cncr.24043] [Medline: 19117349]
24. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228 [FREE Full text] [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
25. Ross WL, Le A, Zheng DJ, Mitchell H, Rotatori J, Li F, et al. Physical activity barriers, preferences, and beliefs in childhood cancer patients. Support Care Cancer 2018 Jul 27;26(7):2177-2184. [doi: 10.1007/s00520-017-4041-9] [Medline: 29383508]
26. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. Computers & Education 2012 Sep;59(2):661-686. [doi: 10.1016/j.compedu.2012.03.004]

27. Stinson J, Wilson R, Gill N, Yamada J, Holt J. A systematic review of internet-based self-management interventions for youth with health conditions. J Pediatr Psychol 2009 Jun 23;34(5):495-510. [doi: 10.1093/jpepsy/jsn115] [Medline: 19029142]
28. Hicks C, von Baeyer CL, McGrath P. Online psychological treatment for pediatric recurrent pain: a randomized evaluation. J Pediatr Psychol 2006 Aug;31(7):724-736. [doi: 10.1093/jpepsy/jsj065] [Medline: 16093516]
29. Joseph CLM, Peterson E, Havstad S, Johnson CC, Hoerauf S, Stringer S, et al. A web-based, tailored asthma management program for urban African-American high school students. Am J Respir Crit Care Med 2007 May 01;175(9):888-895. [doi: 10.1164/rccm.200608-1244oc]
30. Runge C, Lecheler J, Horn M, Tews J, Schaefer M. Outcomes of a web-based patient education program for asthmatic children and adolescents. Chest 2006 Mar;129(3):581-593. [doi: 10.1378/chest.129.3.581] [Medline: 16537855]
31. Bonetti A, Drury D, Danoff J, Miller T. Comparison of acute exercise responses between conventional video gaming and isometric resistance exergaming. The Journal of Strength & Conditioning Research 2010;24(7):1799-1803. [doi: 10.1519/jsc.0b013e3181bab4a8]
32. Gutierrez-Martinez O, Gutierrez-Maldonado J, Cabas-Hoyos K, Loreto D. The illusion of presence influences VR distraction: effects on cold-pressor pain. Stud Health Technol Inform 2010;154:155-159. [Medline: 20543289]
33. Göbel S. Serious games application examples. In: Dörner R, Göbel S, Effelsberg W, Wiemeyer J, editors. Serious Games: Foundations, Concepts and Practice. Cham, Switzerland: Springer; 2016:319-405.
34. Kato P, Cole S, Bradlyn A, Pollock B. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. Pediatrics 2008 Aug;122(2):e305-e317. [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
35. Del Corral T, Cebrià I Iranzo MÀ, López-de-Uralde-Villanueva I, Martínez-Alejos R, Blanco I, Vilaró J. Effectiveness of a home-based active video game programme in young cystic fibrosis patients. Respiration 2018;95(2):87-97 [FREE Full text] [doi: 10.1159/000481264] [Medline: 29045949]
36. Rutkowski S, Czech O, Wrzeciono A, Kiper P, Szczepańska-Gieracha J, Malicka I. Virtual reality as a chemotherapy support in treatment of anxiety and fatigue in patients with cancer: a systematic review and meta-analysis and future research directions. Complement Ther Med 2021 Sep;61:102767 [FREE Full text] [doi: 10.1016/j.ctim.2021.102767] [Medline: 34403772]
37. Cugusi L, Prosperini L, Mura G. Exergaming for quality of life in persons living with chronic diseases: a systematic review and meta-analysis. PM R 2021 Jul;13(7):756-780. [doi: 10.1002/pmrj.12444] [Medline: 32592238]
38. Seiler A, Klaas V, Tröster G, Fagundes CP. eHealth and mHealth interventions in the treatment of fatigued cancer survivors: a systematic review and meta-analysis. Psychooncology 2017 Sep;26(9):1239-1253. [doi: 10.1002/pon.4489] [Medline: 28665554]
39. Escalante C. Cancer-related fatigue: Prevalence, screening, and clinical assessment. UpToDate. 2024. URL: https://www.uptodate.com/contents/cancer-related-fatigue-prevalence-screening-and-clinical-assessment [accessed 2024-09-26]
40. Kumbhare D, Ahmed S, Watter S. A narrative review on the difficulties associated with fibromyalgia diagnosis. Therapeutic Advances in Musculoskeletal 2017 Dec 07;10(1):13-26. [doi: 10.1177/1759720x17740076]
41. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
42. Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive meta-analysis (version 2.2.027) [computer software]. ResearchGate. 2005. URL: https://www.researchgate.net/publication/292393194_Comprehensive_meta-analysis_Version_22027_Computer_software [accessed 2024-09-26]
43. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014 Dec 19;14:135 [FREE Full text] [doi: 10.1186/1471-2288-14-135] [Medline: 25524443]
44. Higgins J, Li T, Deeks J. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions. London, UK: Cochrane; 2019:143-176.
45. Thomas S, Fazakarley L, Thomas PW, Collyer S, Brenton S, Perring S, et al. Mii-vitaliSe: a pilot randomised controlled trial of a home gaming system (Nintendo Wii) to increase activity levels, vitality and well-being in people with multiple sclerosis. BMJ Open 2017 Sep 27;7(9):e016966 [FREE Full text] [doi: 10.1136/bmjopen-2017-016966] [Medline: 28954791]
46. Kobayashi D, Watanabe R, Yamamoto M, Kizaki M. Efficacy of physical exercise using the balance board game on physical and psychological function in patients with hematological malignancies confined to a bioclean room. Phys Ther Res 2020;23(2):172-179. [doi: 10.1298/ptr.e10021]
47. Yazgan YZ, Tarakci E, Tarakci D, Ozdincler AR, Kurtuncu M. Comparison of the effects of two different exergaming systems on balance, functionality, fatigue, and quality of life in people with multiple sclerosis: a randomized controlled trial. Mult Scler Relat Disord 2020 Apr;39:101902. [doi: 10.1016/j.msard.2019.101902] [Medline: 31924591]
48. Field AP. The problems in using fixed-effects models of meta-analysis on real-world data. Understanding Statistics 2003 Apr 03;2(2):105-124. [doi: 10.1207/s15328031us0202_02]
49. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]

50. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an Expanded Disability Status Scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [doi: 10.1212/wnl.33.11.1444] [Medline: 6685237]
51. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
52. Brichetto G, Spallarossa P, de Carvalho MLL, Battaglia MA. The effect of Nintendo® Wii® on balance in people with multiple sclerosis: a pilot randomized control study. Mult Scler 2013 Aug 15;19(9):1219-1221. [doi: 10.1177/1352458512472747] [Medline: 23322502]
53. Cuesta-Gómez A, Sánchez-Herrera-Baeza P, Oña-Simbaña ED, Martínez-Medina A, Ortiz-Comino C, Balaguer-Bernaldo-de-Quirós C, et al. Effects of virtual reality associated with serious games for upper limb rehabilitation in patients with multiple sclerosis: randomized controlled trial. J NeuroEngineering Rehabil 2020 Jul 13;17(1):90 [FREE Full text] [doi: 10.1186/s12984-020-00718-x] [Medline: 32660604]
54. De Giglio L, De Luca F, Prosperini L, Borriello G, Bianchi V, Pantano P, et al. A low-cost cognitive rehabilitation with a commercial video game improves sustained attention and executive functions in multiple sclerosis: a pilot study. Neurorehabil Neural Repair 2015 Jun 14;29(5):453-461. [doi: 10.1177/1545968314554623] [Medline: 25398725]
55. Khalil H, Al-Sharman A, El-Salem K, Alghwiri AA, Al-Shorafat D, Khazaaleh S, et al. The development and pilot evaluation of virtual reality balance scenarios in people with multiple sclerosis (MS): a feasibility study. NRE 2019 Jan 02;43(4):473-482. [doi: 10.3233/nre-182471]
56. Ozdogar AT, Ertekin O, Kahraman T, Yigit P, Ozakbas S. Effect of video-based exergaming on arm and cognitive function in persons with multiple sclerosis: a randomized controlled trial. Mult Scler Relat Disord 2020 May;40:101966. [doi: 10.1016/j.msard.2020.101966] [Medline: 32045868]
57. Ozkul C, Guclu-Gunduz A, Yazici G, Atalay Guzel N, Irkec C. Effect of immersive virtual reality on balance, mobility, and fatigue in patients with multiple sclerosis: a single-blinded randomized controlled trial. European Journal of Integrative Medicine 2020 Apr;35:101092. [doi: 10.1016/j.eujim.2020.101092]
58. Ozdogar AT, Ertekin O, Kahraman T, Dastan S, Ozakbas S. Effect of exergaming in people with restless legs syndrome with multiple sclerosis: a single-blind randomized controlled trial. Mult Scler Relat Disord 2023 Feb;70:104480. [doi: 10.1016/j.msard.2022.104480] [Medline: 36603295]
59. Ozdogar AT, Baba C, Kahraman T, Sagici O, Dastan S, Ertekin O, et al. Effects and safety of exergaming in persons with multiple sclerosis during corticosteroid treatment: a pilot study. Mult Scler Relat Disord 2022 Jul;63:103823. [doi: 10.1016/j.msard.2022.103823] [Medline: 35523062]
60. Hamari L, Järvelä LS, Lähteenmäki PM, Arola M, Axelin A, Vahlberg T, et al. The effect of an active video game intervention on physical activity, motor performance, and fatigue in children with cancer: a randomized controlled trial. BMC Res Notes 2019 Nov 29;12(1):784 [FREE Full text] [doi: 10.1186/s13104-019-4821-z] [Medline: 31783907]
61. Villumsen BR, Jorgensen MG, Frystyk J, Hørdam B, Borre M. Home-based 'exergaming' was safe and significantly improved 6-min walking distance in patients with prostate cancer: a single-blinded randomised controlled trial. BJU Int 2019 Oct;124(4):600-608. [doi: 10.1111/bju.14782] [Medline: 31012238]
62. Cho H, Sohng K. The effect of a virtual reality exercise program on physical fitness, body composition, and fatigue in hemodialysis patients. J Phys Ther Sci 2014 Oct;26(10):1661-1665 [FREE Full text] [doi: 10.1589/jpts.26.1661] [Medline: 25364137]
63. Chou H, Chen S, Yen T, Han H. Effect of a virtual reality-based exercise program on fatigue in hospitalized Taiwanese end-stage renal disease patients undergoing hemodialysis. Clin Nurs Res 2020 Jul 15;29(6):368-374. [doi: 10.1177/1054773818788511] [Medline: 30009636]
64. Ribas CG, Alves da Silva L, Corrêa MR, Teive HG, Valderramas S. Effectiveness of exergaming in improving functional balance, fatigue and quality of life in Parkinson's disease: a pilot randomized controlled trial. Parkinsonism Relat Disord 2017 May;38:13-18. [doi: 10.1016/j.parkreldis.2017.02.006] [Medline: 28190675]
65. de Rooij IJM, van de Port IGL, Punt M, Abbink-van Moorsel PJM, Kortsmit M, van Eijk RPA, et al. Effect of virtual reality gait training on participation in survivors of subacute stroke: a randomized controlled trial. Phys Ther 2021 May 04;101(5):pzab051 [FREE Full text] [doi: 10.1093/ptj/pzab051] [Medline: 33594443]
66. Charlier N, Zupancic N, Fieuws S, Denhaerynck K, Zaman B, Moons P. Serious games for improving knowledge and self-management in young people with chronic conditions: a systematic review and meta-analysis. J Am Med Inform Assoc 2016 Jan;23(1):230-239 [FREE Full text] [doi: 10.1093/jamia/ocv100] [Medline: 26186934]
67. Domhardt M, Schröder A, Geirhos A, Steubl L, Baumeister H. Efficacy of digital health interventions in youth with chronic medical conditions: a meta-analysis. Internet Interv 2021 Apr;24:100373 [FREE Full text] [doi: 10.1016/j.invent.2021.100373] [Medline: 33732626]
68. O'Higgins CM, Brady B, O'Connor B, Walsh D, Reilly RB. The pathophysiology of cancer-related fatigue: current controversies. Support Care Cancer 2018 Oct 30;26(10):3353-3364. [doi: 10.1007/s00520-018-4318-7] [Medline: 29961146]
69. Matura LA, Malone S, Jaime-Lara R, Riegel B. A systematic review of biological mechanisms of fatigue in chronic illness. Biol Res Nurs 2018 Jul 14;20(4):410-421 [FREE Full text] [doi: 10.1177/1099800418764326] [Medline: 29540066]
70. Dendrou CA, Fugger L, Friese MA. Immunopathology of multiple sclerosis. Nat Rev Immunol 2015 Sep 15;15(9):545-558. [doi: 10.1038/nri3871] [Medline: 26250739]

71. Oh J, Vidal-Jordana A, Montalban X. Multiple sclerosis: clinical aspects. Current Opinion in Neurology 2018;31(6):752-759. [doi: 10.1097/wco.0000000000000622]
72. Hedges LV, Pigott TD. The power of statistical tests in meta-analysis. Psychological Methods 2001;6(3):203-217. [doi: 10.1037//1082-989x.6.3.203]
73. Callesen J, Cattaneo D, Brincks J, Kjeldgaard Jørgensen ML, Dalgas U. How do resistance training and balance and motor control training affect gait performance and fatigue impact in people with multiple sclerosis? A randomized controlled multi-center study. Mult Scler 2020 Oct 24;26(11):1420-1432. [doi: 10.1177/1352458519865740] [Medline: 31339460]
74. Hebert J, Corboy J. The association between multiple sclerosis-related fatigue and balance as a function of central sensory integration. Gait Posture 2013 May;38(1):37-42. [doi: 10.1016/j.gaitpost.2012.10.015] [Medline: 23200463]
75. Heitkamp H, Horstmann T, Mayer F, Weller J, Dickhuth H. Gain in strength and muscular balance after balance training. Int J Sports Med 2001 May;22(4):285-290. [doi: 10.1055/s-2001-13819] [Medline: 11414672]
76. Davis MP, Walsh D. Mechanisms of fatigue. J Support Oncol 2010;8(4):164-174. [Medline: 20822034]
77. Drummond D, Hadchouel A, Tesnière A. Serious games for health: three steps forwards. Adv Simul (Lond) 2017 Feb 4;2(1):3 [FREE Full text] [doi: 10.1186/s41077-017-0036-3] [Medline: 29450004]
78. Verschueren S, Buffel C, Vander Stichele G. Developing theory-driven, evidence-based serious games for health: framework based on research community insights. JMIR Serious Games 2019 May 02;7(2):e11565 [FREE Full text] [doi: 10.2196/11565] [Medline: 31045496]
79. Li J, Theng Y, Foo S. Effect of exergames on depression: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2016 Jan;19(1):34-42. [doi: 10.1089/cyber.2015.0366] [Medline: 26716638]
80. Lam JW, Sit CH, McManus AM. Play pattern of seated video game and active "exergame" alternatives. Journal of Exercise Science & Fitness 2011;9(1):24-30. [doi: 10.1016/s1728-869x(11)60003-8]
81. Huang H, Wong M, Yang Y, Chiu H, Teng C. Impact of playing exergames on mood states: a randomized controlled trial. Cyberpsychol Behav Soc Netw 2017 Apr;20(4):246-250. [doi: 10.1089/cyber.2016.0322] [Medline: 28394215]
82. Mullins NM, Tessmer KA, McCarroll ML, Peppel BP. Physiological and perceptual responses to Nintendo® Wii Fit™ in young and older adults. Int J Exerc Sci 2012;5(1):79-92 [FREE Full text] [Medline: 27182377]
83. LeLorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 1997 Aug 21;337(8):536-542. [doi: 10.1056/nejm199708213370806]
84. Sun J, Freeman BD, Natanson C. Chapter 22: Meta-analysis of clinical trials. In: Gallin JI, Ognibene FP, Lee Johnson L, editors. Principles and Practice of Clinical Research (Fourth Edition). Cambridge, MA: Academic Press; 2018:317-327.
85. Nap-van der Vlist MM, Dalmeijer GW, Grootenhuis MA, van der Ent CK, van den Heuvel-Eibrink MM, Wulffraat NM, et al. Fatigue in childhood chronic disease. Arch Dis Child 2019 Nov 07;104(11):1090-1095. [doi: 10.1136/archdischild-2019-316782] [Medline: 31175124]
86. Armijo-Olivo S, Ospina M, da Costa BR, Egger M, Saltaji H, Fuentes J, et al. Poor reliability between Cochrane reviewers and blinded external reviewers when applying the Cochrane risk of bias tool in physical therapy trials. PLoS One 2014 May 13;9(5):e96920 [FREE Full text] [doi: 10.1371/journal.pone.0096920] [Medline: 24824199]
87. Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. J Eval Clin Pract 2012 Feb 05;18(1):12-18. [doi: 10.1111/j.1365-2753.2010.01516.x] [Medline: 20698919]

Abbreviations

EDSS: Expanded Disability Status Scale
MS: multiple sclerosis
POMS-sf: Profile of Mood States—short form
SMD: standardized mean difference
VAS: Visual Analog Scale
VR: virtual reality

Edited by T Leung; submitted 14.12.23; peer-reviewed by X Huang, B Lai; comments to author 27.03.24; revised version received 11.04.24; accepted 31.05.24; published 17.10.24.

Please cite as:
Warlo LS, El Bardai S, de Vries A, van Veelen ML, Moors S, Rings EHHM, Legerstee JS, Dierckx B
Game-Based eHealth Interventions for the Reduction of Fatigue in People With Chronic Diseases: Systematic Review and Meta-Analysis
JMIR Serious Games 2024;12:e55034
URL: https://games.jmir.org/2024/1/e55034
doi:10.2196/55034
PMID:

Review

Sensing In Exergames for Efficacy and Motion Quality: Scoping Review of Recent Publications

Sebastian Dill[1*], MSc; Philipp Niklas Müller[2*], MSc; Polona Caserman[2], PhD; Stefan Göbel[2], PD; Christoph Hoog Antink[1], Prof Dr; Thomas Tregel[2], PhD

[1]KIS*MED (AI Systems in Medicine), Technical University of Darmstadt, Darmstadt, Germany

[2]Serious Games Research Group, Technical University of Darmstadt, Darmstadt, Germany

[*]these authors contributed equally

Corresponding Author:
Sebastian Dill, MSc
KIS*MED (AI Systems in Medicine)
Technical University of Darmstadt
Merckstraße 25
Darmstadt, 64283
Germany
Phone: 49 61511623754
Email: dill@kismed.tu-darmstadt.de

Abstract

Background: Many studies have shown a direct relationship between physical activity and health. It has also been shown that the average fitness level in Western societies is lower than recommended by the World Health Organization. One tool that can be used to increase physical activity for individual people is exergaming, that is, serious games that motivate players to do physical exercises.

Objective: This scoping review of recent studies regarding exergame efficacy aims to evaluate which sensing modalities are used to assess exergame efficacy as well as motion quality. We also analyze how the collected motion sensing data is being leveraged with respect to exergame efficacy and motion quality assessment.

Methods: We conducted 2 extensive and systematic searches of the ACM Digital Library and the PubMed database, as well as a single search of the IEEE Xplore database, all according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement. Overall, 343 studies were assessed for eligibility by the following criteria: The study should be peer-reviewed; the year of publication should be between 2015 and 2023; the study should be available in English or German; the study evaluates the efficacy of at least 1 exergame; sensor data is recorded during the study and is used for evaluation; and the study is sufficiently described to extract information on the exergames, sensors, metrics, and results.

Results: We found 67 eligible studies, which we analyzed with regard to sensor usage for both efficacy evaluation and motion analysis. Overall, heart rate (HR) was the most commonly used vital sign to evaluate efficacy (n=52), while the Microsoft Kinect was the most commonly used exergame sensor (n=26). The results of the analysis show that the sensors used in the exergames and the sensors used in the evaluation are, in most cases, mutually exclusive, with motion quality rarely being considered as a metric.

Conclusions: The lack of motion quality assessment is identified as a problem both for the studies and the exergames themselves since incorrectly executed motions can reduce an exergame's effectiveness and increase the risk of injury. Here we propose how to use the same sensors both as input for the exergame and to assess motion quality by presenting recent developments in motion recognition and sensing.

(JMIR Serious Games 2024;12:e52153) doi:10.2196/52153

KEYWORDS

exergame efficacy; motion quality assessment; vital signs; body sensors; camera; virtual reality

Introduction

Exergames are interactive games with the additional goal of engaging players in physical activity to promote a healthy lifestyle and increase players' physical fitness [1,2]. Increasing availability of commercial sensing technologies has allowed exergames such as Beat Saber (Beat Games) [3] for various virtual reality (VR) systems and Ring Fit Adventure (Nintendo) for Nintendo Switch [4] to reach a large audience and become highly commercially successful, selling more than 4 million and 14 million copies, respectively [5,6]. Typically, these exergames use sensing technologies as input devices to either control an in-game avatar or detect the execution of specified motions. However, previous studies on motion-based games suggest that exergames should also give feedback on the motion's quality and provide specific guidelines to follow [7,8]. Without these measures, players may not perform the demanded motions correctly, leading to diminished health benefits and risk of injury [9].

A common research question with regards to exergames is that of their efficacy. Efficacy refers to the effectiveness or ability of exergames to achieve their intended goals or their desired effect. This goal can be defined as an increase in the participants' performance, physical activity, or motivation [10]. Previous reviews and meta-analyses on exergames generally confirm the existence of positive effects associated with exergames [10-13]. Nevertheless, the reviews primarily assert that exergames are most effective in facilitating light- to moderate-intensity physical activity, and only a small proportion of exergames have demonstrated the ability to significantly increase physical activity levels among users [14-16]. Furthermore, previous reviewers have primarily focused on assessing the efficacy of exergames in specific populations, such as children and adolescents without [11,12,16] or explicitly with adults with overweight [13] or older adults [15,17-19]. Only a few reviews encompass a broader range of participants and do not specifically focus on any particular population [14].

The evaluation process, particularly sensor technology usage, is typically not the focus of existing reviews. Hence, this paper investigates recent studies on the efficacy of exergames to identify which sensing technologies are used both in the exergames themselves and in their evaluation. We further review whether current exergames and their evaluations include any motion quality assessment. Motion Quality as a term is not clearly defined and is often visually evaluated by a professional physiotherapist or sports scientist. One of the most comprehensive approaches to defining the term is given by Skjaerven et al [20], who found that motion quality has many different aspects, including biomechanical as well as physiological and temporal characteristics. Therefore, motion quality assessment requires evaluating the motion of all relevant body parts at every point in time. Since we want to focus on the study design and methodology, with little regard for results or target group, we opt for a scoping review approach. For each study, we assess how the sensing data is being leveraged with respect to exergame efficacy and motion quality assessment. Based on our findings, we discuss how state-of-the-art methods for assessing motion quality and already used sensing technologies could be used to improve the efficacy of exergames and reduce the risk of injury during play.

Methods

Overview

The goal of this scoping review was to identify sensing modalities in recent studies that evaluate exergame efficacy. For this, 3 systematic searches were conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [21] and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [22]. A PRISMA-ScR checklist is provided in Multimedia Appendix 1. The database searches were conducted at different points in time by one of the researchers, each without a detailed review protocol.

Search Strategy and Search Terms

Due to the interdisciplinary nature of the research field, 3 different databases were included in the review. To cover both computer science and clinical research, the ACM Digital Library and the subject-specific PubMed database were searched on September 14, 2022, and again on July 14, 2023. A final search on the engineering-focused IEEE Xplore database was conducted on January 15, 2024.

We defined 4 main requirements for studies to be included in our review. First, studies have to feature the evaluation of an exergame. Second, the evaluation has to be conducted quantitatively with sensors with regard to the game's efficacy. Third, the games should focus on general fitness or sports to avoid studies focusing on activities of daily living. Finally, the studies should be recent, which we defined as being published between 2015 and 2023. Based on these requirements, relevant search terms were identified, combined, and generalized as follows:

(exergam* OR "fitness game")

AND (efficacy OR evaluat* OR "heart rate" OR vo2 OR oxygen)

AND (fitness OR sport)

For ACM Digital Library, the search terms were searched for in the categories "Title," "Abstract," and "Author Keyword," each category connected with an OR operator, whereas for PubMed and IEEE Xplore, the search terms were typed "as is" into the "Query Box" and "Command Search," respectively. A detailed search strategy is given in Multimedia Appendix 2.

Study Eligibility Criteria and Selection

Afterward, the following eligibility criteria based on the previously stated requirements were defined. The results were screened and filtered accordingly:

1. Publication type: Peer-reviewed study
2. Publication year: 2015 to 2023
3. Available in English or German
4. The study evaluates the efficacy of at least one exergame

5. Sensor data is recorded during the study and used for evaluation
6. The study is sufficiently described to extract information on the exergames, sensors, metrics, and results

The studies were split among the authors to be screened by abstract and assessed for eligibility. The results of this screening process were then presented and discussed conjointly.

Data Extraction and Data Analysis

Eligible studies were evaluated with regard to their general study design (including participant numbers and focus group), their evaluation methods (including evaluation metrics, vital signs, and sensor usage), and the exergames used (including exergame sensors as well as additional motion sensing). The results were again discussed, assessed for relevance for the review, and generalized conjointly.

Results

Overview

The first search yielded 253 studies of which 30 were removed before the screening process as they were duplicates or not available in English or German. An additional 31 studies were identified through previous work of included authors and citation searching, resulting in a total of 258 studies included in the screening process. 208 studies were excluded in this screening process based on the eligibility criteria stated above, leaving a total of 49 eligible studies. The second search yielded an additional 40 studies, of which 32 were excluded from the screening process. The final search, conducted on the IEEE Xplore database, yielded another 50 studies. After screening by title, abstract, and eligibility criteria, 10 studies were deemed fitting, resulting in a total of 67 studies that were included in our review.

Table 1 gives a brief summary of the vital sign-sensing statistics. Multimedia Appendix 3 [23-89] and Multimedia Appendix 4 [23,24,26-58,60-90] feature additional tables that present details for all 67 studies included in the review. Multimedia Appendix 3 provides an overview of the study design, participants, and how efficacy was assessed. Participation numbers ranged from 6 [23] to 360 [24], with a median of 28. There was a large variety of different focus groups, with the biggest group being healthy adults (n=19 studies). Multimedia Appendix 4 details the exergames and corresponding sensors used in each study and gives information on additional motion sensing if there were any. Not all studies explicitly mention all the exergames evaluated; some feature the evaluation of a multitude of different games, and some tested games they developed themselves. Overall, the 67 studies included in our review feature the evaluation of approximately 49 different exergames. The overall review process with all screening steps is outlined in the PRISMA flow diagram in Figure 1.

Table 1. Summary of vital sign sensing statistics. Detailed statistics for each study can be found in Multimedia Appendix 3.

	Measured vital signs			Evaluated metrics						Evaluation criteria analyzed over time			
	HR[a]	VO_2	Other	mean HR	peak HR or % max HR	peak VO_2 or % max VO_2	MET[b]	EE[c]	Other	Performance	Intensity or PA[d]	Motivation	Other
Number of studies	52	22	18	35	31	21	16	15	45	18	11	6	10
Proportion of studies, %	78	33	27	52	46	31	24	22	67	27	16	9	15

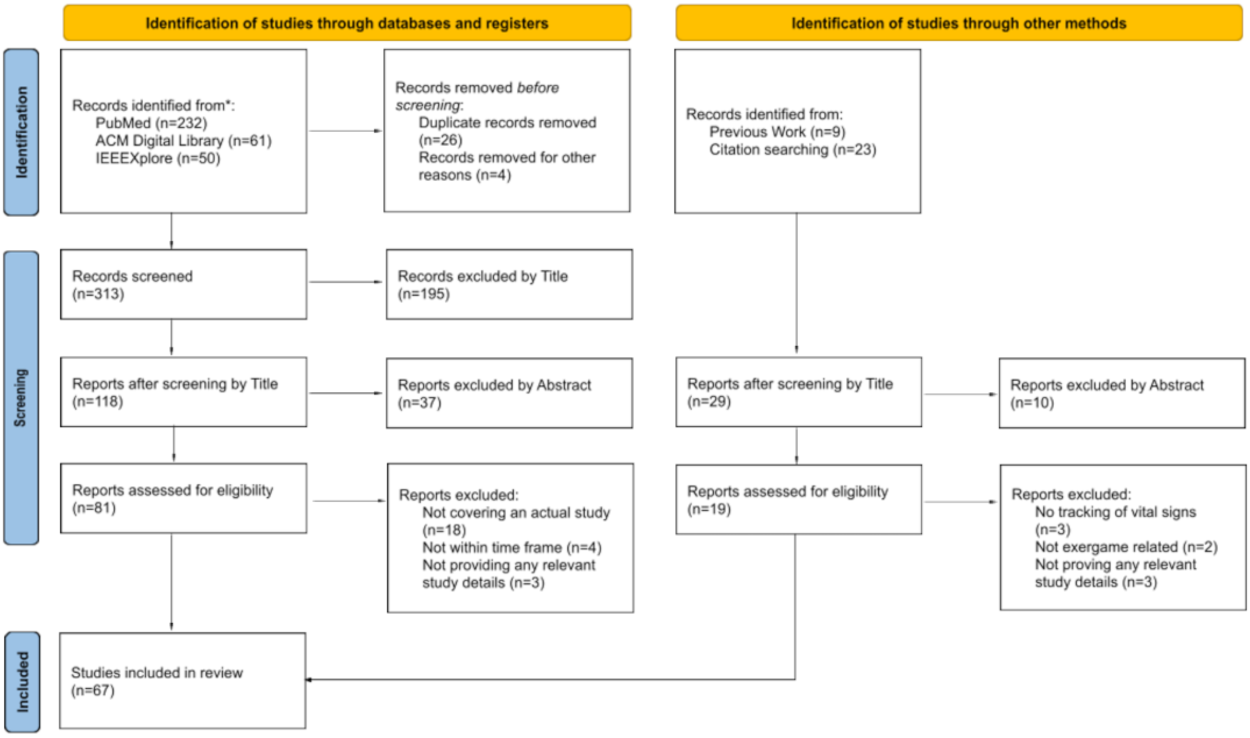
[a]HR: heart rate.

[b]MET: metabolic equivalent of task.

[c]EE: energy expenditure.

[d]PA: physical activity.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [21] flow diagram of the systematic review process.



Vital Sign Sensing Technologies

As per our inclusion criterion, all 67 studies aim to assess efficacy. For that, 54 studies used sensors to measure vital signs, while 11 studies relied only on motion sensing, and 2 studies assessed efficacy without any sensors but with standardized tests only. The majority, that is, 52 of the 54, measured the player's heart rate (HR), and 22 measured the oxygen uptake (VO_2) (more details in Table 1). Furthermore, 18 studies measured other vital signs such as blood pressure, blood lactate concentration, and carbon dioxide production.

Efficacy Assessment

In the context of exergames, efficacy refers to the effectiveness or ability of exergames to achieve their intended goals or their desired effect under optimal or controlled conditions. Efficacy is distinguished from effectiveness, with efficacy focusing on how well an approach works in ideal conditions, while effectiveness considers its performance in real-world or practical situations [91]. However, this distinction is not always clearly made by the researchers [25]. Therefore, in the context of this review, the 2 terms are used interchangeably. Typical goals for exergames can be defined as an increase in the participants' performance, physical activity, or motivation [10,92]. Exergame efficacy is usually evaluated by analyzing these metrics over time or in comparison with conventional physical exercises. Often, they are estimated by measuring vital signs, doing standardized tests, or filling out questionnaires. Efficacy is a highly individual metric and will vary for different activities, different target demographics, and from person to person [10].

When looking at the studies' approaches to assessing exergame efficacy, several different study designs were identified. First of all, a distinction can be made by whether the studies evaluate a single group or compare 2 or more groups. The former is referred to as a within-subject or cross-over study. The latter category is referred to as controlled studies, which again can be split into 2 subcategories. Randomized controlled studies compare randomly split groups, while cross-section studies compare nonrandomized groups based on preexisting conditions. The studies can also be categorized by how many measurements are done. Single-measurement studies use measured metrics and questionnaires to compare the participants' performance, motivation, and general physical activity with nongaming-related exercises. Alternatively, studies with the repeated-measures design, focusing on the long-term effects of exergaming, do the same measurements more than once to observe longitudinal developments. The latter approach was used by 22 studies, while the other 45 studies focused on the first approach. Details on the study design of every included study can be found in Multimedia Appendix 3.

Acute Effects

A total of 53 out of the 67 (79%) studies focus on efficacy evaluation by measuring vital signs in their studies. Out of all analyzed studies (more details in Table 1), 35 reported mean HR, and 31 reported peak HR or proportionate maximum HR (%max HR), with the maximum HR commonly being age-predicted. Ventilation-based evaluation took place in 21 analyzed studies, reporting peak VO_2 or proportionate maximum VO_2 (%max VO_2).

Based on the recorded vital signs, 15 studies processed their participants' data to indicate the recorded energy expenditure (EE) for the exergame activity. In total, 16 studies calculated and reported the metabolic equivalent of task (MET) values for each activity. A total of 8 studies overlapped with studies reporting EE. Furthermore, noteworthy vital-sign-based evaluation metrics are systolic and diastolic blood pressure

[26-29], heart rate reserve [30-33], pulse wave velocity with total peripheral resistance and stroke volume [34], respiratory exchange rate [35,36], lactate concentration [37,38], maximum power output [39], muscle activity [40], concentration through EEG (electroencephalography) [41]. In 3 studies, other derived indicators are used: play intensity and hemodynamic reactivity are calculated based on players' heart rate [37], and the so-called activity counts are based on accelerometer measurements [42,43]. Many studies used subjective questionnaire scales for assessment, which we do not evaluate as they are unrelated to vital sign sensing.

Long-Term Effects

Although many analyzed studies (35 out of 67) comprised more than 1 session of playing exergames, only a minority evaluated changes over time (22 out of 67), and thus qualifies as a longitudinal study in the context of this paper. The evaluation time for longitudinal studies ranged from 4 weeks [44-47] to 6 months [48].

As detailed in Multimedia Appendix 3, most researchers analyzed players' performance. The performance evaluation was predominantly done objectively by tracking changes in anaerobic [46,49,50] or muscular fitness [51,52], functional strength [50], peak VO_2 [28,48], postural control [45], reaction time [24], or cognitive function [53]. The performance was sometimes evaluated subjectively using exercise self-efficacy score [54].

Other evaluation metrics analyzed over time include changes in intensity, usually based on HR or EE [29,45,55,56], whereas further research publications evaluated playtime for different physical activity levels [28,48,51]. In addition to performance and intensity, 6 research publications analyzed changes in motivation, such as play frequency and duration [56] as well as flow [45], player engagement [55] or enjoyment [54], and happiness [57].

Motion Sensing Technologies

The 67 reviewed studies assess various exergames that use different motion sensing technologies, among which the Kinect for Xbox 360 (Microsoft Corp) [93] is the most commonly used (26 out of 67). The Kinect combines a regular RGB (red-green-blue) camera with a depth sensor and can thus be considered a camera-based system. In addition, 1 study each included an exergame using the EyeToy (London Studio) for PlayStation 2 [54] and PS Move (Sony Interactive Entertainment) for PlayStation 3 [44], resulting in 28 out of 67 studies featuring at least one camera-based system. Ergometers, dance mats, balance boards, and other sensing technologies specialized for specific motions are used in 21 out of 67 studies and, thus, make up the second-most common type of motion sensing technology.

VR systems and body-worn inertial measurement units (IMUs) are featured less commonly. Only 5 out of 67 studies feature body-worn IMUs for motion sensing, and only one of them uses IMUs not included in a smartphone or game controller. Furthermore, 7 studies evaluate exergames that use the Nintendo Wii remote controllers held in the player's hands. These controllers are not body-worn and combine an IMU and an infrared sensor with an external emitter for tracking. Together with VR systems, which were used in 16 out of 67 studies as either standalone or within the so-called ExerCube (Sphery AG) [94], they make up the category of hybrid sensing (overall n=22). Finally, one study [58] features a game that does not use any form of motion sensing.

Out of the 67 studies, 21 did additional motion sensing during the exergaming for evaluation purposes. Common motion sensors used for this task are body-worn accelerometers and IMUs (used in 12 studies) and force plates (used in 4 studies). Furthermore, 2 studies used hybrid motion-capture systems, and 2 used camera-based systems. A specialized system to measure reaction times, an exercise bike, surface electromyography, and a handgrip dynamometer were used in 1 study each.

Motion Quality Assessment

Motion quality, as a term, is not clearly defined and is often visually evaluated by a professional physiotherapist or sports scientist. One of the most comprehensive approaches to defining the term is given by Skjaerven et al [20], where they investigated the lived experiences of a group of expert physiotherapists in search of essential characteristics and features of the term. They found that motion quality has many different aspects, including the biomechanical "characteristics of path and form in [motion]" as well as the physiological and temporal "characteristic of flow, elasticity, vitality, and rhythm in [motion]." Therefore, to provide specific feedback on the full motion quality, it would be necessary to assess the whole motion by tracking and evaluating the motion of all relevant body parts at every point in time, independent from the gameplay.

An essential part of exergaming is the game's ability to track and analyze the player's motions to judge the player's gameplay and their exercise quality. Often, exergames focus on the gameplay aspect, only evaluating motions implicitly by requiring players to interact with the virtual environment, for example, slashing, collecting, or avoiding virtual objects [3,39,40,95,96]. Alternatively, some games combine both evaluations by having players fit a predefined shape [59,97,98]. While such an approach might be suitable to verify if players are performing a specific motion at all, it only enables detecting static poses or certain joints without analyzing the holistic motion execution.

None of the 49 games featured in the studies explicitly assess the quality of exercises performed. However, 7 studies use additional motion sensing to do some form of quality assessment during their evaluation. In addition, 4 of these 7 studies focus on specific motion aspects, such as assessing the angular displacement or range of motion of certain joints [41,55], quantifying postural sway [45], and evaluating shoulder flexibility [26]. Only 3 studies [44,51,60] assess the players' motion quality as a whole, using a multitude of sensors and either a standardized test or a professional physiotherapist. However, for 2 of the studies, these assessments do not happen during the exergame but are used to identify general motion quality at times when the subjects are not playing. The final study [44] employs an expert to do an analysis of gross upper body biomechanics based on video recordings of the subjects during the exergames. Therefore, this is the only study in our review that fulfills the condition of holistically evaluating all relevant body parts at every point in time during the exercise, which we consider necessary for motion quality assessment, according to Skjaerven et al [20]. However, this is not an automated process and instead requires manual input from an expert.

Discussion

Overview

To the best of our knowledge, this review represents the first work analyzing sensor usage in studies that examine exergame efficacy. We found 67 eligible studies, which we analyzed with regard to sensor usage for both efficacy evaluation and motion analysis. Overall, HR was the most commonly used vital sign to evaluate efficacy (n=52), while the Microsoft Kinect was the most commonly used exergame sensor (n=26). The results of the analysis show that the sensors used in the exergames and the sensors used in the evaluation are, in most cases, mutually exclusive.

Principal Findings

From the review, 2 main conclusions can be drawn: First, the majority of papers focus on measuring vital signs to infer a physical activity metric. Second, an overall lack of motion quality assessment can be stated among exergame studies. For the studies, this lack of quality assessment can lead to undetected and unwanted influences on the results. For example, it might be hard to differentiate if a lack of efficacy should be attributed to the exergame or to the participants' execution, especially in longitudinal studies. Furthermore, it also poses a problem for the exergames themselves since the players are incentivized to optimize for score, which might lead to optimal gameplay that does not coincide with a correct exercise execution and might lead to injuries instead of increased health [9]. Therefore, in the following outlook, we outline how different types of sensor technologies could be used with existing motion analysis techniques to assess motion quality in exergames.

Outlook

Quality Assessment Based on Cameras

Overview

While the traditional approach of motion analysis by a camera is a hybrid one with visual markers attached to certain body parts, a lot of recent research has focused on markerless methods solely based on single-camera video images [99]. This gives cameras a few advantages over other sensing methods: They are unobtrusive to the players and allow them a full range of motion; they can be easily set up and moved around, and they are comparably cheap and widely available [99]. The recent approaches can be differentiated by their modality: While classic RGB cameras are mostly used when no depth information is needed, so-called RGB-D (red-green-blue-depth) cameras like the Microsoft Kinect [93] are able to also record depth information [100,101].

RGB-D

As stated, 26 of the 28 exergames considered in the review that used cameras for the gaming input relied on the Kinect RGB-D camera. In addition to the advantages of cameras already listed, the Kinect also has its origin in gaming, making it a widespread tool not only for scientists but also for game developers. However, all papers in our review use the Kinect only as input for the games, not for an explicit motion quality assessment. This is surprising as, outside of the exergaming context, RGB-D cameras are used in several motion research areas, such as gait analysis [102-104] and fall detection [105], where they have shown to be a reliable tool to capture and evaluate full-body movements. Sporadically, the Kinect has been used in exergaming-related motion analysis as well, albeit not for quality control in an efficacy assessment study. Examples include motion dissimilarity analysis [106] or interrater reliability evaluation between a Kinect system [107] and a human rater.

RGB

In contrast, classic RGB cameras are not commonly used as motion sensors in exergaming, aside from rare examples [108]. This is also apparent from our review since only 2 exergames used an RGB camera compared to the 26 using the Kinect. However, studies without exergame context have shown that modern RGB-based systems are well-suited to do quantitative motion analysis. Systems like MonoCap (Zhejiang University) [109], OpenPose (Carnegie Mellon University) [110], and MediaPipe Pose (Google) [111] have proven that full-body pose estimation can be done with high accuracy. One of the most well-researched applications of RGB-based motion analysis is gait analysis, which can either be done with a feature-based approach using pose estimation [112-114] or a feature-less approach directly on the images [115,116]. More complex medical applications such as joint load prediction [117] and motion limitation analysis [118] indicate the method's ability to do precise quality assessment of human movements. Furthermore, RGB cameras have already been used in sports analysis [119]. These use cases imply the suitability of RGB cameras as a tool for both gaming input and quality assessment in exergaming.

Quality Assessment Based on Hybrid Sensing Techniques

The release of consumer-grade virtual reality systems contributed to the development of many immersive exergames. An analysis of the 29 top VR exergames from a recent review [120] shows that players prefer games providing a high level of exertion (equivalent to real-world exercise level), whereas a high level of immersion is important for distraction (reducing perceived exertion). Most reviewed exergames using hybrid sensing technology indeed provide a playful fitness experience; nevertheless, existing approaches often fail to analyze motion execution to detect errors or to provide specific feedback on motion quality. For example, approaches letting players fit a predefined shape [59,97,98] might be suitable to verify if players are performing a specific motion at all. However, they only enable the detection of static poses without analyzing motion execution. Furthermore, many VR exergames enabling interaction with VR objects usually depend only on hand-held devices and often lack lower body movement.

To overcome this limitation, other researchers use additional off-the-shelf VR sensors. Previous studies already concluded that VR sensors are feasible for clinical, research, and industry usage [121,122]. For example, Martin-Niedecken et al [90,123] demonstrated how trackers attached to wrists and ankles can be used to recognize different exercises, such as burpees, lunges, and punches. The motion quality, presented by a star ranking, is then assessed according to how well players reached predetermined target points and how quickly they returned to their initial pose. A further study [124] also analyzes the entire motion execution of different yoga poses and thereby identifies execution errors.

Another possibility to assess motion quality is marker-based motion capture systems, for example, OptiTrack (NaturalPoint Inc.) [125] and Vicon Tracker (Vicon Motion Systems Ltd.) [126]. These systems are the gold standard for tracking individual joint positions and angular movements with high accuracy and low latency [127,128]. However, as such systems require several markers and cameras, they can be obtrusive and not suitable for home or clinical environments. Nevertheless, previous research publications have already demonstrated that motion capture suits are reliable for analyzing holistic motion executions during dance [129], identifying exercise execution errors during physical exercise [130], or conducting kinematic trunk motion analysis [60].

Quality Assessment Based on Body-Worn Sensors

IMUs consisting of an accelerometer, a gyroscope, and sometimes a magnetometer are the most common types of body-worn sensors. They are integrated into commercial devices such as smartphones, smartwatches, and fitness trackers. Among other things, they are often used to track fitness activities and activities of daily living. Compared with camera-based or hybrid approaches, they are particularly well suited to track nonstationary activities. In the following, we will not discuss other body-worn sensors, such as EMG sensors, because of their limited availability and applicability.

While regular body-worn IMUs do not provide as extensive motion data as camera-based or hybrid systems, exergames could use them as robust, low-cost sensors to assess specific quality metrics and enable tracking without a stationary setup. An example of this can be seen in the commercially successful exergame Ring Fit Adventure (Nintendo) for the Nintendo Wii, which tracks and assesses performed exercises through IMUs placed at the thigh and the inside of an elastic ring held by the user. Alternatively, commercial IMU-based systems such as Xsens (Movella) [131], Noraxon Ultium Motion (Noraxon) [132], or the Teslasuit (Teslasuit, Deep Divers Ltd) [133] could be used for an extensive motion analysis in a mobile setting [130].

Whereas some exergames in our reviewed papers use controllers that incorporate an IMU, none rely purely on body-worn sensors for tracking. In 7 papers [41,43,51,60-64], body-worn IMUs are used to assess physical activity. Only Ko et al [41] and Mueller et al [60] additionally use body-worn IMU data for quality assessment by determining the users' range of motion and back posture respectively.

In their review, Rana and Mittal [134] show that wearable sensors can be successfully deployed for kinematic analysis in a variety of sports applications. While these included sports applications such as swimming [135,136], which are ill-suited for exergames, most sports applications presented could be integrated into exergames. Particularly noteworthy applications are swing sports such as tennis [137-143] and badminton [140,144], in which IMUs can either be wrist-worn or integrated into the racket to track and assess individual swings, as it can be difficult to track these with stationary setups in a practical setting.

Limitations

This scoping review is prone to the same search-related limitations as other reviews of this type. First, only articles published in international journals and full articles published in conference proceedings written in English or German are considered. Therefore, potentially relevant studies published in other languages may have been missed. Second, some articles were excluded from the analysis because the required information to assess their eligibility based on the inclusion criteria was not provided. Third, we assess a lack of reproducibility for the ACM Digital Library. In general, a very high volatility can be noted in the amount of records ACM's search function returns when changing individual words or operators. Since the current search results in less records than the original search, which were all included in the review, we consider this to be a minor limitation. Finally, out of the 67 studies included in this scoping review, 8 studies were not found firsthand using the search terms but instead through previous work of included authors and citation searching. This may indicate that there would have been even more fitting search terms for the review question. However, due to the author's experience with the topic and the unambiguity of the results, we are confident that we were able to mitigate any possible negative effects and that the included studies present a complete overview of the current state of research.

In addition to limitations related to the search, we can also note 2 limitations related to the scope of evaluation. They are a deliberately chosen result of our research focus, which means to analyze the studies' methodology with regards to sensing instead of their results. First, the definitions for the terms "efficacy" and "motion quality" used in this paper focus on how these 2 metrics are evaluated and do not go into detail on what qualifies as "good" efficacy or motion quality. Instead, we evaluate how well the studies are able to assess efficacy and motion quality with the methodology and especially the sensors they use. Therefore, we do not define what a "high/low quality motion" may look like as this question is highly specific to the individual motion and therefore cannot be answered generally. For the same reason, we also do not go into detail on the cohorts' age, sex, and focus group as they are predominantly relevant to the studies' results.

Conclusions

In this paper, we conduct a scoping review of recent studies that examine exergame efficacy to determine which sensors are being used and how they are being used during gameplay and evaluation. Our results show that most studies evaluate exergame efficacy by measuring vital signs, the most common being heart rate (52 out of 67) and oxygen consumption (22 out of 67). However, motion quality is only assessed in 7 out of 67 studies despite being an important factor in an exergame's effectiveness and risk of injury. Furthermore, out of the 49 exergames evaluated in the reviewed studies, none feature quality assessment during or after gameplay, and only 3 studies feature motion quality assessment beyond the exergame's feedback.

Since exergames already use motion sensing technologies to track the player's motions, they could also be used for external quality assessment. Therefore, we discuss recent advances in the field of motion analysis and potential use cases of different sensors commonly used in exergames. We come to the conclusion that many of the same sensing technologies typically used in exergames and exergame studies are well-suited for additional motion quality assessment to ensure consistent exergame effectiveness and reduce the likelihood of injury while exergaming.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the Hessian Ministry for Digital Strategy and Development (Hessisches Ministerium für Digitale Strategie und Entwicklung, Distr@l-Förderlinie 2, "SG4smartmedication", 21_0038_2A), and the Technische Universität Darmstadt and University and State Library Darmstadt in publishing this work as Open Access.

Authors' Contributions

SD contributed to the investigation, methodology, conceptualization, and writing the original draft. PNM contributed to the investigation, methodology, conceptualization, and writing the original draft. SD and PNM contributed equally as shared first authors. PC contributed to the investigation, writing the original draft, and visualization; SG contributed to conceptualization, writing, reviewing, and editing. CHA contributed to conceptualization, writing, reviewing, and editing. TT contributed to the investigation, methodology, conceptualization, writing the original draft, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preferred reporting items for systematic reviews and meta-analyses extension for scoping reviews (PRISMA-ScR) checklist.
[PDF File (Adobe PDF File), 549 KB - games_v12i1e52153_app1.pdf]

Multimedia Appendix 2
Search strategies and results for individual databases.
[DOCX File , 15 KB - games_v12i1e52153_app2.docx]

Multimedia Appendix 3
Summary of studies with regards to the Evaluation of Exergame Efficacy.
[DOCX File , 51 KB - games_v12i1e52153_app3.docx]

Multimedia Appendix 4
Summary of studies with regards to the Evaluation of Games and Motion Sensing.
[DOCX File , 42 KB - games_v12i1e52153_app4.docx]

References

1. Dörner R, Göbel S, Effelsberg W, Wiemeyer J. Serious Games: Foundations, Concepts and Practice. India: Springer; 2016.
2. Staiano AE. Exergames, energy expenditure, and obesity. In: The International Encyclopedia of Media Psychology. United States: John Wiley & Sons, Ltd; 2020:1-7.
3. Beat Saber - VR Rhythm Game. URL: https://beatsaber.com/ [accessed 2022-11-11]
4. Ring Fit AdventureTM for Nintendo SwitchTM. URL: https://ringfitadventure.nintendo.com/ [accessed 2022-12-22]
5. IR Information: Sales Data - Top Selling Title Sales Units, Nintendo Co., Ltd. URL: http://www.nintendo.co.jp/ir/en/finance/software/index.html [accessed 2022-11-11]
6. Verdu M. From Bear to Bull: How Oculus Quest 2 Is Changing the Game for VR. URL: https://www.oculus.com/blog/from-bear-to-bull-how-oculus-quest-2-is-changing-the-game-for-vr/ [accessed 2022-11-11]

7. Caserman P, Hoffmann K, Müller P, Schaub M, Straßburg K, Wiemeyer J, et al. Quality criteria for serious games: serious part, game part, and balance. JMIR Serious Games 2020;8(3):e19037 [FREE Full text] [doi: 10.2196/19037] [Medline: 32706669]
8. Mueller F, Isbister K. Movement-based game guidelines. : Association for Computing Machinery; 2014 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2014 April 26; New York, NY, USA p. 2191-2200. [doi: 10.1145/2556288.2557163]
9. Slobounov SM. Injuries in Athletics: Causes and Consequences. Germany: Springer Science & Business Media; 2008.
10. Whitehead A, Johnston H, Nixon N, Welch J. Exergame effectiveness: what the numbers can tell us. : Association for Computing Machinery; 2010 Presented at: Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games; 2010 July 28; New York, NY, USA p. 55-62. [doi: 10.1145/1836135.1836144]
11. Gao Z, Chen S, Pasco D, Pope Z. A meta-analysis of active video games on health outcomes among children and adolescents. Obes Rev 2015;16(9):783-794. [doi: 10.1111/obr.12287] [Medline: 25943852]
12. Joronen K, Aikasalo A, Suvitie A. Nonphysical effects of exergames on child and adolescent well-being: a comprehensive systematic review. Scand J Caring Sci 2017;31(3):449-461. [doi: 10.1111/scs.12393] [Medline: 27862145]
13. Andrade A, Correia CK, Coimbra DR. The psychological effects of exergames for children and adolescents with obesity: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2019;22(11):724-735. [doi: 10.1089/cyber.2019.0341] [Medline: 31697604]
14. Peng W, Crouse JC, Lin JH. Using active video games for physical activity promotion: a systematic review of the current state of research. Health Educ Behav 2013;40(2):171-192. [doi: 10.1177/1090198112444956] [Medline: 22773597]
15. Larsen LH, Schou L, Lund HH, Langberg H. The physical effect of exergames in healthy elderly-a systematic review. Games Health J 2013;2(4):205-212. [doi: 10.1089/g4h.2013.0036] [Medline: 26192224]
16. LeBlanc AG, Chaput JP, McFarlane A, Colley RC, Thivel D, Biddle SJH, et al. Active video games and health indicators in children and youth: a systematic review. PLoS One 2013;8(6):e65351 [FREE Full text] [doi: 10.1371/journal.pone.0065351] [Medline: 23799008]
17. Cacciata M, Stromberg A, Lee JA, Sorkin D, Lombardo D, Clancy S, et al. Effect of exergaming on health-related quality of life in older adults: a systematic review. Int J Nurs Stud 2019;93:30-40 [FREE Full text] [doi: 10.1016/j.ijnurstu.2019.01.010] [Medline: 30861452]
18. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016;85(1):1-16 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
19. Li J, Erdt M, Chen L, Cao Y, Lee SQ, Theng YL. The social effects of exergames on older adults: systematic review and metric analysis. J Med Internet Res 2018;20(6):e10486 [FREE Full text] [doi: 10.2196/10486] [Medline: 29954727]
20. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. Physiother Theory Pract 2008;24(1):13-27. [doi: 10.1080/01460860701378042] [Medline: 18300105]
21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021;10(1):89 [FREE Full text] [doi: 10.1186/s13643-021-01626-4] [Medline: 33781348]
22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
23. Wünsche BC, Abernethy M, Hyniewska S, Luo Z, Pawar RP, Cao S, et al. Rift racers - effect of balancing and competition on exertion, enjoyment, and motivation in an immersive exergame. 2021 Presented at: Proceedings of the 36th International Conference on Image and Vision Computing New Zealand (IVCNZ); 2021 December 9-10; Tauranga, New Zealand. [doi: 10.1109/ivcnz54163.2021.9653159]
24. Badau D, Badau A, Ene-Voiculescu C, Larion A, Ene-Voiculescu V, Mihaila I, et al. The impact of implementing an exergame program on the level of reaction time optimization in handball, volleyball, and basketball players. Int J Environ Res Public Health 2022;19(9):5598 [FREE Full text] [doi: 10.3390/ijerph19095598] [Medline: 35564993]
25. Singal AG, Higgins PDR, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol 2014;5(1):e45 [FREE Full text] [doi: 10.1038/ctg.2013.13] [Medline: 24384867]
26. Li J, Li L, Huo P, Ma C, Wang L, Theng YL. Wii or kinect? A pilot study of the exergame effects on older adults' physical fitness and psychological perception. Int J Environ Res Public Health 2021;18(24):12939 [FREE Full text] [doi: 10.3390/ijerph182412939] [Medline: 34948547]
27. Neves LEDS, Cerávolo MPDS, Silva E, De Freitas WZ, Da Silva FF, Higino WP, et al. Cardiovascular effects of zumba(®) performed in a virtual environment using XBOX Kinect. J Phys Ther Sci 2015;27(9):2863-2865 [FREE Full text] [doi: 10.1589/jpts.27.2863] [Medline: 26504312]
28. Berg J, Haugen G, Wang AI, Moholdt T. High-intensity exergaming for improved cardiorespiratory fitness: a randomised, controlled trial. Eur J Sport Sci 2022;22(6):867-876 [FREE Full text] [doi: 10.1080/17461391.2021.1921852] [Medline: 33944698]

29. Mologne MS, Hu J, Carrillo E, Gomez D, Yamamoto T, Lu S, et al. The efficacy of an immersive virtual reality exergame incorporating an adaptive cable resistance system on fitness and cardiometabolic measures: a 12-week randomized controlled trial. Int J Environ Res Public Health 2022;20(1):210 [FREE Full text] [doi: 10.3390/ijerph20010210] [Medline: 36612530]
30. Lisón JF, Cebolla A, Guixeres J, Álvarez-Pitti J, Escobar P, Bruñó A, et al. Competitive active video games: physiological and psychological responses in children and adolescents. Paediatr Child Health 2015;20(7):373-376 [FREE Full text] [doi: 10.1093/pch/20.7.373] [Medline: 26526217]
31. Polechoński J, Dębska M, Dębski PG. Exergaming can be a health-related aerobic physical activity. Biomed Res Int 2019;2019:1890527 [FREE Full text] [doi: 10.1155/2019/1890527] [Medline: 31275964]
32. Çakir-Atabek H, Aygün C, Dokumacı B. Active video games versus traditional exercises: energy expenditure and blood lactate responses. Res Q Exerc Sport 2020;91(2):188-196. [doi: 10.1080/02701367.2019.1653431] [Medline: 31625817]
33. Evans E, Naugle KE, Kaleth AS, Arnold B, Naugle KM. Physical activity intensity, perceived exertion, and enjoyment during head-mounted display virtual reality games. Games Health J 2021;10(5):314-320. [doi: 10.1089/g4h.2021.0036] [Medline: 34449262]
34. Kircher E, Ketelhut S, Ketelhut K, Röglin L, Martin-Niedecken AL, Hottenrott K, et al. Acute effects of heart rate-controlled exergaming on vascular function in young adults. Games Health J 2022;11(1):58-66. [doi: 10.1089/g4h.2021.0196] [Medline: 34986027]
35. McGuire S, Willems ME. Physiological responses during multiplay exergaming in young adult males are game-dependent. J Hum Kinet 2015;46:263-271 [FREE Full text] [doi: 10.1515/hukin-2015-0054] [Medline: 26240669]
36. Ciążyńska J, Maciaszek J. Effects of low-immersive vs. High-immersive exercise environment on postural stability and reaction and motor time of healthy young adults. J Clin Med 2023;12(1):389 [FREE Full text] [doi: 10.3390/jcm12010389] [Medline: 36615191]
37. Ketelhut S, Ketelhut RG, Kircher E, Röglin L, Hottenrott K, Martin-Niedecken AL, et al. Gaming instead of training? Exergaming induces high-intensity exercise stimulus and reduces cardiovascular reactivity to cold pressor test. Front Cardiovasc Med 2022;9:798149 [FREE Full text] [doi: 10.3389/fcvm.2022.798149] [Medline: 35155627]
38. Chung LMY, Sun FH, Cheng CTM. Physiological and perceived responses in different levels of exergames in elite athletes. Games Health J 2017;6(1):57-60. [doi: 10.1089/g4h.2016.0074] [Medline: 28192034]
39. Farrow M, Lutteroth C, Rouse PC, Bilzon JLJ. Virtual-reality exergaming improves performance during high-intensity interval training. Eur J Sport Sci 2019;19(6):719-727 [FREE Full text] [doi: 10.1080/17461391.2018.1542459] [Medline: 30403927]
40. Feodoroff B, Konstantinidis I, Froböse I. Effects of full body exergaming in virtual reality on cardiovascular and muscular parameters: cross-sectional experiment. JMIR Serious Games 2019;7(3):e12324 [FREE Full text] [doi: 10.2196/12324] [Medline: 31464194]
41. Ko J, Jang SW, Lee HT, Yun HK, Kim YS. Effects of virtual reality and non-virtual reality exercises on the exercise capacity and concentration of users in a ski exergame: comparative study. JMIR Serious Games 2020;8(4):e16693 [FREE Full text] [doi: 10.2196/16693] [Medline: 33112240]
42. Sousa CV, Hwang J, Cabrera-Perez R, Fernandez A, Misawa A, Newhook K, et al. Active video games in fully immersive virtual reality elicit moderate-to-vigorous physical activity and improve cognitive performance in sedentary college students. J Sport Health Sci 2022;11(2):164-171 [FREE Full text] [doi: 10.1016/j.jshs.2021.05.002] [Medline: 34004390]
43. Pasco D, Roure C, Kermarrec G, Pope Z, Gao Z. The effects of a bike active video game on players' physical activity and motivation. J Sport Health Sci 2017;6(1):25-32 [FREE Full text] [doi: 10.1016/j.jshs.2016.11.007] [Medline: 30356595]
44. Rosly MM, Halaki M, Rosly HM, Hasnan N, Husain R, Davis GM. Arm exercises for individuals with spinal cord injury: exergaming versus arm cranking. 2019 Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); 2019 August 05-07; Kyoto, Japan p. 1-7. [doi: 10.1109/segah.2019.8882460]
45. Barry G, van Schaik P, MacSween A, Dixon J, Martin D. Exergaming (XBOX Kinect™) versus traditional gym-based exercise for postural control, flow and technology acceptance in healthy adults: a randomised controlled trial. BMC Sports Sci Med Rehabil 2016;8(1):25 [FREE Full text] [doi: 10.1186/s13102-016-0050-0] [Medline: 27555917]
46. Wu YS, Wang WY, Chan TC, Chiu YL, Lin HC, Chang YT, et al. Effect of the nintendo ring fit adventure exergame on running completion time and psychological factors among university students engaging in distance learning during the COVID-19 pandemic: randomized controlled trial. JMIR Serious Games 2022;10(1):e35040 [FREE Full text] [doi: 10.2196/35040] [Medline: 35315780]
47. Sheu FR, Lee YL, Hsu HT, Chen NS. Effects of gesture-based fitness games on functional fitness of the elders. 2015 Presented at: 2015 IEEE 15th International Conference on Advanced Learning Technologies; 2015 July 06-09; Hualien, Taiwan. [doi: 10.1109/icalt.2015.35]
48. Berg J, Wang AI, Lydersen S, Moholdt T. Can gaming get you fit? Front Physiol 2020;11:1017 [FREE Full text] [doi: 10.3389/fphys.2020.01017] [Medline: 32973553]
49. Ketelhut S, Röglin L, Martin-Niedecken AL, Nigg CR, Ketelhut K. Integrating regular exergaming sessions in the ExerCube into a school setting increases physical fitness in elementary school children: a randomized controlled trial. J Clin Med 2022;11(6):1570 [FREE Full text] [doi: 10.3390/jcm11061570] [Medline: 35329896]

50. Smits-Engelsman BCM, Jelsma LD, Ferguson GD. The effect of exergames on functional strength, anaerobic fitness, balance and agility in children with and without motor coordination difficulties living in low-income communities. Hum Mov Sci 2017;55:327-337. [doi: 10.1016/j.humov.2016.07.006] [Medline: 27423302]
51. Comeras-Chueca C, Villalba-Heredia L, Perez-Lasierra JL, Marín-Puyalto J, Lozano-Berges G, Matute-Llorente A, et al. Active video games improve muscular fitness and motor skills in children with overweight or obesity. Int J Environ Res Public Health 2022;19(5):2642 [FREE Full text] [doi: 10.3390/ijerph19052642] [Medline: 35270330]
52. Lin CC, Lin YS, Yeh CH, Huang CC, Kuo LC, Su FC. An exergame-integrated IoT-based ergometer system delivers personalized training programs for older adults and enhances physical fitness: a pilot randomized controlled trial. Gerontology 2023;69(6):768-782. [doi: 10.1159/000526951] [Medline: 36617413]
53. Gouveia É, Smailagic A, Ihle A, Marques A, Gouveia BR, Cameirão M, et al. The efficacy of a multicomponent functional fitness program based on exergaming on cognitive functioning of healthy older adults: a randomized controlled trial. J Aging Phys Act 2021;29(4):586-594 [FREE Full text] [doi: 10.1123/japa.2020-0083] [Medline: 33361495]
54. Dos Santos H, Bredehoft MD, Gonzalez FM, Montgomery S. Exercise video games and exercise self-efficacy in children. Glob Pediatr Health 2016;3:2333794X16644139 [FREE Full text] [doi: 10.1177/2333794X16644139] [Medline: 27336015]
55. Bronner S, Pinsker R, Naik R, Noah JA. Physiological and psychophysiological responses to an exer-game training protocol. J Sci Med Sport 2016;19(3):267-271. [doi: 10.1016/j.jsams.2015.03.003] [Medline: 25824058]
56. Kaos MD, Beauchamp MR, Bursick S, Latimer-Cheung AE, Hernandez H, Warburton DER, et al. Efficacy of online multi-player versus single-player exergames on adherence behaviors among children: a nonrandomized control trial. Ann Behav Med 2018;52(10):878-889. [doi: 10.1093/abm/kax061] [Medline: 30212848]
57. Hastürk G, Akyıldız Munusturlar M. The effects of exergames on physical and psychological health in young adults. Games Health J 2022;11(6):425-434. [doi: 10.1089/g4h.2022.0093] [Medline: 36126314]
58. Patten JW, Iarocci G, Bojin N. A pilot study of children's physical activity levels during imagination-based mobile games. J Child Health Care 2017;21(3):292-300. [doi: 10.1177/1367493517708477] [Medline: 29119828]
59. Born F, Abramowski S, Masuch M. Exergaming in VR: the impact of immersive embodiment on motivation, performance, and perceived exertion. 2019 Presented at: 2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games); 2019 September 04-06; Vienna, Austria. [doi: 10.1109/vs-games.2019.8864579]
60. Mueller J, Niederer D, Tenberg S, Oberheim L, Moesner A, Mueller S. Acute effects of game-based biofeedback training on trunk motion in chronic low back pain: a randomized cross-over pilot trial. BMC Sports Sci Med Rehabil 2022;14(1):192 [FREE Full text] [doi: 10.1186/s13102-022-00586-z] [Medline: 36372870]
61. Mackintosh KA, Standage M, Staiano AE, Lester L, McNarry MA. Investigating the physiological and psychosocial responses of single- and dual-player exergaming in young adults. Games Health J 2016;5(6):375-381 [FREE Full text] [doi: 10.1089/g4h.2016.0015] [Medline: 27782766]
62. McDonough DJ, Pope ZC, Zeng N, Lee JE, Gao Z. Comparison of college students' energy expenditure, physical activity, and enjoyment during exergaming and traditional exercise. J Clin Med 2018;7(11):433 [FREE Full text] [doi: 10.3390/jcm7110433] [Medline: 30423805]
63. McDonough DJ, Pope ZC, Zeng N, Lee JE, Gao Z. Retired elite athletes' physical activity, physiological, and psychosocial outcomes during single- and double-player exergaming. J Strength Cond Res 2019;33(12):3220-3225. [doi: 10.1519/JSC.0000000000003386] [Medline: 31567792]
64. Lin Y, Wang J, Luo Z, Li S, Zhang Y, Wünsche BC. Dragon hunter: loss aversion for increasing physical activity in AR exergames. 2023 Presented at: Proceedings of the 2023 Australasian Computer Science Week, in ACSW; 2023 March 13; New York, NY, USA p. 212-221. [doi: 10.1145/3579375.3579403]
65. Wu PT, Wu WL, Chu IH. Energy expenditure and intensity in healthy young adults during exergaming. Am J Health Behav 2015;39(4):556-561. [doi: 10.5993/AJHB.39.4.12] [Medline: 26018104]
66. Hoffmann K, Sportwiss D, Hardy S, Wiemeyer J, Göbel S. Personalized adaptive control of training load in cardio-exergames--a feasibility study. Games Health J 2015;4(6):470-479. [doi: 10.1089/g4h.2014.0073] [Medline: 26484388]
67. Rodrigues GAA, Felipe DDS, Silva E, De Freitas WZ, Higino WP, Da Silva FF, et al. Acute cardiovascular responses while playing virtual games simulated by Nintendo wii(®). J Phys Ther Sci 2015;27(9):2849-2851 [FREE Full text] [doi: 10.1589/jpts.27.2849] [Medline: 26504308]
68. Benzing V, Heinks T, Eggenberger N, Schmidt M. Acute cognitively engaging exergame-based physical activity enhances executive functions in adolescents. PLoS One 2016;11(12):e0167501 [FREE Full text] [doi: 10.1371/journal.pone.0167501] [Medline: 28030542]
69. Moholdt T, Weie S, Chorianopoulos K, Wang AI, Hagen K. Exergaming can be an innovative way of enjoyable high-intensity interval training. BMJ Open Sport Exerc Med 2017;3(1):e000258 [FREE Full text] [doi: 10.1136/bmjsem-2017-000258] [Medline: 29021913]
70. Tietjen AMJ, Devereux GR. Physical demands of exergaming in healthy young adults. J Strength Cond Res 2019;33(7):1978-1986. [doi: 10.1519/JSC.0000000000002235] [Medline: 29189579]

71. Rodrigues GAA, Rodrigues PC, da Silva FF, Nakamura PM, Higino WP, de Souza RA. Mini-trampoline enhances cardiovascular responses during a stationary running exergame in adults. Biol Sport 2018;35(4):335-342 [FREE Full text] [doi: 10.5114/biolsport.2018.78052] [Medline: 30765918]
72. Viana RB, Gentil P, Andrade MS, Vancini RL, de Lira CAB. Is the energy expenditure provided by exergames valid? Int J Sports Med 2019;40(9):563-568. [doi: 10.1055/a-0955-9394] [Medline: 31295745]
73. Roure C, Pasco D, Benoît N, Deldicque L. Impact of a design-based bike exergame on young adults' physical activity metrics and situational interest. Res Q Exerc Sport 2020;91(2):309-315. [doi: 10.1080/02701367.2019.1665621] [Medline: 31718499]
74. Berg J, Moholdt T. Game on: a cycling exergame can elicit moderate-to-vigorous intensity. a pilot study. BMJ Open Sport Exerc Med 2020;6(1):e000744 [FREE Full text] [doi: 10.1136/bmjsem-2020-000744] [Medline: 32341801]
75. Martin-Niedecken AL, Schwarz T, Schättin A. Comparing the impact of heart rate-based in-game adaptations in an exergame-based functional high-intensity interval training on training intensity and experience in healthy young adults. Front Psychol 2021;12:572877 [FREE Full text] [doi: 10.3389/fpsyg.2021.572877] [Medline: 34234705]
76. Aygün C, Çakir-Atabek H. Alternative model for physical activity: active video games lead to high physiological responses. Res Q Exerc Sport 2022;93(3):447-456. [doi: 10.1080/02701367.2020.1864258] [Medline: 34252344]
77. Soria Campo A, Wang AI, Moholdt T, Berg J. Physiological and perceptual responses to single-player vs. Multiplayer exergaming. Front Sports Act Living 2022;4:903300 [FREE Full text] [doi: 10.3389/fspor.2022.903300] [Medline: 35784804]
78. Ketelhut S, Röglin L, Kircher E, Martin-Niedecken A, Ketelhut R, Hottenrott K, et al. The new way to exercise? Evaluating an innovative heart-rate-controlled exergame. Int J Sports Med 2022;43(1):77-82. [doi: 10.1055/a-1520-4742] [Medline: 34255326]
79. Marks DW, Rispen L, Calara G. Greater physiological responses while playing XBox kinect compared to nintendo wii. Int J Exerc Sci 2015;8(2) [FREE Full text] [doi: 10.13140/RG.2.1.3302.5446]
80. Monedero J, Murphy EE, O'Gorman DJ. Energy expenditure and affect responses to different types of active video game and exercise. PLoS One 2017;12(5):e0176213 [FREE Full text] [doi: 10.1371/journal.pone.0176213] [Medline: 28459835]
81. Dębska M, Polechoński J, Mynarski A, Polechoński P. Enjoyment and intensity of physical activity in immersive virtual reality performed on innovative training devices in compliance with recommendations for health. Int J Environ Res Public Health 2019;16(19):3673 [FREE Full text] [doi: 10.3390/ijerph16193673] [Medline: 31574911]
82. Stewart TH, Villaneuva K, Hahn A, Ortiz-Delatorre J, Wolf C, Nguyen R, et al. Actual vs. perceived exertion during active virtual reality game exercise. Front Rehabil Sci 2022;3:887740 [FREE Full text] [doi: 10.3389/fresc.2022.887740] [Medline: 36189005]
83. Wouda MF, Gaupseth JA, Bengtson EI, Johansen T, Brembo EA, Lundgaard E. Exercise intensity during exergaming in wheelchair-dependent persons with SCI. Spinal Cord 2023;61(6):338-344 [FREE Full text] [doi: 10.1038/s41393-023-00893-3] [Medline: 37012335]
84. Muñoz JE, Goncalves A, Cameirao MS, Bermúdez i Badia S, Gouveia ER. Measured and perceived physical responses in multidimensional fitness training through exergames in older adults. 2018 Presented at: 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games); 2018 September 05-07; Würzburg, Germany p. 1-4. [doi: 10.1109/vs-games.2018.8493433]
85. Munoz Cardona JE, Cameirao MS, Paulino T, Bermudez i Badia S, Rubio E. Modulation of physiological responses and activity levels during exergame experiences. 2016 Presented at: 2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES); 2016 September 07-09; Barcelona, Spain p. 1-8. [doi: 10.1109/vs-games.2016.7590353]
86. Liu H, Wang Z, Mousas C, Kao D. Virtual reality racket sports: virtual drills for exercise and training. 2020 Presented at: 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 2020 November 09-13; Porto de Galinhas, Brazil p. 566-576. [doi: 10.1109/ismar50242.2020.00084]
87. Martin Dantas EH, Bastos Andrade K, Silva Bezerra MR, Ramos Coelho R, Scudese E, Nunes Calvacante EM, et al. Neuromotor and functional adaptations in schoolchildren practicing exergames. 2022 Presented at: 2022 International Conference on Technology Innovations for Healthcare (ICTIH); 2022 September 14-16; Magdeburg, Germany p. 42-46. [doi: 10.1109/ictih57289.2022.10111946]
88. Han L, Pan Z, Zhang M, Tian F. A pleasurable persuasive model for e-fitness system. 2016 Presented at: 2016 International Conference on Cyberworlds (CW); 2016 September 28-30; Chongqing, China p. 89-96. [doi: 10.1109/cw.2016.20]
89. Julianjatsono R, Ferdiana R, Hartanto R. Development and evaluation of a low cost music based exergame using microsoft kinect. 2016 Presented at: 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE); 2016 October 05-06; Yogyakarta, Indonesia p. 1-4. [doi: 10.1109/iciteed.2016.7863261]
90. Martin-Niedecken AL, Rogers K, Turmo Vidal L, Mekler ED, Márquez Segura E. ExerCube vs. Personal trainer: evaluating a holistic, immersive, and adaptive fitness game setup. 2019 Presented at: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; 2019 May 02; New York, NY, USA p. 1-15. [doi: 10.1145/3290605.3300318]
91. Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. Pharmacoeconomics 1999;15(5):423-434. [doi: 10.2165/00019053-199915050-00001] [Medline: 10537960]

92. Lima A, Moreira MT, Ferreira MS, Parola V, Sampaio F, Nóbrega MDP, et al. Efficacy of the use of exergames in promoting the mental health of the elderly: protocol of a systematic review. JAL 2023;3(3):191-202. [doi: 10.3390/jal3030015]
93. Kinect for Windows. URL: https://learn.microsoft.com/en-us/windows/apps/design/devices/kinect-for-windows [accessed 2023-01-20]
94. The ExerCube, Sphery. URL: https://sphery.ch/exercube/ [accessed 2023-01-20]
95. Audioshield on Steam. URL: https://store.steampowered.com/app/412740/Audioshield/ [accessed 2022-08-30]
96. BOXVR on Steam. URL: https://store.steampowered.com/app/641960/BOXVR/ [accessed 2022-10-24]
97. OhShape on Steam. URL: https://store.steampowered.com/app/1098100/OhShape/ [accessed 2022-08-30]
98. Hot Squat on Steam. URL: https://store.steampowered.com/app/553590/Hot_Squat/ [accessed 2022-08-30]
99. Gong W, Zhang X, Gonzàlez J, Sobral A, Bouwmans T, Tu C, et al. Human pose estimation from monocular images: a comprehensive survey. Sensors (Basel) 2016;16(12):1966 [FREE Full text] [doi: 10.3390/s16121966] [Medline: 27898003]
100. Zhang S, Wei Z, Nie J, Huang L, Wang S, Li Z. A review on human activity recognition using vision-based method. J Healthc Eng 2017;2017:3090343 [FREE Full text] [doi: 10.1155/2017/3090343] [Medline: 29065585]
101. Colyer SL, Evans M, Cosker DP, Salo AIT. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. Sports Med Open 2018;4(1):24 [FREE Full text] [doi: 10.1186/s40798-018-0139-y] [Medline: 29869300]
102. Gabel M, Gilad-Bachrach R, Renshaw E, Schuster A. Full body gait analysis with Kinect. Annu Int Conf IEEE Eng Med Biol Soc 2012;2012:1964-1967. [doi: 10.1109/EMBC.2012.6346340] [Medline: 23366301]
103. Albert JA, Owolabi V, Gebel A, Brahms CM, Granacher U, Arnrich B. Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: a pilot study. Sensors (Basel) 2020;20(18):5104 [FREE Full text] [doi: 10.3390/s20185104] [Medline: 32911651]
104. Bari ASMH, Gavrilova ML. Artificial neural network based gait recognition using kinect sensor. IEEE Access 2019;7:162708-162722. [doi: 10.1109/access.2019.2952065]
105. Eichler N, Raz S, Toledano-Shubi A, Livne D, Shimshoni I, Hel-Or H. Automatic and efficient fall risk assessment based on machine learning. Sensors (Basel) 2022;22(4):1557 [FREE Full text] [doi: 10.3390/s22041557] [Medline: 35214471]
106. Passos P, Campos T, Diniz A. Quantifying the degree of movement dissimilarity between two distinct action scenarios: an exploratory approach with procrustes analysis. Front Psychol 2017;8:640 [FREE Full text] [doi: 10.3389/fpsyg.2017.00640] [Medline: 28503159]
107. Rosenberg M, Thornton AL, Lay BS, Ward B, Nathan D, Hunt D, et al. Development of a kinect software tool to classify movements during active video gaming. PLoS One 2016;11(7):e0159356 [FREE Full text] [doi: 10.1371/journal.pone.0159356] [Medline: 27442437]
108. Assad O, Hermann R, Lilla D, Mellies B, Meyer R, Shevach L, et al. Motion-based games for parkinson's disease patients. In: Anacleto JC, Fels S, Graham N, Kapralos B, Saif El-Nasr M, Stanley K, editors. Entertainment Computing - ICEC 2011. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011.
109. Zhou X, Zhu M, Pavlakos G, Leonardos S, Derpanis KG, Daniilidis K. MonoCap: monocular human motion capture using a CNN coupled with a geometric prior. IEEE Trans Pattern Anal Mach Intell 2019;41(4):901-914. [doi: 10.1109/TPAMI.2018.2816031] [Medline: 29993801]
110. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. OpenPose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 2021;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257] [Medline: 31331883]
111. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. BlazePose: on-device real-time body pose tracking. Computer Science Computer Vision and Pattern Recognition 2020. [doi: 10.48550/ARXIV.2006.10204]
112. Viswakumar A, Rajagopalan V, Ray T, Parimi C. Human gait analysis using OpenPose. 2019 Presented at: 2019 Fifth International Conference on Image Information Processing (ICIIP); 2019 November 15-17; Shimla, India p. 310-314. [doi: 10.1109/iciip47207.2019.8985781]
113. Zago M, Luzzago M, Marangoni T, De Cecco M, Tarabini M, Galli M. 3D tracking of human motion using visual skeletonization and stereoscopic vision. Front Bioeng Biotechnol 2020;8:181 [FREE Full text] [doi: 10.3389/fbioe.2020.00181] [Medline: 32195243]
114. Albuquerque P, Verlekar TT, Correia PL, Soares LD. A spatiotemporal deep learning approach for automatic pathological gait classification. Sensors (Basel) 2021;21(18):6202 [FREE Full text] [doi: 10.3390/s21186202] [Medline: 34577408]
115. Kidziński L, Yang B, Hicks JL, Rajagopal A, Delp SL, Schwartz MH. Deep neural networks enable quantitative movement analysis using single-camera videos. Nat Commun 2020;11(1):4054 [FREE Full text] [doi: 10.1038/s41467-020-17807-z] [Medline: 32792511]
116. Leu A, Ristić-Durrant D, Gräser A. A robust markerless vision-based human gait analysis system. 2011 Presented at: 2011 6th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI); 2011 May 19-21; Timisoara, Romania p. 415-420. [doi: 10.1109/saci.2011.5873039]
117. Mehrizi R, Peng X, Metaxas DN, Xu X, Zhang S, Li K. Predicting 3-D lower back joint load in lifting: a deep pose estimation approach. IEEE Trans Hum-Mach Syst 2019;49(1):85-94. [doi: 10.1109/thms.2018.2884811]

118. Trinidad-Fernández M, Cuesta-Vargas A, Vaes P, Beckwée D, Moreno FA, González-Jiménez J, et al. Human motion capture for movement limitation analysis using an RGB-D camera in spondyloarthritis: a validation study. Med Biol Eng Comput 2021;59(10):2127-2137 [FREE Full text] [doi: 10.1007/s11517-021-02406-x] [Medline: 34467447]
119. Wang T, Goto D, Manno M, Okada S, Shiozawa N, Ueta K. Movement coordination during forward and backward rope jumping: a relative phase study. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:4627-4630. [doi: 10.1109/EMBC46164.2021.9630017] [Medline: 34892245]
120. Faric N, Potts HWW, Hon A, Smith L, Newby K, Steptoe A, et al. What players of virtual reality exercise games want: thematic analysis of web-based reviews. J Med Internet Res 2019;21(9):e13833 [FREE Full text] [doi: 10.2196/13833] [Medline: 31538951]
121. Borrego A, Latorre J, Alcañiz M, Llorens R. Comparison of oculus rift and HTC vive: feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation. Games Health J 2018;7(3):151-156. [doi: 10.1089/g4h.2017.0114] [Medline: 29293369]
122. Jost TA, Drewelow G, Koziol S, Rylander J. A quantitative method for evaluation of 6 degree of freedom virtual reality systems. J Biomech 2019;97:109379. [doi: 10.1016/j.jbiomech.2019.109379] [Medline: 31679757]
123. Martin-Niedecken AL, Mekler ED. The ExerCube: participatory design of an immersive fitness game environment. In: Göbel S, Garcia-Agundez A, Tregel T, Ma M, Baalsrud Hauge J, Oliveira M, et al, editors. Serious Games. Cham: Springer International Publishing; 2018:263-275.
124. Caserman P, Liu S, Gobel S. Full-body motion recognition in immersive- virtual-reality-based exergame. IEEE Trans Games 2022;14(2):243-252. [doi: 10.1109/tg.2021.3064749]
125. OptiTrack. URL: http://optitrack.com/index.html [accessed 2023-02-03]
126. Vicon. URL: https://www.vicon.com/ [accessed 2023-02-03]
127. Nagymáté G, Kiss RM. Application of OptiTrack motion capture systems in human movement analysis: a systematic literature review. RIiM 2018;5(1):1. [doi: 10.17667/riim.2018.1/13]
128. Windolf M, Götzen N, Morlock M. Systematic accuracy and precision analysis of video motion capturing systems--exemplified on the vicon-460 system. J Biomech 2008;41(12):2776-2780. [doi: 10.1016/j.jbiomech.2008.06.024] [Medline: 18672241]
129. Chan JCP, Leung H, Tang JKT, Komura T. A virtual reality dance training system using motion capture technology. IEEE Trans Learn Technol 2011;4(2):187-195. [doi: 10.1109/tlt.2010.27]
130. Caserman P, Krug C, Göbel S. Recognizing full-body exercise execution errors using the teslasuit. Sensors (Basel) 2021;21(24):8389 [FREE Full text] [doi: 10.3390/s21248389] [Medline: 34960481]
131. Xsens. URL: https://www.movella.com/products/xsens [accessed 2023-01-20]
132. Measure 3D body movement using inertial measurement units (IMUs). Noraxon. URL: https://www.noraxon.com/our-products/3d-motion-capture-imu/ [accessed 2023-01-20]
133. Meet our haptic VR suit and glove with force feedback. Teslasuit. URL: https://teslasuit.io/ [accessed 2023-01-20]
134. Rana M, Mittal V. Wearable sensors for real-time kinematics analysis in sports: a review. IEEE Sensors J 2021;21(2):1187-1207. [doi: 10.1109/jsen.2020.3019016]
135. Lee J, Leadbetter R, Ohgi Y, Thiel D, Burkett B, James DA. Quantifying and assessing biomechanical differences in swim turn using wearable sensors. Sports Technology 2012;4(3-4):128-133. [doi: 10.1080/19346182.2012.725171]
136. Davey N, Anderson M, James DA. Validation trial of an accelerometer-based sensor platform for swimming. Sports Technol 2008;1(4-5):202-207. [doi: 10.1002/jst.59]
137. Büthe L, Blanke U, Capkevics H, Tröster G. A wearable sensing system for timing analysis in tennis. 2016 Presented at: 2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN); 2016 June 14-17; San Francisco, CA, USA p. 43-48. [doi: 10.1109/bsn.2016.7516230]
138. Takano K, Li KF. A multimedia tennis instruction system: tracking and classifying swing motions. IJSSC 2013;3(3):155-168. [doi: 10.1504/ijssc.2013.056406]
139. Clarke SR, Norman JM. Optimal challenges in tennis. J Oper Res Soc 2012;63(12):1765-1772. [doi: 10.1057/jors.2011.147]
140. Anand A, Sharma M, Srivastava R, Kaligounder L, Prakash D. Wearable motion sensor based analysis of swing sports. 2017 Presented at: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA); 2017 December 18-21; Cancun, Mexico p. 261-267. [doi: 10.1109/icmla.2017.0-149]
141. Ahmadi A, Rowlands D, James DA. Towards a wearable device for skill assessment and skill acquisition of a tennis player during the first serve. Sports Technol 2009;2(3-4):129-136. [doi: 10.1080/19346182.2009.9648510]
142. Benages Pardo L, Buldain Perez D, Orrite Uruñuela C. Detection of tennis activities with wearable sensors. Sensors (Basel) 2019;19(22):5004 [FREE Full text] [doi: 10.3390/s19225004] [Medline: 31744136]
143. Srivastava R, Patwari A, Kumar S, Mishra G, Kaligounder L, Sinha P. Efficient characterization of tennis shots and game analysis using wearable sensors data. : IEEE; 2015 Presented at: 2015 IEEE SENSORS; 2015 November 01-04; Busan, Korea (South). [doi: 10.1109/icsens.2015.7370311]
144. Lin J, Chang CW, Wang CH, Chi HC, Yi CW, Tseng YC, et al. Design and implement a mobile badminton stroke classification system. 2017 Presented at: 2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS); 2017 September 27-29; Seoul, Korea (South) p. 235-238. [doi: 10.1109/apnoms.2017.8094141]

Abbreviations

EE: energy expenditure
EEG: electroencephalography
HR: heart rate
IMU: inertial measurement unit
MET: metabolic equivalent of task
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RGB: red-green-blue
RGB-D: red-green-blue-depth
VO2: oxygen uptake
VR: virtual reality

Edited by T Leung, A Coristine; submitted 24.08.23; peer-reviewed by N Spicher; comments to author 07.01.24; revised version received 28.03.24; accepted 03.09.24; published 05.11.24.

Please cite as:
Dill S, Müller PN, Caserman P, Göbel S, Hoog Antink C, Tregel T
Sensing In Exergames for Efficacy and Motion Quality: Scoping Review of Recent Publications
JMIR Serious Games 2024;12:e52153
URL: https://games.jmir.org/2024/1/e52153
doi:10.2196/52153
PMID:

Interactive Serious Game to Teach Basic Life Support Among Schoolchildren in Brazil: Design and Rationale

Uri Adrian Prync Flato[1,2], MD, PhD; Emilio José Beffa dos Santos[3], MD; Isabella Bispo Diaz T Martins[3], MD; Vinicius Gazin Rossignoli[3], BM; Thais Dias Midega[1], MD; Lucas Kallas-Silva[2], MD; Ricardo Ferreira Mendes de Oliveira[2], MD; Adriana do Socorro Lima Figueiredo Flato[4], BMS, MSc; Mario Vicente Guimarães[1], MD, PhD; Hélio Penna Guimarães[1], MD, PhD

1

2

3

4

Corresponding Author:
Uri Adrian Prync Flato, MD, PhD

(JMIR Serious Games 2024;12:e55333) doi:10.2196/55333

KEYWORDS
cardiopulmonary resuscitation; basic life support; serious game; CPR training; usability; cardiopulmonary; emergency; life support; CPR; training; education; game; gaming; educational; resuscitation; survey; satisfaction; SUS; user experience; System Usability Scale

Introduction

Cardiovascular diseases are among the leading causes of death and morbidity worldwide [1,2]. Cardiopulmonary resuscitation (CPR) and early defibrillation increase survival chances [3]. Serious games (SGs) are tools used to enhance the learning process through entertainment. Current strategies focus on teaching CPR to the community and schoolchildren [4].

While other games exist for teaching basic life support (BLS), no studies have validated these for children in low- to middle-income settings. The SG Children Save Hearts teaches the 5 resuscitation steps per International Liaison Committee on Resuscitation (ILCOR) guidelines. Before use in schools, it requires a formal usability assessment by game developers and health care professionals to ensure ease of use, learning, and interaction.

The primary objective was to evaluate the usability of the SG Children Save Hearts among health care and IT professionals using the System Usability Scale (SUS) [5], a validated usability assessment tool.

Methods

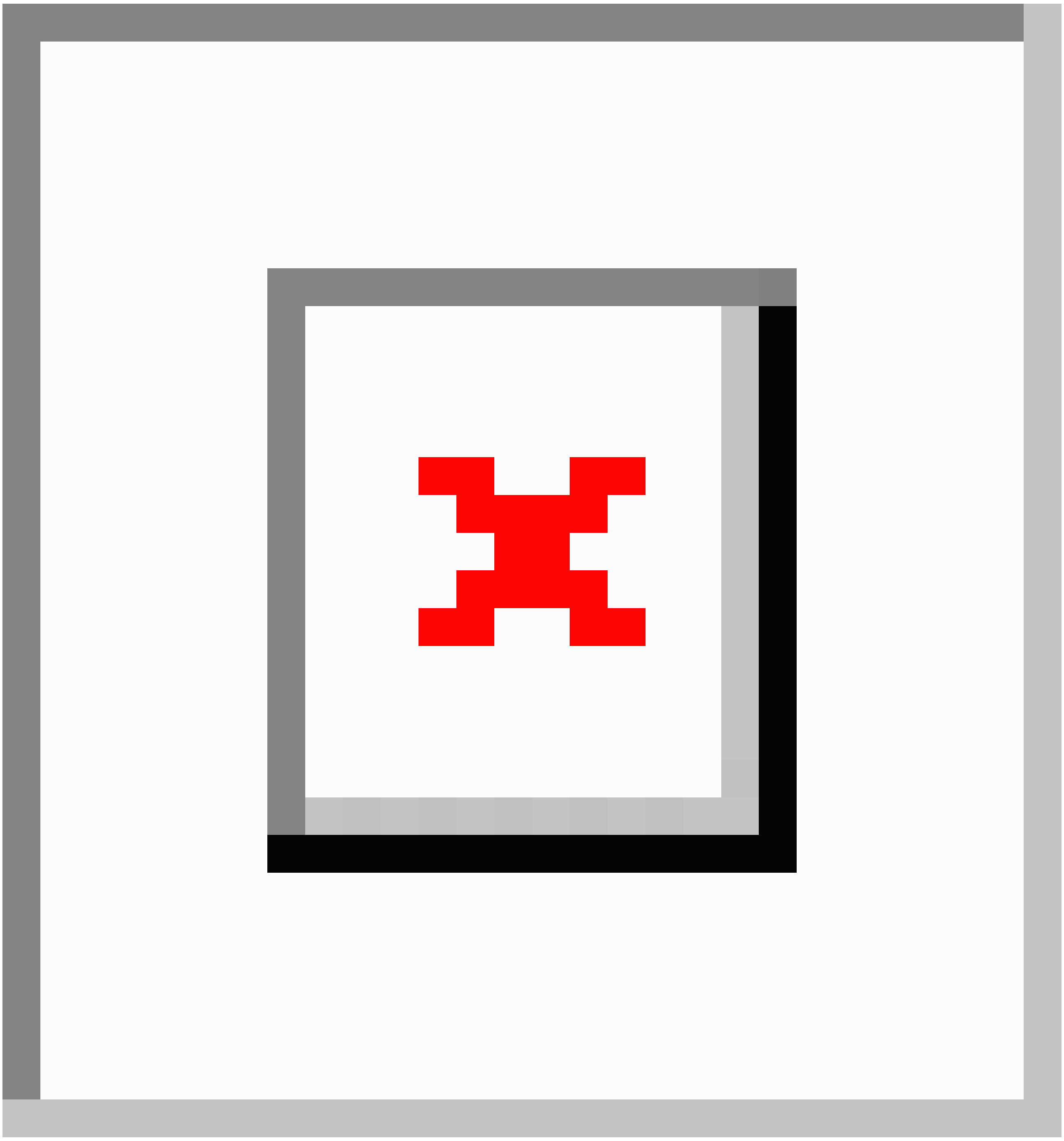
Ethical Considerations

The study protocol was approved by the ethics committee of the University of Marília, Brazil (CAAE: 57160121400005496). All participants signed an informed consent form.

Study Design

We used a nonprobabilistic casual sample to include IT and health care professionals. The usability test was conducted in August 2022 in the university's IT department after a 10-minute lecture on the SG's purpose. Children Save Hearts was developed on the Microsoft Smile Game Builder platform and implemented on Microsoft Windows (versions 7 to 11), targeting schoolchildren aged 7 to 17 years. The script and storytelling are based on the ILCOR 2020 guidelines. The game uses a joystick and simple commands to simplify the user experience (Multimedia Appendix 1). The design and testing process is illustrated in Figure 1.

Figure 1. Design of a serious game for teaching cardiopulmonary resuscitation among schoolchildren. BLS: basic life support; MVP: minimum viable product; SAMU: Serviço de Atendimento Móvel de Urgência; SUS: System Usability Scale.



After completing the SG, participants answered a 10-question survey on its usability using a Likert-type scale. The final grade was converted to a 0 to 100 scale (Multimedia Appendix 2). A grade above 70 was considered acceptable to proceed to a minimum viable product. A sample size of 17 users was required, based on a 10% estimated probability of encountering an interface error, to identify 85% of the problems.

Statistical Analysis

Categorical variables are presented as absolute and relative frequencies. Continuous variables are presented as medians with IQRs. Normality was assessed with the Shapiro-Wilk test. Comparisons were made between IT and health care professionals. Continuous variables were compared using an independent 2-tailed *t* test (for normal distributions) or the Mann-Whitney *U* test (for nonnormal distributions). All analyses were performed using R (version 4.1.0; R Foundation for Statistical Computing).

Results

Children Save Hearts was used by 17 volunteers with a median age of 22 (IQR 20 - 26) years; 8 (47%) were male. Regarding professional training, 8 (47%) held a bachelor's degree in IT and 9 (53%) were health care professionals. All participants played the game and answered the questionnaire. The median SUS score was 75 (IQR 72.5 - 87.5; Table 1). Questions 2 and 4 had the lowest median scores, and questions 7 and 9 had the highest. Health care professionals gave higher grades to all 5 domains when compared to IT professionals. The average time spent in the game was 3.2 (SD 0.4) minutes.

Table . Participant characteristics and System Usability Scale (SUS) scores by profession.

	All (N=17)	IT professionals (n=8)	HCPs[a] (n=9)	*P* value
Participant characteristics				
Age (years), median (IQR)	22.00 (20.00-26.00)	21.00 (18.75-22.75)	24.00 (22.00-27.00)	.07[b]
Male sex, n (%)	8 (47)	4 (50)	4 (44)	>.99[c]
SUS scores, median (IQR)				
Question 1	3.00 (3.00-4.00)	3.00 (2.00-3.25)	4.00 (3.00-4.00)	.08[b]
Question 2	2.00 (2.00-3.00)	2.00 (1.75-3.00)	3.00 (2.00-3.00)	.47[b]
Question 3	3.00 (3.00-4.00)	3.00 (3.00-4.00)	4.00 (3.00-4.00)	.36[b]
Question 4	2.00 (2.00-3.00)	2.00 (1.75-2.25)	2.00 (2.00-3.00)	.61[b]
Question 5	3.00 (3.00-4.00)	3.00 (3.00-3.00)	4.00 (3.00-4.00)	.07[b]
Question 6	4.00 (3.00-4.00)	3.00 (2.00-3.25)	4.00 (4.00-4.00)	.002[b]
Question 7	3.00 (3.00-4.00)	3.00 (3.00-3.00)	4.00 (4.00-4.00)	.009[b]
Question 8	4.00 (3.00-4.00)	3.00 (2.75-3.25)	4.00 (4.00-4.00)	.02[b]
Question 9	4.00 (3.00-4.00)	3.00 (3.00-3.25)	4.00 (4.00-4.00)	.03[b]
Question 10	3.00 (3.00-4.00)	3.50 (3.00-4.00)	3.00 (3.00-4.00)	.68[b]
Analysis, median (IQR)				
Total score	30.00 (29.00-35.00)	27.50 (25.50-30.50)	34.00 (30.00-35.00)	.05[d]
SUS grade	75.00 (72.50-87.50)	68.75 (63.75-76.25)	85.00 (75.00-87.50)	.05[d]
Score for ease of learning domain (Q3, Q4, Q7, Q10)	3.25 (2.75-3.25)	3.00 (2.50-3.25)	3.25 (2.75-3.50)	.54[d]
Score for efficiency domain (Q5, Q6, Q8)	3.67 (3.00-4.00)	3.00 (2.50-3.17)	4.00 (3.67-4.00)	.02[b]
Score for ease of memorization domain (Q2)	2.00 (2.00-3.00)	2.00 (1.75-3.00)	3.00 (2.00-3.00)	.47[b]
Score for minimization of errors domain (Q6)	4.00 (3.00-4.00)	3.00 (2.00-3.25)	4.00 (4.00-4.00)	.002[b]
Score for satisfaction domain (Q1, Q4, Q9)	2.67 (2.67-3.33)	2.67 (2.33-2.83)	3.33 (2.67-3.67)	.08[b]

[a]HCP: health care professional.

[b]Mann-Whitney *U* test.

[c]Fisher test.

[d]2-sample Student *t* test.

Discussion

We developed an SG, Children Save Hearts, to teach BLS to schoolchildren. When tested on 17 IT and health care professionals, it achieved an overall mean SUS score of 75, suitable for implementation.

Novel technologies like virtual reality (VR) have been successfully used in Europe to teach CPR to schoolchildren [6]. However, transferring this technology to limited-income countries faces challenges, such as language barriers, VR device acquisition, cultural context, and technical support. Previous SGs for teaching CPR were developed and tested in high-income countries [7]. Educational strategies for teaching CPR in limited-income countries have focused on health care professionals and students [8], not schoolchildren, highlighting a significant gap in the literature. This is the first SG developed in Brazil in Portuguese for schoolchildren.

Our study has some limitations. First, we had a small sample size due to insufficient data to calculate sample size in usability tests and financial constraints in contracting a software house. Continuous usability monitoring with larger sample sizes is needed to maintain external validation. Further studies should target schoolchildren to assess the effectiveness of teaching BLS in schools and explore user experiences to gain insights into how users feel about SGs.

Active teaching methods are crucial to improving survival rates and translating accessible knowledge into practice. Programs like Kids Save Lives [9] and World Restart a Heart Day [10] are teaching schoolchildren that CPR is vital. Despite some usability issues, the game is adequate for testing in schoolchildren.

Acknowledgments

We would like to thank the University of Marília for the use of their facilities and resources. We also thank the participants for their time and dedication to our study.

Data Availability

Due to the nature and design of this study, the raw data supporting our findings cannot be made publicly available. The data set contains detailed information that could compromise the privacy and confidentiality of the participants involved. Protecting participant confidentiality is a priority for us, and thus, in adherence to ethical considerations and participant consent agreements, we are unable to deposit our data in publicly accessible repositories or present it within the manuscript or supplementary files.

Authors' Contributions

UAPF contributed to conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing (original draft), visualization, supervision, project administration, and funding acquisition. EJBdS contributed to conceptualization, validation, investigation, resources, writing (review and editing), visualization, and supervision. IBDTM contributed to software, validation, formal analysis, data curation, and writing (review and editing). VGR contributed to software, validation, investigation, data curation, and writing (review and editing). TDM contributed to methodology, validation, investigation, resources, data curation, writing (review and editing), and visualization. LK–S contributed to writing (original draft, review, and editing) and supervision. RFMdO contributed to writing (original draft, review, and editing) and supervision. AdSLFF contributed to validation, formal analysis, investigation, resources, data curation, and writing (original draft). MVG contributed to methodology, software, validation, investigation, and writing (review and editing). HPG contributed to conceptualization, methodology, writing (original draft), and project administration.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Serious game screenshots and setup.
[PPTX File, 8718 KB - games_v12i1e55333_app1.pptx]

Multimedia Appendix 2
System Usability Scale (SUS) questions and domains.
[DOCX File, 15 KB - games_v12i1e55333_app2.docx]

References

1. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics-2023 update: a report from the American Heart Association. Circulation 2023 Feb 21;147(8):e93-e621. [doi: 10.1161/CIR.0000000000001123] [Medline: 36695182]
2. Oliveira GD, Brant LCC, Polanczyk CA, et al. Cardiovascular statistics - Brazil 2020. Arq Bras Cardiol 2020 Sep;115(3):308-439. [doi: 10.36660/abc.20200812] [Medline: 33027364]
3. Wyckoff MH, Greif R, Morley PT, et al. 2022 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science With Treatment Recommendations: summary from the basic life support; advanced life support; pediatric life support; neonatal life support; education, implementation, and teams; and first aid task forces. Circulation 2022 Dec 20;146(25):e483-e557. [doi: 10.1161/CIR.0000000000001095] [Medline: 36325905]
4. Berglund A, Jaarsma T, Berglund E, Strömberg A, Klompstra L. Understanding and assessing gamification in digital healthcare interventions for patients with cardiovascular disease. Eur J Cardiovasc Nurs 2022 Aug 29;21(6):630-638. [doi: 10.1093/eurjcn/zvac048] [Medline: 35709297]
5. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]

6. Yeung J, Kovic I, Vidacic M, et al. The school Lifesavers study—a randomised controlled trial comparing the impact of Lifesaver only, face-to-face training only, and Lifesaver with face-to-face training on CPR knowledge, skills and attitudes in UK school children. Resuscitation 2017 Nov;120:138-145. [doi: 10.1016/j.resuscitation.2017.08.010] [Medline: 28974344]
7. Cheng P, Huang Y, Yang P, et al. The effects of serious games on cardiopulmonary resuscitation training and education: systematic review with meta-analysis of randomized controlled trials. JMIR Serious Games 2024 Feb 6;12:e52990. [doi: 10.2196/52990] [Medline: 38319697]
8. Siqueira TV, Nascimento JDSG, de Oliveira JLG, Regino DDSG, Dalri MCB. The use of serious games as an innovative educational strategy for learning cardiopulmonary resuscitation: an integrative review. Rev Gaucha Enferm 2020;41:e20190293. [doi: 10.1590/1983-1447.2020.20190293] [Medline: 33111759]
9. Nakagawa NK, Salles IC, Semeraro F, Böttiger BW. KIDS SAVE LIVES: a narrative review of associated scientific production. Curr Opin Crit Care 2021 Dec 1;27(6):623-636. [doi: 10.1097/MCC.0000000000000872] [Medline: 34495875]
10. Horriar L, Rott N, Semeraro F, Böttiger BW. A narrative review of European public awareness initiatives for cardiac arrest. Resusc Plus 2023 Jun;14:100390. [doi: 10.1016/j.resplu.2023.100390] [Medline: 37128626]

Abbreviations

BLS: basic life support
CPR: cardiopulmonary resuscitation
ILCOR: International Liaison Committee on Resuscitation
SG: serious game
SUS: System Usability Scale
VR: virtual reality

Edited by T Leung; submitted 09.12.23; peer-reviewed by I Straand, N Cavalcante, V Beninca; revised version received 14.05.24; accepted 12.07.24; published 09.10.24.

Please cite as:
Flato UAP, Beffa dos Santos EJ, Bispo Diaz T Martins I, Rossignoli VG, Dias Midega T, Kallas-Silva L, Ferreira Mendes de Oliveira R, do Socorro Lima Figueiredo Flato A, Vicente Guimarães M, Penna Guimarães H
Interactive Serious Game to Teach Basic Life Support Among Schoolchildren in Brazil: Design and Rationale
JMIR Serious Games 2024;12:e55333
URL: https://games.jmir.org/2024/1/e55333
doi:10.2196/55333

Research Letter

Health Care Professional–Supported Co-Design of a Mime Therapy–Based Serious Game for Facial Rehabilitation

Daniel Lima Sousa[1], MSc; Silmar Teixeira[1], PhD; José Everton Fontenele[1], MSc; Renato Mendes Santos[1], MSc; Leynilson Pereira[1], MSc; Rodrigo Baluz[1,2], PhD; Victor Hugo Bastos[1], PhD; Ariel Soares Teles[1,3], PhD

[1]Parnaíba Delta Federal University, Parnaíba, Brazil

[2]State University of Piauí, Campus Parnaíba, Parnaíba, Brazil

[3]Federal Institute of Maranhão, Campus Araioses, Araioses, Brazil

Corresponding Author:
Ariel Soares Teles, PhD
Federal Institute of Maranhão, Campus Araioses
Rua José de Alencar, S/N
Bairro Cumprida
Araioses, 65570-000
Brazil
Phone: 55 86995501313
Email: ariel.teles@ifma.edu.br

Abstract

This research letter presents the co-design process for RG4Face, a mime therapy–based serious game that uses computer vision for human facial movement recognition and estimation to help health care professionals and patients in the facial rehabilitation process.

(JMIR Serious Games 2024;12:e52661) doi:10.2196/52661

KEYWORDS

serious game; serious games; facial recognition; face estimation; computer vision; facial rehabilitation; face; rehabilitation; physiotherapy; mime therapy; co-design; human face estimation; facial palsy; facial paralysis; motor rehabilitation; exergame; physiotherapists; psychologists; participation

Introduction

Facial paralysis is a consequence of damage or injury to the facial nerve, resulting in functional impairments. A challenge of rehabilitation through exercise repetition is maintaining patients' engagement and motivation in the intensive and repetitive execution of the exercises necessary for successful rehabilitation [1]. Repetitive and intensive movements are recommended for progress in treatment [2], and the variety of movements has significant effects on patient recovery [3].

In motor rehabilitation, exergames—serious games that require physical exercise to play—add fun to exercises and allow patients to forget about their condition and focus on the game [4]. Studies conducted with games for motor rehabilitation have achieved promising results [5] on patient motivation and engagement [4]. This study aimed to co-design *RG4Face*—an exergame for facial rehabilitation.

Methods

Ethical Considerations

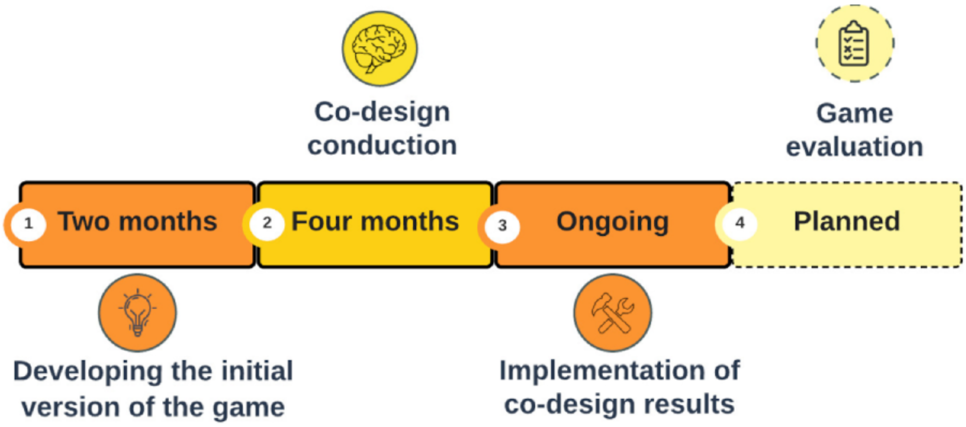
This study was approved by the Research Ethics Committee of Universidade Federal do Delta do Parnaíba (5.632.311). The first author (DLS) provided explicit consent for use of his image in Multimedia Appendices 1 and 2.

Study Design

To develop *RG4Face*, a co-design procedure (Figure 1) was conducted with physiotherapists (n=16) and psychologists (n=5; Multimedia Appendix 3) to obtain the necessary knowledge on the game requirements.

In the first stage, a version of the game was developed with an initial idea (Multimedia Appendix 1). In the second, we recruited physiotherapists and psychologists to participate in co-design meetings (August to November 2022) and answer a questionnaire. We then presented the game to the participants and allowed them to make suggestions. The prototype was essential to encouraging participation during meetings. In total, 5 meetings were held—4 with physiotherapists and 1 with psychologists. The main activities of the meetings were brainstorming sessions, in which the generation of game requirements was encouraged for their incorporation into visual elements, gamification, and game mechanics. Meeting results allowed for the creation of a list of requirements. As a third stage, we are concluding the implementation of *RG4Face* based on the produced requirements. The game code was implemented in JavaScript to provide new features for facial rehabilitation via the Rehabilite Game platform [6].

Figure 1. Co-design timeline.



Results

Per its initial conception, *RG4Face* uses computer vision (via a camera) for capturing, recognizing, and estimating human facial movements. The game prototype was implemented via the MediaPipe face mesh [7] to enable the recognition and use of 1 movement (eg, raising eyebrows; ie, frontal muscle) to control game elements. The game involves a spaceship moving horizontally across the bottom of the captured video window and firing a projectile when face movement is detected. The main objective is to hit triangles that randomly appear on the player's face.

Table 1 presents participants' suggestions during co-design, game requirements, and rationales.

RG4Face is in the testing phase and, prior to evaluations, can recognize 6 movements used in mime therapy to improve facial muscle strength and mobility (Multimedia Appendix 2). To implement the recognition of these movements, MediaPipe was used [7]. The face mesh model allows for the real-time tracking of 468 3D landmarks on the human face that represent important facial features (eg, eyes, eyebrows, nose, and mouth). Distances between landmarks are calculated to recognize movements.

RG4Face provides a mirror therapy feature [8], which can mirror the healthy side of the face to create a visual illusion that can help reduce pain and improve function. *RG4Face* allows for parameter adjustment on the Rehabilite Game platform. Health care professionals can choose specific game mechanics for each rehabilitation case, thereby customizing the game according to patients' needs and difficulties.

Table 1. Functional and nonfunctional game requirements from the co-design procedure.

Participants' suggestions	Refined requirement	Rationale
• Improve the game scenario • Improve the representation of the ship and projectiles • Choose attractive colors and contrasts	• Improve game colors and elements: border, ship, projectiles, and collision	• Enable the game to become more attractive and stimulating
• Include levels with difficulty levels	• Provide difficulty levels	• Gamification for each level, depending on the patient's condition
• Provide an option of mirror therapy for the game	• Implement a mirror therapy simulation	• Patients with Bell palsy can benefit from it
• Implement better game mechanics for rewards • Promote progression in the game	• Create a scoring and bonus system	• Increase patients' adherence to and engagement with treatment
• Movement sensitivity must be customized according to the patient's degree of disability	• Implementation of sensitivity levels for motion recognition	• The level of sensitivity respects the movement capacity of each patient
• Create metrics on the game platform to monitor the rehabilitation process	• Provide in-game metrics	• They are interesting for the health care professional to follow the patient's progress
• To avoid causing botheration to some types of patients, the sound should be optional	• Allow game sound to be optional (ie, turn off the sound)	• The sound may be unnecessary for some patients
• Consider visual acuity of the players • The game scenario should be full screen	• Make game screen full, automatically adjusting to the aspect ratio	• Game elements should be clearly visible
• Head movement should not influence the game • Calibration is essential to avoid false positives and false negatives of movements	• Perform a prior calibration of the player's face	• Adjustment of the distance between player's face and screen, in addition to improving movement recognition

Discussion

We co-designed a serious game for facial rehabilitation that represents a potential new approach to improving patients' adherence to facial rehabilitation. The co-design procedure allowed stakeholders to participate in defining game requirements, thereby empowering the tool to meet the needs and expectations of patients and be more engaging and motivating.

Although there are studies that focus on games for rehabilitating specific parts of the face (eg, eyes [9] and mouth [10]), to our knowledge, no serious game for facial rehabilitation has been proposed that can recognize the face movements used in mime therapy. This study proposes the first such exergame.

Our results demonstrate that the co-design approach was effective for creating a serious game with the potential to meet patients' needs. We plan to evaluate the game with health care professionals, healthy participants, and patients with facial paralysis.

Acknowledgments

This work was supported by the Federal Institute of Maranhão, Coordination for the Improvement of Higher Education Personnel (CAPES; finance code 001), National Council for Scientific and Technological Development (CNPq; grants 308736/2022-2 and 308059/2022-0), and State Funding Agency of Maranhão (FAPEMA; grant UNIVERSAL-06123/22).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Video presentation with the game prototype before the co-design procedure.

[MP4 File (MP4 Video), 7354 KB - games_v12i1e52661_app1.mp4]

Multimedia Appendix 2

Video presentation with the game after implementing requirements from the co-design procedure.
[MP4 File (MP4 Video), 55743 KB - games_v12i1e52661_app2.mp4]

Multimedia Appendix 3
Demographic characteristics of participants.
[DOCX File , 15 KB - games_v12i1e52661_app3.docx]

References

1. Khan AJ, Szczepura A, Palmer S, Bark C, Neville C, Thomson D, et al. Physical therapy for facial nerve paralysis (Bell's palsy): an updated and extended systematic review of the evidence for facial exercise therapy. Clin Rehabil 2022 Nov;36(11):1424-1449 [FREE Full text] [doi: 10.1177/02692155221110727] [Medline: 35787015]
2. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. Lancet Neurol 2009 Aug;8(8):741-754. [doi: 10.1016/S1474-4422(09)70150-4] [Medline: 19608100]
3. Lang CE, MacDonald JR, Gnip C. Counting repetitions: an observational study of outpatient therapy for people with hemiparesis post-stroke. J Neurol Phys Ther 2007 Mar;31(1):3-10. [doi: 10.1097/01.npt.0000260568.31746.34] [Medline: 17419883]
4. Ambros-Antemate JF, Beristain-Colorado MDP, Vargas-Treviño M, Gutiérrez-Gutiérrez J, Hernández-Cruz PA, Gallegos-Velasco IB, et al. Software engineering frameworks used for serious games development in physical rehabilitation: systematic review. JMIR Serious Games 2021 Nov 11;9(4):e25831 [FREE Full text] [doi: 10.2196/25831] [Medline: 34762052]
5. Chen Y, Abel KT, Janecek JT, Chen Y, Zheng K, Cramer SC. Home-based technologies for stroke rehabilitation: a systematic review. Int J Med Inform 2019 Mar;123:11-22 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.12.001] [Medline: 30654899]
6. Baluz R, Teles A, Fontenele JE, Moreira R, Fialho R, Azevedo P, et al. Motor rehabilitation of upper limbs using a gesture-based serious game: evaluation of usability and user experience. Games Health J 2022 Jun;11(3):177-185. [doi: 10.1089/g4h.2022.0005] [Medline: 35294849]
7. Lugaresi C, Tang J, Nash H, McClanahan C, Uboweja E, Hays M, et al. MediaPipe: a framework for building perception pipelines. arXiv Preprint posted online on Jun 14, 2019. [FREE Full text] [doi: 10.48550/arXiv.1906.08172]
8. Gandhi DB, Sterba A, Khatter H, Pandian JD. Mirror therapy in stroke rehabilitation: current perspectives. Ther Clin Risk Manag 2020 Feb 7;16:75-85 [FREE Full text] [doi: 10.2147/TCRM.S206883] [Medline: 32103968]
9. Tasneem T, Shome A, Hossain SKA. A gaming approach in physical therapy for facial nerve paralysis patient. In: 16th Int'l Conf. Computer and Information Technology. New York, NY: Institute of Electrical and Electronics Engineers; 2014:345-349.
10. Wang YX, Lo LY, Hu MC. Eat as much as you can: a Kinect-based facial rehabilitation game based on mouth and tongue movements. In: MM '14: Proceedings of the 22nd ACM International Conference on Multimedia. New York, NY: Association for Computing Machinery; Nov 2014:743-744.

Edited by T Leung; submitted 11.09.23; peer-reviewed by A Junus, Y Asada; comments to author 03.11.23; revised version received 09.11.23; accepted 29.12.23; published 24.01.24.

Please cite as:
Sousa DL, Teixeira S, Fontenele JE, Santos RM, Pereira L, Baluz R, Bastos VH, Teles AS
Health Care Professional–Supported Co-Design of a Mime Therapy–Based Serious Game for Facial Rehabilitation
JMIR Serious Games 2024;12:e52661
URL: https://games.jmir.org/2024/1/e52661
doi:10.2196/52661
PMID:38265856

Developing an Automated Virtual Reality Therapy for Improving Positive Self-Beliefs and Psychological Well-Being (Phoenix VR Self-Confidence Therapy): Tutorial

Laina Rosebrock[1,2], PhD; Jason Freeman[1], MA; Aitor Rovira[1,2], PhD; Andre Lages Miguel[1], MSc; Rupert Ward[1], PhD; Matthew Bousfield[1], BA; Ludovic Riffiod[1]; Roya Kamvar[3], BA; Thomas Kabir[1,3], PhD; Felicity Waite[1,2], DClinPsych; Daniel Freeman[1,2], DClinPsy

1

2

3

Corresponding Author:
Laina Rosebrock, PhD

Abstract

Virtual reality (VR) is an immersive technology in which delivery of psychological therapy techniques can be automated. Techniques can be implemented similarly to real-world delivery or in ways that are not possible in the real world to enhance efficacy. The potential is for greater access for patients to effective therapy. Despite an increase in the use of VR for mental health, there are few descriptions of how to build and design automated VR therapies. We describe the development of Phoenix VR Self-Confidence Therapy, designed to increase positive self-beliefs in young patients diagnosed with psychosis in order to improve psychological well-being. A double-diamond, user-centered design process conducted over the course of 18 months was used, involving stakeholders from multiple areas: individuals with lived experience of psychosis, clinical psychologists, treatment designers, and VR software developers. Thirteen meetings were held with young patients diagnosed with psychosis to increase the understanding and improve the assessment of positive self-beliefs, help design the scenarios for implementing therapeutic techniques, and conduct user testing. The resulting Phoenix therapy is a class I United Kingdom Conformity Assessed (UKCA)–certified medical device designed to be used on the standalone Meta Quest 2 (Meta Platforms) headset. Phoenix aims to build up 3 types of positive self-beliefs that are connected to psychological well-being. In a community farm area, tasks are designed to increase a sense of mastery and achievement ("I can make a difference"); in a TV studio, users complete an activity with graded levels of difficulty to promote success in the face of a challenge ("I can do this"); and in a forest by a lake, activities are designed to encourage feelings of pleasure and enjoyment ("I can enjoy things"). Phoenix is delivered over the course of approximately 6 weekly sessions supported by a mental health provider. Patients can take the headsets home to use in between sessions. Usability testing with individuals with lived experience of psychosis, as well as patients in the National Health Service (aged 16‑26 years), demonstrated that Phoenix is engaging, easy to use, and has high levels of satisfaction.

(JMIR Serious Games 2024;12:e51512) doi:10.2196/51512

KEYWORDS

virtual reality; VR; psychosis; cognitive behavioural therapy; psychological well-being; design process; self-beliefs; psychological therapy; real-world; efficacy; well-being; mental health; participant; stakeholder; user; Phoenix VR Self-Confidence Therapy

Introduction

Virtual reality (VR) has the potential to automate the delivery of powerful psychological therapy techniques. This includes direct translation of evidence-based techniques that are used in traditional face-to-face therapies as well as techniques that are not possible or difficult in the real world, potentially enhancing efficacy. The automation of such techniques means therapies can be more widely available in clinical services. However, automation of psychological therapy in VR requires careful execution to ensure there is precise translation of effective techniques that link to clear clinical targets. In order for an automated VR therapy to be effective, the design process must be methodically and rigorously conducted.

We previously published 2 papers on the design process of gameChange [1,2], a VR therapy targeting agoraphobic avoidance in patients with psychosis. Most VR therapies have been used to target avoidance of feared situations and reduce anxiety [3]. We describe in this paper the development of a new VR therapy with a different clinical target: the improvement of positive self-beliefs and psychological well-being. Phoenix VR Self-Confidence Therapy uniquely combines techniques derived from existing therapies with techniques that differ from those used in the real world to target key processes that impact psychological well-being. Additionally, since the development of gameChange, there have been improvements in VR hardware and programming capabilities. Primarily, this has included the use of standalone headsets (ie, a gaming laptop is not required to run the program, nor external cameras) such as the Meta Quest 2 (Meta Platforms). Standalone headsets are more affordable and accessible for patients and do not require a staff member to set them up and be present throughout use.

VR for Psychosis

VR allows for the creation of 3D computer simulations of environments that can be interacted with in a similar way to the real world. Virtual environments elicit similar responses to their real-world counterparts [4]. This makes it possible for people to enter environments in the virtual world and learn new things that can then apply to real-world situations. VR has a number of advantages. It allows for full control over the environment, including specific stimuli. Scenarios can be repeated and entered gradually. People are more willing to enter virtual scenarios precisely because they know the situations are not real; indeed, one study indicated that three-quarters of people would rather participate in VR-based exposures than real-world exposures [5]. Lastly, VR therapies can be highly engaging; therefore, they are more appealing for patients [6].

The use of VR among patients with psychosis has included aided therapies (ie, using VR alongside traditional face-to-face therapy). For example, Pot-Kolder and colleagues [7] compared participants who received VR-based cognitive behavioral therapy (VR-CBT) to those in a wait-list control group and found that those who received VR-CBT had reductions in paranoid ideation and anxiety. Patients received 16 sessions with a psychologist, and VR was used together with face-to-face therapy: psychologists provided in-the-moment feedback and guidance on how to respond to situations differently while the patient was in the headset. Another study examined the initial feasibility and efficacy of VR therapy to target the impact of negative symptoms (eg, poor motivation, blunted affect, and difficulties experiencing pleasure). Patients received 12 sessions of VR therapy using principles based on existing evidence-based therapies (CBT and cognitive remediation). Environments were simulated real-world scenes (eg a factory, pub, and rooms that may be found in one’s house, such as a TV room). The VR therapy was found to be acceptable and engaging, with patients indicating increased levels of goal attainment [8]. As with the Pot-Kolder et al [7] study, all sessions were completed with a graduate-level psychologist, and tasks in VR were discussed directly with the therapist in each session.

A significant advantage of VR is that therapy can be automated and psychological principles embedded within the VR program, including through the use of a virtual coach. This allows the therapy to be supported by a range of staff members, including peer support workers and graduate mental health workers, thereby making it accessible to many more people. An automated VR therapy developed by our team was shown to significantly reduce fear of heights over the course of 2 weeks [9]. In the largest trial of VR with patients to date, an automated treatment developed by our team (gameChange) was found to be effective at reducing agoraphobic avoidance in patients with psychosis, particularly for people with severe levels of avoidance [10,11]. VR scenarios were simulated everyday situations, such as a physician’s clinic, bus, and café. Both gameChange and Fear of Heights include a virtual coach who describes the treatment principles and guides patients through the program. Another study conducted by our team compared an automated VR cognitive intervention with automated VR mental relaxation to reduce paranoia and persecutory beliefs in patients with psychosis. Both therapies included 4 sessions including anxiety reduction techniques. The cognitive intervention was delivered directly in VR by a virtual coach while the mental relaxation intervention included audio files of relaxation techniques. Both interventions were supported by a clinical or assistant psychologist and produced large reductions in paranoia [12].

The studies mentioned above all use exposure-based techniques to target anxiety, social avoidance, and paranoia, which are typical treatment targets in traditional face-to-face therapies for psychosis. They focus on encouraging patients to enter simulations of real-world scenes such as a café, shopping center, or pub in order to learn that they are safe and capable of coping with anxiety. However, there are other important treatment targets for patients with psychosis that may benefit from interventions that encourage users to enter situations they may not encounter in everyday life.

Positive Self-Beliefs and Psychological Well-Being

Our new treatment focused on a different clinical issue in psychosis: low self-confidence and psychological well-being. Patients with psychosis can be self-critical, see themselves as inferior to others, and feel a sense of failure [13]. Psychological well-being is often low [14]. Negative self-beliefs exacerbate the experience of other psychotic symptoms, including paranoia, voice-hearing, and lack of motivation. The result is often a withdrawal from everyday activities. Withdrawal further increases the sense of defeat and means there are fewer opportunities for people to have positive experiences, thereby lowering self-confidence even more. A survey of 1800 patients with schizophrenia found that increasing self-confidence is 1 of the top 3 treatment goals [14].

Psychosis often first presents in late adolescence or the early twenties [15,16]. This is an age at which many people are in school or university, beginning employment, and experiencing other significant life events (eg, moving out of the family home, entering into a significant romantic relationship, and making new friends). Such changes can affect one’s self-beliefs and psychological well-being and have a lasting impact. Therefore, this is a key time to intervene as there are greater connections on which to build (ie, colleagues and friends, hobbies and interests, a sense of purpose through employment or study, and family relationships). Indeed, intervening early during the experience of psychosis has been shown to improve clinical outcomes [17,18].

Effective face-to-face interventions, drawing on CBT and positive psychology techniques, can improve self-confidence

[19-22]. Some of these techniques involve helping patients learn to increase positive self-beliefs—a mechanism for improving psychological well-being—through direct experience (ie, being in situations in which these positive self-beliefs can be activated and maintained). It is these techniques that lend themselves particularly well to implementation in VR therapies.

Design Process

Overview

An automated VR program with embedded therapeutic principles requires a design process that ensures these principles are correctly targeted and implemented. This means input from end users is imperative. The overall aim is to create a program that is easy to use, meaningful, accessible, and highly satisfactory. The design process for Phoenix specifically used a double diamond approach to achieve this aim. Similar to gameChange [1], we used a person-centered design process within the double diamond approach. This is an iterative approach that ensures that the end user's perspective and needs are incorporated into all stages (in this case, the end users are patients with psychosis).

The double-diamond design approach, developed by the Design Council [23], involves exploring an issue widely then taking focused action. It is organized into 4 phases: discover (understanding lived experience of the clinical problem and enhancing understanding and assessment of the key target), define (set out the design brief), develop (refine storyboards, program the application, and perform user testing), and deliver (finalize implementation, perform debugging, and conduct final usability testing). The design process for Phoenix will therefore be described within each of these phases. The double diamond approach was used for the overall look, feel, and design of Phoenix, as well as for individual elements within the program itself (eg, the virtual coach). Prior to beginning the design process, a young person's advisory group (YPAG) was formed, facilitated by the McPin Foundation (a foundation that aims to include people with lived experience of mental health difficulties in mental health research). The YPAG included 14 members aged 16 - 30 years with lived experience of the clinical problem (ie, low levels of positive self-beliefs in the context of the experience of psychosis).

The research team at Oxford included clinical psychologists, a treatment designer, and the software development team (3 VR programmers, a 3D artist, and a lead computer scientist). As with gameChange, the overall responsibility for final design decisions was held by the chief investigator at Oxford.

Discover

Before the first meetings with the YPAG, clinical psychologists and the treatment designer met weekly over the course of 2 months (January and February 2021) to develop a basic design brief that would form the basis for initial discussions with the YPAG. This design brief set the boundaries for the treatment and how techniques could be implemented in VR. The initial aim of the discover phase was therefore to further elaborate this design brief. Using clinical expertise and findings from previous face-to-face studies [19,24], we set the key mechanisms to be targeted as well as the psychological techniques used to target these mechanisms. The resulting design brief outlined the clinical problem (low self confidence and psychological well-being), the target audience (young people with a diagnosis of psychosis), the clinical target (positive self-beliefs), the key function of VR (provide experiences of positive beliefs in VR to stimulate engagement in real-world activities), essential elements (eg, a virtual coach who would guide users through the program), and how it might fit in existing services (a discrete form of standalone therapy that patients could take home and use with the support of a mental health provider and supplementary materials). The script for the program needed to effectively communicate psychological principles using accessible language and ensure that the content was appropriate for the users. Additionally, the program needed to be certified as a medical device and therefore developed with regulatory and safety procedures in mind.

The first meetings with the YPAG primarily centered around clarifying the terminology used to describe self-confidence, as well as specific activities and situations in everyday life that increase self-confidence and psychological well-being. At least one member of the software development team, as well as all members of the research team, attended these meetings. Given the clinical target (positive self-beliefs), an important aspect of these initial meetings with the YPAG therefore included the identification of specific positive self-beliefs associated with increased self-confidence and psychological well-being. It became clear that existing measures did not provide a wide enough range of positive self-beliefs, did not use clear and precise wording, and would not select those in greatest need of interventions to improve such beliefs. It was agreed that a well-validated measure of positive self-beliefs—informed by people with lived experience—should include key beliefs highly connected to psychological well-being that can be targeted by psychological approaches.

Therefore, as part of this discovery phase, we developed such a measure: the Oxford Positive Self Scale (OxPos) [25]. Through discussions with the YPAG, it was agreed this measure should assess the key types of positive self-beliefs to be targeted in VR. Three key types were identified: those relating to mastery and achievement ("I have a purpose," "I am capable," and "I can achieve things"); those relating to strength and resilience in the face of a challenge ("I am strong," "I am resilient," and "I don't give up"), and those relating to enjoyment ("I can relax," "I can have fun," and "I can switch off"). These beliefs fed into all aspects of programming. The validation study for the OxPos began during this phase and resulted in a 24-item, easy to use, and psychometrically robust assessment of positive self-beliefs linked to psychological well-being.

The YPAG also discussed their views on the virtual coach who would guide users through Phoenix. Different ideas about the coach were shared, including the possibility of a coach that was nonhuman (ie, a robot or creature). Ultimately, it was agreed that a human coach that was friendly, kind, and was ethnically ambiguous and relatively gender neutral in appearance would be maximally relatable and therefore most effective for guiding patients through Phoenix. Emotional attributes such as the warmth of facial expressions and animations such as head nodding were highlighted as particularly important for therapeutic alliance [26]. Furthermore, following discussions on the importance of terminology used to describe self-confidence and positive self-beliefs, initial ideas on the script for the coach were developed during meetings among the research team. During this phase, how the coach would describe the rationale behind Phoenix was discussed: that often, feelings of vulnerability and difficult experiences can allow negative self-beliefs to flourish and make it difficult to access positive self-beliefs. Therefore, Phoenix is about redressing this balance.

Given that any suggestions for activities and environments in VR needed to be programmable, weekly meetings among the clinical psychologists, treatment designer, and software development team at the University of Oxford were held in parallel to meetings with the YPAG. Additionally, at least one member of the software development team attended meetings with the YPAG to provide input on what was programmable. Due to the COVID-19 pandemic, meetings were conducted remotely over Zoom (Zoom Video Communications Inc).

Define

The design brief was used as a living document that, in the define phase, was elaborated in greater detail to set the framework for the software development team. This included identifying the broad virtual environments linked to the 3 key types of positive self-beliefs and the activities within these environments. For mastery and achievement, YPAG members particularly identified activities that would tap into a sense of responsibility, such as caring for animals. They also highlighted the importance of simply completing tasks (ie, via a to-do list). For strength and resilience, it was agreed that patients should enter a challenging situation that would allow them to learn that they could complete a task despite feeling anxious and therefore feel a sense of accomplishment. Finally, for enjoyment, members described wanting a situation that was relaxing, where patients could immerse themselves in their surroundings and enjoy the experience in the moment. It was agreed that this situation should include an outdoor space.

Three environments were chosen: a community farm for mastery and achievement, a TV studio for strength and succeeding in the face of a challenge, and a forest by a lake for pleasure and enjoyment. Initial sketches were drawn by the 3D artist and provided to YPAG members in both 2D and 3D format for feedback. When discussing activities to include in each environment, the importance of multiple types of positive feedback was emphasized (eg, animals' facial expressions or behavior, comments from the coach, and positive sounds when tasks were completed). Therefore, these were incorporated in all activities. In addition to the 3 environments where users would complete activities, a fourth area was highlighted as important: the welcome room where users were initially greeted by the virtual coach. It was decided that this room should be a calming, indoor space with elements of nature (ie, plants, a room overlooking mountains). Similar to the other environments, sketches were drawn by the 3D artist and provided to the YPAG members for feedback. We also discussed having uplifting music playing in the background; the music chosen was also provided to the YPAG for feedback.

In this phase, the features and appearance of the coach were also finalized with feedback provided by the YPAG on the look and feel (ie, hair color, body shape, and clothing) as well as the voice (eg, male vs female and young vs old). The YPAG indicated that they would prefer a female coach. They wanted a body shape that was tending more to gender neutral (eg, not too feminine or masculine) and average in terms of size and height (eg, not too thin). Her name (Farah) and voice were therefore also finalized. Several voice actors were identified from the Royal Academy of Dramatic Art in London, United Kingdom, and clips were sent to the YPAG for feedback. A voice actor was chosen and recording sessions for the script took place in later phases of the design process. This occurred concurrently with the development of the script.

The 3D artist also created an overall visual design manual detailing the general look and feel of each element of the program (including textures, lighting, color, animation, and stylization) to ensure consistency across the environments. The name of the therapy (Phoenix) and logo were also designed during this phase, and feedback was provided by the YPAG. Phoenix was suggested by the research team given that, in Greek mythology, a phoenix is associated with the sun and a sense of renewal. This fitted with the aim of the therapy—to encourage patients to build up positive self-beliefs and rediscover themselves. The logo was drawn by a consultant graphic designer and included suggestions of color schemes, images of the phoenix bird, and different fonts for the name.

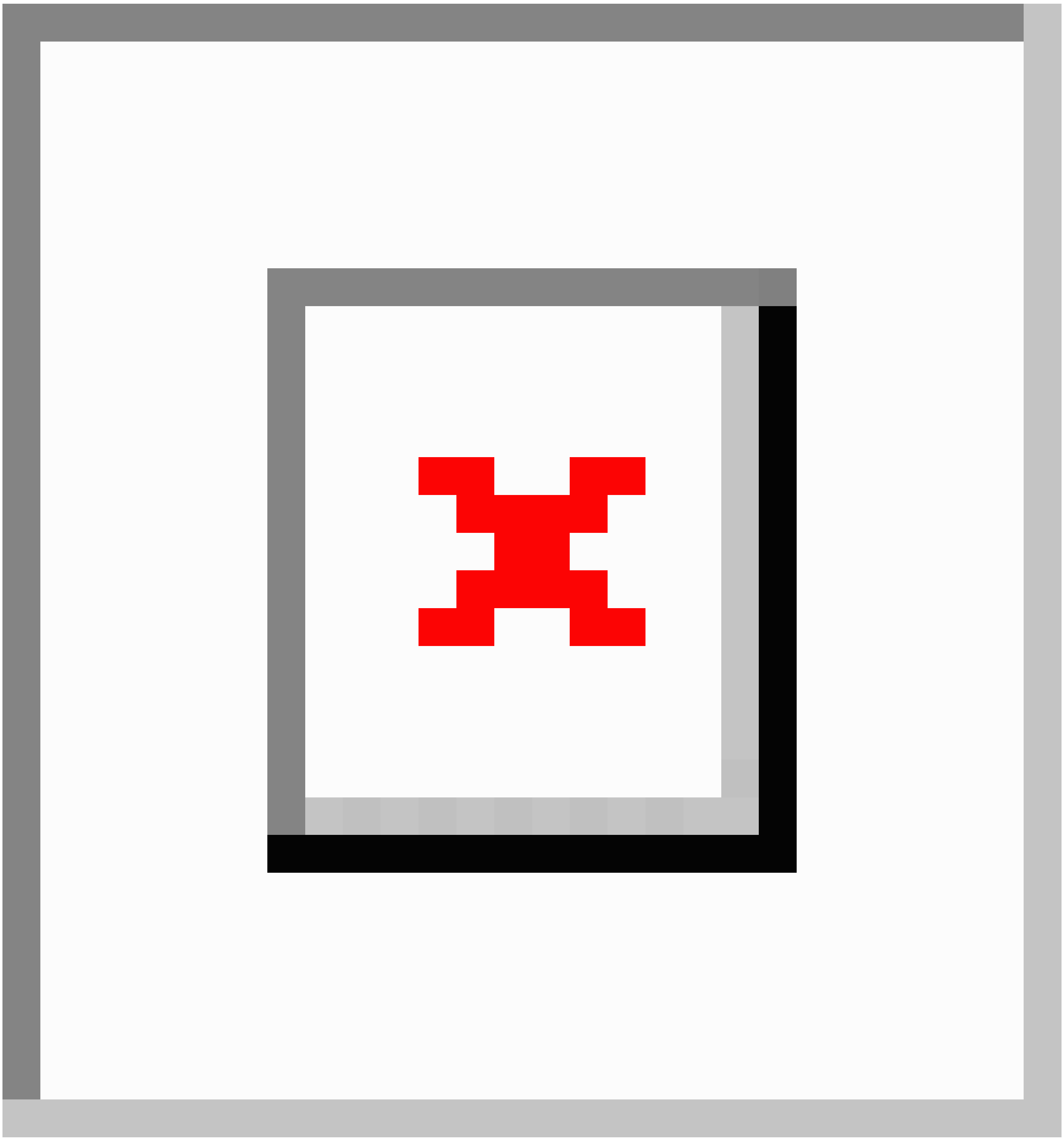
Altogether, across both the discover and define phases, 7 virtual meetings were held with the YPAG over the course of 3 months (February-May 2021), totaling 76.5 YPAG person-hours. Email consultations on the final version of the OxPos questionnaire, logo, and project and coach names totaled an additional 19.75 YPAG person-hours.

Develop

Overview

The develop phase (June 2021-August 2022) included finalizing and programming the specific activities within each of the 3 main environments, as well as the welcome room (Figure 1 provides screenshots). The welcome room includes large glass-fronted windows looking out onto waterfalls, a lake, and mountains, with the room itself containing plants and a fireplace. Users can hear the sounds of birds and water splashing and uplifting music. The coach greets users in front of the window while describing the purpose of Phoenix and providing a tutorial on how to use the equipment to navigate through the program.

Figure 1. Pictures of the areas within the Phoenix VR Self-Confidence Therapy program.



In the mastery and achievement scenario (the community farm), 3 key activities were chosen, comprising 10 steps: caring for animals (feeding one animal, building a hutch, and then feeding multiple animals), growing plants (preparing the bed, planting the seeds, watering the seeds, and picking crops), and redecorating a house (hanging and populating items on shelves, hanging pictures on the wall, and building a bird feeder). Rabbits were chosen as the animals due to patient preference and design complexity (ie, rabbits are relatively easy to design, animate, and program in VR compared to other animals).

In the scenario designed to promote success in the face of a challenge (the TV studio), the key element was being the center of attention and having to perform a specific task, where an audience would gradually increase in size and provide positive feedback. The resulting activity requires users to read out preexisting text from a teleprompter, with the scene being gradually populated by an increasing number of characters (eg, crew and audience). The audience (when present) claps at the end of each reading. Other elements to increase difficulty include the addition of cameras (being both off and on), spotlights, and "on air" signs. The resulting scenario includes 10 levels of difficulty, with the final 2 levels requiring users to make up a weather report rather than read a preexisting one from the teleprompter.

For the pleasurable scenario (the forest), the main programming focus was to ensure the scene was rich in terms of audio (eg, birds singing, owls hooting, water splashing) and visual detail in order to ensure full immersion. Four activities were ultimately chosen: 2 exercises to promote relaxation (a savoring exercise, whereby users are encouraged to focus on different elements of the environment, and a progressive muscle relaxation exercise, both delivered by the coach) and 2 games purely for fun (a game whereby users throw pinecones into holes in trees for points and a drumming game with different modes). The YPAG also suggested giving users choice about whether they complete activities in daylight or at night—therefore, this is built into each exercise and each mode has different sounds and sights.

The script for the dialogue of the virtual coach evolved as activities were programmed. Preferences for wording were elicited through feedback from the YPAG at all stages of the design process. As the scenarios in Phoenix do not include social interaction with other virtual characters, the script only includes lines delivered by the virtual coach. Initial voice recording for the script took place at a recording studio in Oxford over 2 sessions in June and August 2022.

Software Development

The development of the Phoenix software was carried out by the software development team (3 VR programmers, a 3D artist, and a lead computer scientist) within the research team at the University of Oxford. A key requirement was for Phoenix to run on a standalone headset to allow for greater accessibility and portability (ie, ensuring that participants could take the headsets home and run Phoenix themselves, without the need for additional hardware such as a laptop and link cable). Therefore, Phoenix was developed to run on a Meta Quest 2 headset.

The 3D model for the virtual coach was created by the 3D artist using Character Creator (Reallusion)and ZBrush (Maxon). This model was then animated using iClone (Reallusion) software, including blinking and lip-sync. The coach was programmed to look at the user when speaking to them (or at other objects in the scene at relevant moments). Virtual environments and interactive objects were modeled using 3ds Max and Maya (both Autodesk) and textured using Substance Painter (Adobe).

The 3D models were imported into a project in the Unity game engine (Unity) where they were assembled into individual scenes. The programmers added behaviors to create the various activities and other scripts to control positioning and navigation. The player rig and object interactions were based on Oculus Unity Integration (Meta Platforms), with enhancements provided by HurricaneVR (CloudWalkin Games).

A database was created in the cloud using Azure (Microsoft) to facilitate identification of users, store their progress, and receive a log of each session generated in real time by the application. The session logs contained a chronological list of details of significant events occurring during the session to facilitate future analyses. This system allowed for offline use, with data generated while the headset was not connected to the internet being stored on the headset so that it could be synced with the server later when an internet connection was available.

For efficiency of development, an iterative approach was used whereby initial prototypes of environments and activities were checked by the research team at the University of Oxford for usability and effectiveness. Thus, the initial design and script were continuously honed as each build of the application was distributed and reviewed. Some decisions regarding scoping and descoping of design elements became clearer only once a prototype containing the whole experience in some form was available. Once elements and activities were more fully programmed, we ran usability testing sessions with members of the YPAG.

Usability Testing

Usability sessions were completed in person with members of the YPAG as well as 2 members of another project's patient involvement group to ensure a diversity of users. Testing sessions took place at the Warneford Hospital site at the University of Oxford in Oxford, United Kingdom, and the McPin Foundation offices in London, United Kingdom. At least one member of the research team (clinical psychologist or treatment designer) and one member of the software development team were present for all sessions. Sessions lasted approximately 2 hours, with the first 1 - 1.5 hours spent using Phoenix (with breaks) and approximately 30 minutes for debriefing. A usability questionnaire, created by our team, was given to all users during the final user testing session. Users were offered the choice to complete the questionnaire on their own or with help from the research team. Additional feedback was elicited during the debrief using specific questions about what they had done in VR (ie, how they found interacting with the objects, what they thought of the activities, and whether the coach was helpful).

The first usability session was held with 4 users in May 2022 to provide feedback on the farm activities, welcome room, TV environment, and forest environment (particularly the pine cone game). The second usability session was held with one user in June 2022 and the third and final usability session was held in August 2022 with 7 users (3 of whom had attended the earlier user testing sessions and 4 of whom had not) to review all activities and provide feedback on Farah, the virtual coach. At this point (August 2022), only 50% of the coach's lines had been animated. Table 1 and Table 2 show the results from the usability questionnaire given during the final user testing session. These results indicated that all users found it easy to understand the coach's instructions, carry out actions, and move through the program. All users except one found it easy to remember how to do things a second time. Four users found it easy to know what to do in any given situation while 3 found it difficult. This was likely because the coach's lines had not been fully programmed and a textbox with instructions was used in its place. All users felt immersed in VR, were satisfied with the experience, and enjoyed using the therapy. Two users felt mildly nauseous while using VR while 5 did not.

Table . First part of usability questionnaire completed by members of the young person's advisory group (n=7) during Phoenix VR Self-Confidence Therapy testing in August 2022.

Questions	Very easy, n	Easy, n	Difficult, n	Very difficult, n
Knowing what to do in any given VR[a] situation was…	0	4	3	0
Understanding the coach's instructions was…	2	5	0	0
Carrying out an action (for example, planting the seeds, throwing the pinecones) was…	2	5	0	0
Moving through the program was…	2	5	0	0
Learning what to do and how to do it was…	1	4	1	1
Remember how to do things a second time was…	4	2	1	0

[a]VR: virtual reality.

Table . Second part of usability questionnaire completed by members of the young person's advisory group (n=7) during Phoenix VR Self-Confidence Therapy testing in August 2022.

Statements	Agree, n	Agree a little, n	Disagree a little, n	Disagree, n
When using the VR[a], I felt like I was in the situation.	4	3	0	0
I was satisfied with the overall VR experience.	7	0	0	0
I enjoyed using the treatment.	5	2	0	0
The VR made me feel sick.	0	2	0	5

[a]VR: virtual reality.

Deliver

This phase of the design process (August 2022-May 2023) included finalizing all activities, debugging, and conducting final usability testing (as part of a cohort study with patients in clinical services). As in the develop phase, an iterative approach was used whereby finalized activities were checked by the research team at the University of Oxford for usability and effectiveness. The script was also finalized during this phase after feedback from users in the cohort study and a last recording with the voice actor for the coach took place in February 2023. The final script is approximately 500 lines. When the design and script were locked in, there was a pure bug-testing phase involving the whole team (clinical psychologists and software developers), during which the application was polished to remove glitches and add final enhancements.

The cohort study for Phoenix examined usability, satisfaction, and initial effectiveness of Phoenix in 12 young people with a diagnosis of psychosis across 3 different mental health trusts in the south of England. Patients were recruited over a 4-month period (December 2022-March 2023). All patients were being seen in National Health Service (NHS) services. The primary outcome was usability (Table 3 and Table 4) and satisfaction with the Phoenix program, with secondary outcomes of positive self-beliefs (measured by the OxPos) and psychological well-being (a full description of this cohort study was provided by Freeman et al [27]). The same usability questionnaire administered during user testing in the develop phase was also used with patients in the case series. Compared with the first usability testing session conducted in the develop phase, all users found it easy to move through the program, understand the coach's instructions, carry out an action, learn what to do, and remember how to do things a second time. Only one person found it difficult to know what to do in any given situation. This improvement in user experience is likely due to the coach's lines being fully programmed. Similarly, all users enjoyed the treatment, all but one user were satisfied with the overall VR experience, and all but one user felt immersed in VR when they were using it. Only one user reported that the VR made them feel sick.

Table . First part of usability questionnaire completed by patients (n=11) in the case series for Phoenix VR Self-Confidence Therapy.

	Very easy, n	Easy,n	Difficult, n	Very difficult, n
Knowing what to do in any given VR[a] situation was…	7	3	1	0
Understanding the coach's instructions was…	10	1	0	0
Carrying out an action (for example, planting the seeds, throwing the pinecones) was…	7	4	0	0
Moving through the program was…	7	4	0	0
Learning what to do and how to do it was…	7	4	0	0
Remember how to do things a second time was…	8	3	0	0

[a]VR: virtual reality.

Table . Second part of usability questionnaire completed by patients (n=11) in the case series for Phoenix VR Self-Confidence Therapy.

	Agree, n	Agree a little, n	Disagree a little, n	Disagree, n
7. When using the VR[a], I felt like I was in the situation.	6	4	1	0
8. I was satisfied with the overall VR experience.	9	1	1	0
9. I enjoyed using the treatment.	9	2	0	0
10. The VR made me feel sick.	0	1	1	9

[a]VR: virtual reality.

Regulatory Process

As described in the original design brief, Phoenix is certified as a class I medical device in conformity with the essential requirements of Directive 93/42/EEC. The certification process was run in parallel with the design process for Phoenix and involved evaluating risks, software, usability, clinical safety, and efficacy of the application. A medical device consultant provided feedback on how to document the software development to ensure it complied with requirements. This was done via online project management and bug tracking software (Jira/Confluence; Atlassian Corporation). The research team, along with the consultant, developed and completed documentation regarding potential hazards of use and embedded mitigating factors, instructions for use, a data processing impact assessment, software specifications, and a clinical evaluation report (which included data on usability testing, described in the Design Process section).

Phoenix VR Self-Confidence Therapy

The resulting Phoenix VR Self-Confidence Therapy program is designed to be delivered over the course of approximately 6 weekly sessions. The activities within Phoenix spark positive self-beliefs that are consolidated through real-world activities. Although key positive self-beliefs are embedded within each broad area and explicitly named by the coach, patients are also encouraged to identify additional positive self-beliefs on which to focus. Key techniques, aided by the mental health provider, include helping patients identify their strengths and the activities linked to those strengths, develop the ability to fully immerse themselves in and therefore savor positive activities, and increase engagement in meaningful activities.

Users start out in the welcome room and, after an introduction to Phoenix and tutorial on how to use the program, choose from the 3 environments. They can repeat activities and return to the welcome room at any time. Prior to each activity, they are provided with a description of the purpose and what steps need to be taken to complete it. After completion, Farah provides positive feedback and encouragement. Farah directs users to pay attention to the feelings of enjoyment and achievement elicited by the activities in Phoenix and encourages them to elicit these same feelings in their daily life. Handouts with positive psychology techniques supplement the delivery of Phoenix.

Users also have the option to take the headset home with them to use in between sessions. This means that in-person sessions may or may not involve the use of VR depending on patient preference. The mental health provider helps users set practices to complete between sessions aimed at continuing to build positive self-beliefs, with sessions focused on reviewing these practices and consolidating the learning made from using Phoenix.

Clinical Evaluation

Ethical Considerations

Ethical approval for the cohort study and the randomized controlled trial (described below) was obtained from an NHS research ethics committee (22/LO/0273). The study abided by the Ethical Principles of Psychologists and Code of Conduct. The McPin Foundation followed all appropriate safeguarding policies and ethical considerations during the design process.

Next Steps

Twelve patients completed Phoenix as part of the cohort study and 11 provided outcome data [27]. Uptake of Phoenix was high: 9 users had 6 sessions and 8 patients consistently used Phoenix on their own in between sessions. Satisfaction was high and there were few side effects. All users rated the quality of the VR therapy as good or excellent. An initial evaluation of positive self-beliefs and psychological well-being showed large improvements. Therefore, the cohort study is now being followed up with a randomized controlled trial in 4 mental health trusts in the south of England.

Eighty patients will take part [28]. Half of the participants will receive Phoenix in addition to their usual treatment and half will continue with their usual treatment only. There will be 2 follow-up assessments: one at 6 weeks (after treatment completion) and one at 12 weeks. Assessments will be completed by research assistants who are blind to the treatment allocation. The primary outcome is positive self-beliefs (measured by the OxPos), with secondary outcomes of well-being, everyday confidence, depression, anxiety, hopelessness, activity levels, and quality of life. Service use will also be measured for an analysis of health economics. For those who receive Phoenix, satisfaction and side effects will also be assessed.

Acknowledgments

Phoenix VR Self-Confidence Therapy is funded by an award to DF from the International Foundation, an independent and not-for-profit foundation. The project is also supported by the National Institute for Health and Care Research (NIHR) Oxford Health Biomedical Research Centre. DF is an NIHR senior investigator. FW is funded by a Wellcome Trust Clinical Doctoral Fellowship (102176/B/13/Z).

The project is sponsored by the University of Oxford. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, the Department of Health, the University of Oxford, or the International Foundation. We wish to thank all the members of the young person’s advisory group.

Conflicts of Interest

DF is the scientific founder of Oxford VR, a University of Oxford spin-out company. DF and JF hold equity in Oxford VR. Oxford VR was not involved in the development of Phoenix VR Self-Confidence Therapy. L Rosebrock, RW, MB, AR, ALM, L Riffiod, RK, TK, and FW have no conflicts of interest.

References

1. Lambe S, Knight I, Kabir T, et al. Developing an automated VR cognitive treatment for psychosis: gameChange VR therapy. J Behav Cogn Ther 2020 Apr;30(1):33-40. [doi: 10.1016/j.jbct.2019.12.001]
2. Knight I, West J, Matthews E, et al. Participatory design to create a VR therapy for psychosis. Design Health 2021 May;5(1):98-119. [doi: 10.1080/24735132.2021.1885889]
3. Freeman D, Reeve S, Robinson A, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400. [doi: 10.1017/S003329171700040X] [Medline: 28325167]
4. Martens MA, Antley A, Freeman D, Slater M, Harrison PJ, Tunbridge EM. It feels real: physiological responses to a stressful virtual reality environment and its impact on working memory. J Psychopharmacol 2019 Oct;33(10):1264-1273. [doi: 10.1177/0269881119860156] [Medline: 31294651]
5. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007 Oct;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
6. Bond J, Kenny A, Pinfold V, et al. A safe place to learn: peer research qualitative investigation of gameChange virtual reality therapy. JMIR Serious Games 2023 Jan 16;11:e38065. [doi: 10.2196/38065] [Medline: 36645707]
7. Pot-Kolder R, Geraets CNW, Veling W, et al. Virtual-reality-based cognitive behavioural therapy versus waiting list control for paranoid ideation and social avoidance in patients with psychotic disorders: a single-blind randomised controlled trial. Lancet Psychiatry 2018 Mar;5(3):217-226. [doi: 10.1016/S2215-0366(18)30053-1] [Medline: 29429948]

8. Cella M, Tomlin P, Robotham D, et al. Virtual reality therapy for the negative symptoms of schizophrenia (V-NeST): a pilot randomised feasibility trial. Schizophr Res 2022 Oct;248:50-57. [doi: 10.1016/j.schres.2022.07.013] [Medline: 35939920]
9. Freeman D, Haselton P, Freeman J, et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. Lancet Psychiatry 2018 Aug;5(8):625-632. [doi: 10.1016/S2215-0366(18)30226-8] [Medline: 30007519]
10. Freeman D, Lambe S, Kabir T, et al. Automated virtual reality therapy to treat agoraphobic avoidance and distress in patients with psychosis (gameChange): a multicentre, parallel-group, single-blind, randomised, controlled trial in England with mediation and moderation analyses. Lancet Psychiatry 2022 May;9(5):375-388. [doi: 10.1016/S2215-0366(22)00060-8] [Medline: 35395204]
11. Freeman D, Lambe S, Galal U, et al. Agoraphobic avoidance in patients with psychosis: severity and response to automated VR therapy in a secondary analysis of a randomised controlled clinical trial. Schizophr Res 2022 Dec;250:50-59. [doi: 10.1016/j.schres.2022.10.008] [Medline: 36343472]
12. Freeman D, Lister R, Waite F, et al. Automated virtual reality cognitive therapy versus virtual reality mental relaxation therapy for the treatment of persistent persecutory delusions in patients with psychosis (THRIVE): a parallel-group, single-blind, randomised controlled trial in England with mediation analyses. Lancet Psychiatry 2023 Nov;10(11):836-847. [doi: 10.1016/S2215-0366(23)00257-2] [Medline: 37742702]
13. Collett N, Pugh K, Waite F, Freeman D. Negative cognitions about the self in patients with persecutory delusions: an empirical study of self-compassion, self-stigma, schematic beliefs, self-esteem, fear of madness, and suicidal Ideation. Psychiatry Res 2016 May 30;239:79-84. [doi: 10.1016/j.psychres.2016.02.043] [Medline: 27137965]
14. Freeman D, Taylor KM, Molodynski A, Waite F. Treatable clinical intervention targets for patients with schizophrenia. Schizophr Res 2019 Sep;211:44-50. [doi: 10.1016/j.schres.2019.07.016] [Medline: 31326234]
15. Häfner H, Maurer K, Löffler W, et al. The epidemiology of early schizophrenia. Br J Psychiatry 1994 Apr;164(S23):29-38. [doi: 10.1192/S0007125000292714]
16. Gogtay N, Vyas NS, Testa R, Wood SJ, Pantelis C. Age of onset of schizophrenia: perspectives from structural neuroimaging studies. Schizophr Bull 2011 May;37(3):504-513. [doi: 10.1093/schbul/sbr030] [Medline: 21505117]
17. Marshall M, Rathbone J. Early intervention for psychosis. Schizophr Bull 2011 Nov;37(6):1111-1114. [doi: 10.1093/schbul/sbr110] [Medline: 21908851]
18. McGorry PD, Killackey E, Yung A. Early intervention in psychosis: concepts, evidence and future directions. World Psychiatry 2008 Oct;7(3):148-156. [doi: 10.1002/j.2051-5545.2008.tb00182.x] [Medline: 18836582]
19. Freeman D, Pugh K, Dunn G, et al. An early phase II randomised controlled trial testing the effect on persecutory delusions of using CBT to reduce negative cognitions about the self: the potential benefits of enhancing self confidence. Schizophr Res 2014 Dec;160(1-3):186-192. [doi: 10.1016/j.schres.2014.10.038] [Medline: 25468186]
20. Freeman D, Freeman J. You Can Be Happy: The Scientifically Proven Way to Change: Pearson; 2012.
21. Forkert A, Brown P, Freeman D, Waite F. A compassionate imagery intervention for patients with persecutory delusions. Behav Cogn Psychother 2022 Jan;50(1):15-27. [doi: 10.1017/S1352465821000229] [Medline: 34078499]
22. Seligman MEP, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
23. Make life better by design. Design Council. URL: https://www.designcouncil.org.uk/ [accessed 2023-07-31]
24. Freeman D, Emsley R, Diamond R, et al. Comparison of a theoretically driven cognitive therapy (the Feeling Safe Programme) with befriending for the treatment of persistent persecutory delusions: a parallel, single-blind, randomised controlled trial. Lancet Psychiatry 2021 Aug;8(8):696-707. [doi: 10.1016/S2215-0366(21)00158-9] [Medline: 34246324]
25. Freeman D, Rosebrock L, Loe BS, Saidel S, Freeman J, Waite F. The Oxford Positive Self Scale: psychometric development of an assessment of cognitions associated with psychological well-being. Psychol Med 2023 Nov;53(15):7161-7169. [doi: 10.1017/S0033291723000624] [Medline: 36927521]
26. Wei S, Freeman D, Rovira A. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. Sci Rep 2023 Jul 17;13(1):11517. [doi: 10.1038/s41598-023-38499-7] [Medline: 37460586]
27. Freeman D, Freeman J, Ahmed M, et al. Automated VR therapy for improving positive self-beliefs and psychological well-being in young patients with psychosis: a proof of concept evaluation of Phoenix VR self-confidence therapy. Behav Cogn Psychother 2024 May;52(3):277-287. [doi: 10.1017/S1352465823000553] [Medline: 37942541]
28. Freeman D, Freeman J, Rovira A, et al. Randomised controlled trial of automated VR therapy to improve positive self-beliefs and psychological well-being in young people diagnosed with psychosis: a study protocol for the Phoenix VR self-confidence therapy trial. BMJ Open 2023 Dec 7;13(12):e076559. [doi: 10.1136/bmjopen-2023-076559] [Medline: 38149422]

Abbreviations

NHS: National Health Service
OxPos: Oxford Positive Self Scale
VR: virtual reality

VR-CBT: virtual reality–based cognitive behavioral therapy
YPAG: young person’s advisory group

Edited by G Eysenbach, T Leung; submitted 18.08.23; peer-reviewed by AG Bergin, E Kawamoto; revised version received 21.01.24; accepted 31.05.24; published 07.08.24.

Please cite as:
Rosebrock L, Freeman J, Rovira A, Miguel AL, Ward R, Bousfield M, Riffiod L, Kamvar R, Kabir T, Waite F, Freeman D
Developing an Automated Virtual Reality Therapy for Improving Positive Self-Beliefs and Psychological Well-Being (Phoenix VR Self-Confidence Therapy): Tutorial
JMIR Serious Games 2024;12:e51512
URL: https://games.jmir.org/2024/1/e51512
doi:10.2196/51512

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org