JMIR Serious Games

Journal Impact Factor (JIF) (2023): 3.8
Volume 10 (2022), Issue 2 ISSN 2291-9279 Editor-in-Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

The Effectiveness of a Serious Game (MemoreBox) for Cognitive Functioning Among Seniors in Care Facilities: Field Study (e33169)
Jana Kleschnitzki, Luzi Beyer, Reinhard Beyer, Inga Großmann. 4

A Mobile-based Virtual Reality Speech Rehabilitation App for Patients With Aphasia After Stroke: Development and Pilot Usability Study (e30196)
Xiaofan Bu, Peter Ng, Ying Tong, Peter Chen, Rongrong Fan, Qingping Tang, Qinqin Cheng, Shuangshuang Li, Andy Cheng, Xiangyu Liu. . . 17

Augmented Reality-Based Surgery on the Human Cadaver Using a New Generation of Optical Head-Mounted Displays: Development and Feasibility Study (e34781)
Behrus Puladi, Mark Ooms, Martin Bellgardt, Mark Cesov, Myriam Lipprandt, Stefan Raith, Florian Peters, Stephan Möhlhenrich, Andreas Prescher, Frank Hölzle, Torsten Kuhlen, Ali Modabber. 31

HelperFriend, a Serious Game for Promoting Healthy Lifestyle Behaviors in Children: Design and Pilot Study (e33412)
Ismael Espinosa-Curiel, Edgar Pozas-Bogarin, Maryleidi Hernández-Arvizu, Maria Navarro-Jiménez, Edwin Delgado-Pérez, Juan Martínez-Miranda, Humberto Pérez-Espinosa. 44

Virtual Reality Training Using Nintendo Wii Games for Patients With Stroke: Randomized Controlled Trial (e29830)
Naveed Anwar, Hossein Karimi, Ashfaq Ahmad, Syed Gilani, Kehkshan Khalid, Ahmed Aslam, Asif Hanif. 57

The Role of Agency and Threat Immediacy in Interactive Digital Narrative Fear Appeals for the Prevention of Excessive Alcohol Use: Randomized Controlled Trial (e32218)
Hendrik Engelbrecht, Laura van der Laan, Renske van Enschot, Emiel Krahmer. 65

Identifying Player Types to Tailor Game-Based Learning Design to Learners: Cross-sectional Survey using Q Methodology (e30464)
A Van Gaalen, J Schönrock-Adema, R Renken, A Jaarsma, J Georgiadis. 79

Using Virtual Reality to Improve Classroom Behavior in People With Down Syndrome: Within-Subjects Experimental Design (e34373)
Stefan Michalski, Ancret Szpak, Caroline Ellison, Rowena Cornish, Tobias Loetscher. 95

Adapting the Use of Digital Content to Improve the Learning of Numeracy Among Children With Autism Spectrum Disorder in Rwanda: Thematic Content Analysis Study (e28276)
Theoneste Ntalindwa, Mathias Nduwingoma, Alphonse Uworwabayeho, Pascasie Nyirahabimana, Evariste Karangwa, Tanjir Rashid Soron, Thomas Westin, Thashmee Karunaratne, Henrik Hansson. 104

Gamification in Diplomacy Studies as an Effective Tool for Knowledge Transfer: Questionnaire Study (e32996)
Mihai Cercel. 114

Effectiveness and Utility of Virtual Reality Infection Control Simulation for Children With COVID-19: Quasi-Experimental Study (e36707)
Mi Yu, Mi Yang. 126

Instructor Development Workshops for Advanced Life Support Training Courses Held in a Fully Virtual Space: Observational Study (e38952)
Tetsuro Kiyozumi, Norio Ishigami, Daisuke Tatsushima, Yoshiyuki Araki, Yuya Yoshimura, Daizoh Saitoh. 139

A Photography-based, Social Media Walking Intervention Targeting Autonomous Motivations for Physical Activity: Semistructured Interviews With Older Women (e35511)
Michael Robertson, Maria Swartz, Ursela Christopherson, Jason Bentley, Karen Basen-Engquist, Debbe Thompson, Elena Volpi, Elizabeth Lyons. 149

Effect of an Active Video Game Intervention Combined With Multicomponent Exercise for Cardiorespiratory Fitness in Children With Overweight and Obesity: Randomized Controlled Trial (e33782)
Cristina Comeras-Chueca, Lorena Villalba-Heredia, Jose Perez-Lasierra, Gabriel Lozano-Berges, Angel Matute-Llorente, German Vicente-Rodriguez, Jose Casajus, Alex Gonzalez-Aguero. 161

User Experience and Usability of Neumorphism and Gamification User Interface Designs in an HIV Self-Test Referral Program for Men Who Have Sex With Men: Prospective Open-Label Parallel-Group Randomized Controlled Trial (e35869)
Tsz Kwan, Denise Chan, Shui Lee. 175

Impact of Visual Game-Like Features on Cognitive Performance in a Virtual Reality Working Memory Task: Within-Subjects Experiment (e35295)
Eric Redlinger, Bernhard Glas, Yang Rong. 187

Interactive Digital Game for Improving Visual–Perceptual Defects in Children With a Developmental Disability: Randomized Controlled Trial (e34756)
Wen-Lan Wu, Yu-Ling Huang, Jing-Min Liang, Chia-Hsin Chen, Chih-Chung Wang, Wen-Hsien Ho. 201

Effects of Virtual Reality–Based Multimodal Audio-Tactile Cueing in Patients With Spatial Attention Deficits: Pilot Usability Study (e34884)
Samuel Knobel, Brigitte Kaufmann, Nora Geiser, Stephan Gerber, René Müri, Tobias Nef, Thomas Nyffeler, Dario Cazzoli. 210

Perceptions of Cognitive Training Games and Assessment Technologies for Dementia: Acceptability Study With Patient and Public Involvement Workshops (e32489)
Kyle Harrington, Michael Craven, Max Wilson, Aleksandra Landowska. 224

Feedback on Trunk Movements From an Electronic Game to Improve Postural Balance in People With Nonspecific Low Back Pain: Pilot Randomized Controlled Trial (e31685)
Anita Meinke, Rick Peters, Ruud Knols, Jaap Swanenburg, Walter Karlen. 238

Designing a Serious Game (Above Water) for Stigma Reduction Surrounding Mental Health: Semistructured Interview Study With Expert Participants (e21376)
Rina Wehbe, Colin Whaley, Yasaman Eskandari, Ally Suarez, Lennart Nacke, Jessica Hammer, Edward Lank. 260

Associations Between Addictive Behaviors, Individual Characteristics, and the Use of Gambling Services Within the World of Gaming: Cross-sectional Survey Study (e29077)
Mark Kisch, Anders Håkansson. 281

Video Game Addictive Symptom Level, Use Intensity, and Hedonic Experience: Cross-sectional Questionnaire Study (e33661)
Bhavneet Walia, Jeeyoon Kim, Ignatius Ijere, Shane Sanders. 292

A Web-Based Escape Room to Raise Awareness About Severe Mental Illness Among University Students: Randomized Controlled Trial (e34222)
Jose Rodriguez-Ferrer, Ana Manzano-León, Adolfo Cangas, Jose Aguilar-Parra. 371

Modification in the Motor Skills of Seniors in Care Homes Using Serious Games and the Impact of COVID-19: Field Study (e36768)
Jana Kleschnitzki, Inga Grossmann, Reinhard Beyer, Luzi Beyer. 382

Reviews

Outcomes, Measurement Instruments, and Their Validity Evidence in Randomized Controlled Trials on Virtual, Augmented, and Mixed Reality in Undergraduate Medical Education: Systematic Mapping Review (e29594)
Lorainne Tudor Car, Bhone Kyaw, Andrew Teo, Tatiana Fox, Sunitha Vimalesvaran, Christian Apfelbacher, Sandra Kemp, Niels Chavannes. . .
3 0 3

Virtual Reality Applications in Chronic Pain Management: Systematic Review and Meta-analysis (e34402)
Lisa Goudman, Julie Jansen, Maxime Billot, Nieke Vets, Ann De Smedt, Manuel Roulaud, Philippe Rigoard, Maarten Moens. 320

Using Video Games to Improve the Sexual Health of Young People Aged 15 to 25 Years: Rapid Review (e33207)
Ignacio Franco Vega, Anastasia Eleftheriou, Cynthia Graham. 341

Innovative Technology–Based Interventions to Reduce Stigma Toward People With Mental Illness: Systematic Review and Meta-analysis (e35099)
Matías Rodríguez-Rivas, Adolfo Cangas, Laura Cariola, Jorge Varela, Sara Valdebenito. 359

Original Paper

The Effectiveness of a Serious Game (MemoreBox) for Cognitive Functioning Among Seniors in Care Facilities: Field Study

Jana Marina Kleschnitzki[1], MSc; Luzi Beyer[2], Prof Dr; Reinhard Beyer[1], Prof Dr; Inga Großmann[3], DrPH

[1]Department of Psychology, Humboldt-University of Berlin, Berlin, Germany

[2]Department of Social Work, Alice-Salomon-Hochschule, Berlin, Germany

[3]Department of Business Psychology, University of Applied Sciences, Berlin, Germany

Corresponding Author:
Jana Marina Kleschnitzki, MSc
Department of Psychology
Humboldt-University of Berlin
Rudower Chaussee 18
Berlin, 12489
Germany
Phone: 49 15231964606
Email: janakleschnitzki@posteo.de

Abstract

Background: Serious games have been found to have enhancing and preventative effects on cognitive abilities in healthy older adults. Yet, there are few results on the effects in older seniors with age-related low cognitive impairments. Their special needs were considered when designing and using innovate technology in the area of prevention, which is especially relevant owing to the continuously aging population.

Objective: The objective of this study was to evaluate the impact of a serious game on the cognitive abilities of seniors in order to potentially implement innovative resource-oriented technological interventions that can help to meet future challenges.

Methods: In this controlled trial, we tested the serious game MemoreBox, which features modules specifically designed for seniors in nursing homes. Over a period of 1 year, we tested the cognitive abilities of 1000 seniors at 4 time points using the Mini-Mental Status Test. Only half of the participating seniors engaged with the serious game.

Results: The study included an intervention group (n=56) and a control group (did not play; n=55). Based on the in-game data collection, a second intervention group (n=38) was identified within the original intervention group, which exactly followed the planned protocol. There were no noteworthy differences between the demographic and main variables of the overall sample. The large reduction in the sample size was due to the effects of the COVID-19 pandemic (drop-out rate: 88.9%). The CI was set at 5%. Mixed analysis of variance (ANOVA) between the cognitive abilities of the intervention and control groups did not show a statistically significant difference between time and group ($F_{2.710,295.379}$=1.942; *P*=.13; partial η^2=0.018). We noted approximately the same findings for mixed ANOVA between the cognitive abilities of the second intervention and control groups ($F_{3,273}$=2.574; *P*=.054; partial η^2=0.028). However, we did observe clear tendencies and a statistically significant difference between the 2 groups after 9 months of the intervention ($t_{88.1}$=−2.394; *P*=.02).

Conclusions: The results of this study show similarities with the current research situation. Moreover, the data indicate that the intervention can have an effect on the cognitive abilities of seniors, provided that they regularly play the serious game of MemoreBox. The small sample size means that the tendency toward improvement cannot be proven as statistically significant. However, the tendency shown warrants further research. Establishing an effective prevention tool as part of standard care in nursing homes by means of an easy-to-use serious game would be a considerable contribution to the weakened health care system in Germany as it would offer a means of activating senior citizens in partially and fully inpatient care facilities.

Trial Registration: German Clinical Trials Register DRKS00016633; https://tinyurl.com/2e4765nj

(JMIR Serious Games 2022;10(2):e33169) doi:10.2196/33169

KEYWORDS

serious game; cognitive function; mental health; seniors; care facilities; aging; cognitive impairments; health technology

Introduction

The Social Challenge of an Aging Population

The demographic changes in mainly industrialized societies include aging of the population because of falling birth rates, increased life expectancy, improved medical care, and adequate nutrition [1,2]. The proportion of people aged 65 years is forecasted to increase from 17% in 2008 to an estimated 30% by 2060, and the number of people aged over 80 years would triple [3]. This age shift is leading to a rapidly growing proportion of the population in need of medical and general care given that the risk of acute and chronic diseases increases with age [4-6]. At the same time, the proportion of healthy caregivers is decreasing [7]. The aging process and the reasons for entry into the need for long-term care are highly individualized, and in Germany, they usually lead to accommodation in nursing homes [8].

Besides losses in physical abilities, cognitive functions in particular are subject to strong age-associated degradation processes in people over 65 years. Population-based epidemiological studies report that 3% to 19% of adults over 65 years of age experience mild cognitive impairment, with more than half of these adults developing dementia within 5 years [9]. Both age-related physical and cognitive impairments lead to increased social isolation in older people in nursing homes [10,11], which in turn has a negative impact on mental health [12]. Empirical research shows that cognitive deterioration processes can be counter-balanced; cognitive training can help healthy older people to significantly improve their cognitive performance and maintain these gains in the long term [13,14].

Different areas of cognitive functioning, such as memory [15], processing speed [16], executive functions [17], and attention [18], can be successfully trained in older people. These findings highlight the relevance and need of prevention, as well as the early detection of cognitive decline [19].

Opportunities of Prevention

The World Health Organization [20] defines health promotion (prevention) as “[...] a process that enables all people to have a greater degree of self-determination about their health and thus empowers them to strengthen their health.” Numerous studies on demographic change point out that the intensity of prevention and health promotion measures, especially in nursing homes, must increase [21]. Research findings provide a clear indication that health-promoting interventions in long-term inpatient care should not only counteract existing physical and cognitive limitations but also promote remaining skills and health resources [22]. Finally, such interventions would also have positive effects on the number of cases and thus the health care system overall [23].

Serious Games for Prevention

Serious games are games which, in addition to the fun of the game, have a serious added value. They are suitable for not only imparting knowledge, but also prevention, therapy, and use in care [24]. Although studies have shown that seniors play computer games [25], it is noticeable that there are only a few studies on serious games and older people, with an even more limited number of studies including older adults in care facilities (also limited availability) and even fewer studies on nursing home residents with cognitive impairments [26]. Serious games for seniors are developed with special features. For example, attention must be paid to simple operation, customizable development potential, intuitive and easy-to-remember game mechanics, and game principles [27,28]. Lau et al [29] also emphasized the combination of cognitive challenge and stimulation for physical activity, which makes serious games a multifactorial and extensive prevention medium. Thus, training with serious games provides synergy effects that cannot be achieved with separate training of physical and cognitive skills [30]. Despite the limited research, Chao et al [31] found 22 studies for their meta-analysis that investigated the effects of serious gaming on older people. The results showed promising indications that serious gaming in older people has a positive impact on physical skills, cognition, and psychosocial experience. There are also some studies specifically on serious gaming, cognition, and older people. So far, these studies have yielded diverse results. A very recent meta-analysis [32], which examined 18 randomized studies about the effectiveness of serious gaming on the cognition of older people, reported moderate effects on general cognitive functions and memory. In addition, Mura et al [33] showed in their meta-analysis that people with neurological disabilities benefit from serious gaming in their cognitive functions, especially in executive functions and visuospatial perceptions, but there was no effect on attention and global cognition. These results demonstrate that there are few effects for special cognitive functions, but the effect sizes are often not large enough to be beneficial. Moreover, Sala et al [34] conducted publication-bias analyses in meta-analyses regarding this topic, which suggested that the actual effect of exergames on overall cognitive function is small or even nonexistent.

Pilot Study on MemoreBox: A Serious Health Game

Launched in 2014, MemoreBox is a science-based, computer-based, gesture-controlled game console, specifically adapted to the needs and abilities of seniors in nursing homes. It aims to promote, activate, and reactivate the seniors’ cognitive, motor, and psychosocial skills in a preventive and therapeutic fashion (RetroBrain R&D, developing company of MemoreBox). In a German pilot study with 72 seniors [35], which was conducted over a period of 6 months, MemoreBox players showed more beneficial development of their cognitive performance, among other health effects, compared with nonplayers (effect of time: playing intervention group, n=28 [χ^2_2=12.653; P=.002; r=2.39]; nonplaying control group, n=29 [χ^2_2=2.495; P=.29]). The positive pilot led to a longitudinal nationwide German study on MemoreBox to evaluate its physical, cognitive, and mental health effects. A detailed methodological discussion of the longitudinal study has been provided elsewhere [36]. This work will evaluate the data from this study relative to cognitive aspects.

The goal of the study is to investigate the impact of the serious game MemoreBox on the cognitive abilities of seniors. Previous considerations with regard to demographic change highlight the need for knowledge about the preventive effect of interventions (physical and psychological). Thus, studies, such as the one presented in this paper, are key to contribute to a growing knowledge base in this area. In order to meet future challenges, it will be essential to turn to novel, resource-oriented, and technology-based ideas that offer possible solutions. One of these ideas could be serious games, and they have the potential to reactivate existing cognitive resources in senior citizens, which is a topic that we aim to explore further in this paper.

Methods

Study Design

The current intervention study has a quasiexperimental design with 8 repeated measures. To evaluate the effectiveness of MemoreBox, a large-scale study was conducted in 100 German care facilities, recording and examining a total of 1000 seniors in a playing intervention group and a nonplaying control group over the course of a year. Seniors were allocated to either the intervention group or control group based on their preference, that is, if they wanted to play serious games or just participate in the study without playing.

Intervention

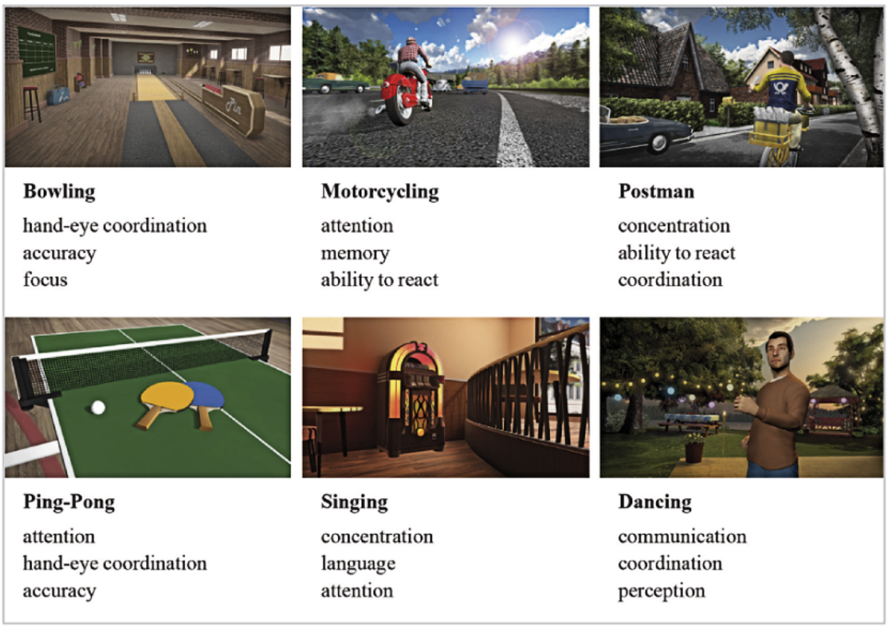
MemoreBox (see detailed description above), a serious game that was specifically designed for seniors in care facilities, constituted the focus of the intervention study. MemoreBox was developed a few years ago by health care professionals (physical therapists, medical doctors, and psychologists) based on empirical research. It is a purely computer-based digital intervention that requires the handling of technology, but was designed in a way that it would be as accessible as possible through pure gesture control. Its technical features allow the recording of a player's movements, including, for instance, inclinations, movement angles, body balance, and movement radius. These movement data can be used in future studies to analyze potential changes in movement or agility. In addition, there is a point system included in the games to promote motivation, and the level of difficulty (sensitivity) can be adjusted.

The intervention included 6 coordinated movement games, which combine preventive, therapeutic, and rehabilitative aspects, aimed at playfully training the physical, mental, and social skills of seniors. Every senior has an individual QR code that records their (movement) data by means of a Kinect camera (Microsoft Corp). The games can be played while sitting or standing, individually or in groups.

The intervention for this study, training through games, was carried out 3 times a week for 1 hour each in a group using a fixed training plan that was developed in advance by therapists. This ensured that each participant used each game (with their different therapeutic foci) once a week. Each training session was supervised by a trained therapist. The preventive training program can be used independently of age-related indications. As shown in Figure 1, there are currently 6 games (motorcycling, bowling, table tennis, singing, postman, and dancing) that are based on everyday activities and focus on different combinations of balance, memory, ability to react, hand-eye coordination, and flexibility.

In addition, MemoreBox provides a record of the amount, duration, and regularity of the game recorded by each person. The data allowed us to differentiate people in the intervention group and subdivide them further into those who regularly played according to the training plan and those who belonged to the intervention group but only exercised irregularly (see the Participants section below).

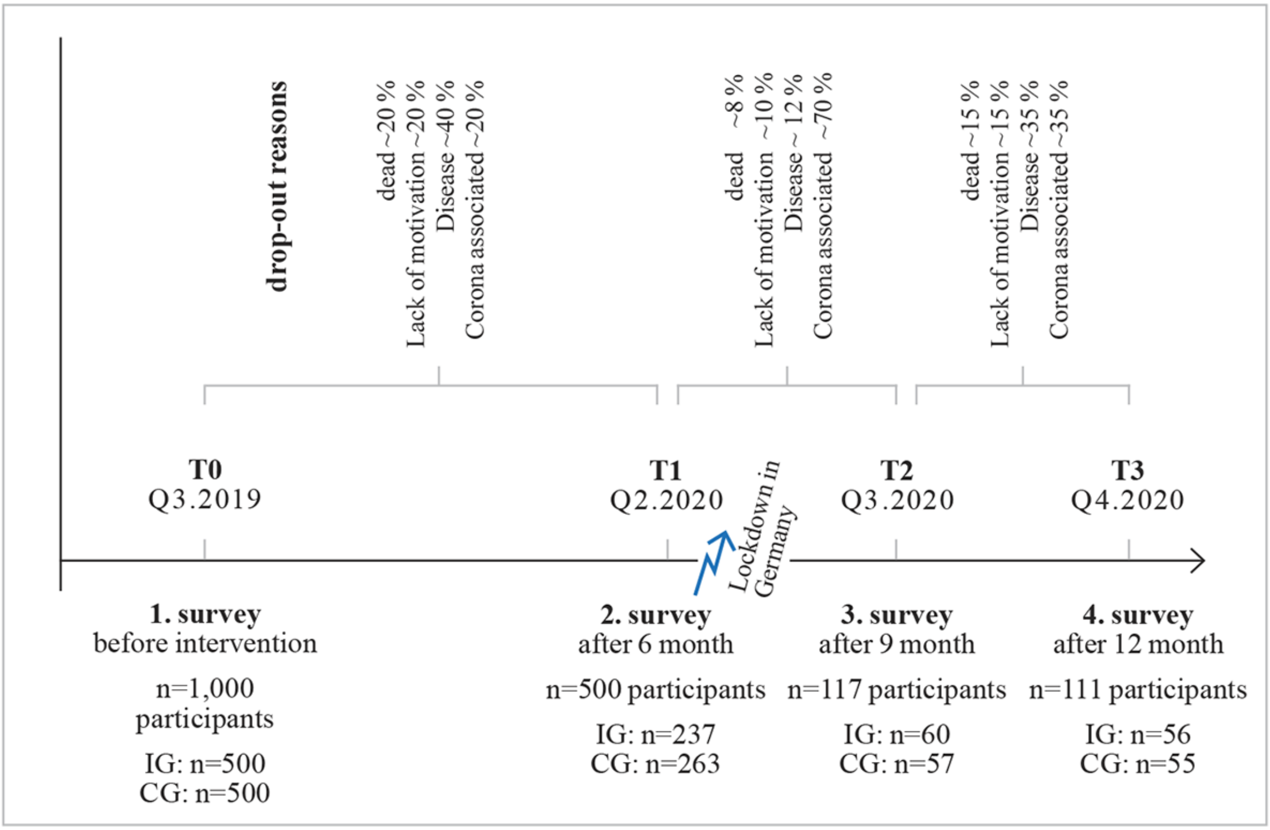
Figure 1. Game modules, and some of their potential therapeutic characters.



Operationalization

Standardized questionnaires on cognition, motor skills, and psychosocial health were administered every 3 to 6 months and were thus completed by participants at 4 to 5 points in time. Given that only cognitive skills are addressed in this study, the corresponding Mini-Mental Status Test (MMST), which has been described in more detail elsewhere [37], was used. It is a screening method for identifying cognitive deficits and monitoring progress. The German-language version was created by Kessler et al [38]. With 30 items, the MMST checks cognitive performance in the following 5 dimensions: temporal and spatial orientation, attention and arithmetic ability, memory, language comprehension, and ability to act. The reliabilities provided in the manual for observer agreement (*r*=0.83) and retest (*r*=0.89; 24 hours apart) are high. There were 4 measurements (Figure 2), where trained nursing staff carried out the MMST. In addition, the objective technical data of the gaming behavior (duration and frequency) were present and were needed.

Figure 2. Survey times, periods of the operationalization and drop-out reasons. CG: controls group; IG: intervention group.



Participants

The sample consisted of an intervention group and a control group. The intervention group included 56 participants (38 females and 18 males) aged between 62 and 96 years (mean age 81.84 years, SD 6.78 years). For this study, the technical data from MemoreBox on gaming behavior were used to identify those within the intervention group who played regularly (according to the training plan) in order to obtain more precise insights into the effects of the training. Another intervention group (second intervention group) of 38 senior citizens, who trained regularly for 1 year, was identified. Of the 38 participants in the second intervention group (regular end of the game), 28 (74%) were females and 10 (26%) were males, with an age range of 71 to 96 years (mean 83.61 years, SD 5.38 years). The control group included 55 participants (41 females and 14 males) aged between 60 and 107 years (mean age 84.24 years, SD 9.33 years). The distributions of age (mean 83.03 years, SD 8.2 years) and gender (66 females, 71%; 28 males, 29%) of the total sample roughly corresponded to the current findings on the need for care in Germany [39].

All participants were residents of care facilities in Germany. A total of 100 care facilities were included. When selecting the facilities, care was taken to ensure that the facility structure was very diverse (eg, private, state, church, city, and country). For both groups, severe mental or neurological illnesses and age below 60 years were exclusion criteria. In addition, the health status, comorbidities, and medications in both groups were surveyed.

Recruitment

After the nursing staff of the facilities passed on the information materials to the residents, the seniors could volunteer to participate in the study. A similar procedure was used to assign participants to the intervention or control groups. The participants voluntarily assigned themselves to 1 of the 2 groups. Given that this voluntary group assignment led to unequal groups with different parameters, the group data were parallelized for comparison purposes after data collection was completed.

Dropout and Missing Data Resulting From the COVID-19 Pandemic

The global COVID-19 pandemic, which has caused major changes in care facilities in Germany at least since March 2020, led to the inevitable interruption of the study and thus of the planned surveys in most of the participating care facilities. Due to the large number of nationally participating institutions, the individual on-site situation and the local restrictions imposed because of COVID-19 differed greatly from one care facility to the next. In 11 nursing homes, the study design could be continued unhindered based on the training plan because of special facility conditions, which ultimately led to a sample size of 111 (56 in the intervention group and 55 in the control group) in this study (Figure 2). The surveys were subject to a very brief interruption (3 months), which led to an analyzable and sufficiently interpretable data set in the sense of the originally intended study design for 4 measurement times with different intervals. Additionally, the intervention group (n=56) was further divided into the following 2 groups: a group of participants for whom the data in MemoreBox indicated that they had participated in the training relatively regularly despite the pandemic conditions (second intervention group, n=38) and a group of participants for whom the data showed that they had participated irregularly (n=18).

The MMST was originally planned to be administered every 6 months. However, due to the changes in the nursing homes as a result of COVID-19, an additional measurement was carried out after the first easing of the lockdown in the summer of 2020, to have a kind of “second start value.” This step resulted in 4 MMST measurement times, with the interval between the first 2 measurements being 6 months and the further 2 intervals being 3 months each (Figure 2).

Statistical Analysis

To ensure comparability, the intervention and control groups were compared and parallelized in terms of the characteristics of the dependent variables at baseline (T0). For the statistical significance test, the confidence interval was set at a level of α=.05. To examine the 2 intervention groups in comparison with the control group, a mixed analysis of variance (ANOVA) was used. The dependent variable was cognitive impairment, the time factor with 4 values (T0 [Q3 2019], T1 [Q2 2020], T2 [Q3 2020], and T3 [Q4 2020]) functioned as an inner-subject factor, and group membership functioned as a between-subject factor (2 factor levels).

Ethics Approval

This study was approved by the Ethics Committee of Charite Berlin (Ethikausschuss am Campus Benjamin Franklin; review number: EA4/035/19).

Results

Overview

We outline the results and separate the intervention group into the original intervention group and the second intervention group (ie, those seniors who completed the training for a year according to the plan).

Baseline Comparison/Descriptive Statistics

There were no noteworthy differences between the demographic and main variables in the overall sample. Moreover, in the newly created classification, the groups at time T0 (baseline) did not differ significantly in demographic and main variables (Tables 1 and 2).

Table 1. Baseline measurement of variables and their differences between the intervention group and control group.

Variable	Intervention group (N=56), mean (SD)	Control group (N=55), mean (SD)	*t (df)*	*P* value	95% CI
Age	81.84 (6.78)	84.24 (9.33)	−1.55 (98.58)	.13	−0.67 to 0.08
Care level[a]	2.49 (0.89)	2.67 (0.83)	−1.02 (99)	>.99	−0.60 to 0.19
Financial medium score[b]	2.03 (0.88)	2.00 (1.04)	0.121 (68)	.90	−0.44 to 0.50
Health condition score[c]	3.28 (1.07)	3.10 (1.27)	0.79 (103)	.43	−0.23 to 0.54
Health behavior score[d]	2.34 (0.79)	2.53 (0.97)	−1.11 (109)	.27	−0.58 to 0.16
Health assessment score[e]	2.84 (0.80)	2.89 (0.83)	−0.33 (109)	.74	−0.44 to 0.31
MMST[f] mean score	1.32 (0.25)	1.29 (0.26)	0.536 (109)	.59	−0.27 to 0.47
MMST total score	24.77 (4.60)	24.47 (4.86)	0.329 (109)	.74	−0.31 to 0.43

[a]0 (no need for care) to 5 (most severe impairments).

[b]0=<€1000; 1=€1000-€1500; 2=€1500-€2000; >€2000.

[c]0 (healthy) to 5 (chronically ill).

[d]0 (not taking care of their health) to 5 (taking great care of their health).

[e]0 (“I rate my health as very bad”) to 5 (“I rate my health as very good”).

[f]MMST: Mini-Mental Status Test.

Table 2. Baseline measurement of variables and their differences between the second intervention group (regular players) and control group.

Variable	Second intervention group (N=38), mean (SD)	Control group (N=55), mean (SD)	*t* (*df*)	*P* value	95% CI
Age	83.61 (5.38)	84.24 (9.33)	−0.412 (88.54)	.68	−0.49 to 0.34
Care level[a]	2.42 (0.94)	2.67 (0.83)	−1.30 (82)	.20	−0.72 to 0.15
Financial medium score[b]	2.00 (0.83)	2.00 (1.04)	0.00 (58)	>.99	−0.51 to −0.51
Health condition score[c]	3.41 (1.04)	3.10 (1.27)	1.22 (86)	.23	−0.16 to 0.68
Health behavior score[d]	2.37 (0.75)	2.53 (0.97)	−0.84 (91)	.40	−0.59 to 0.24
Health assessment score[e]	2.89 (0.73)	2.89 (0.83)	0.02 (91)	.98	−0.41 to 0.42
MMST[f] mean score	1.32 (0.23)	1.29 (0.26)	0.60 (91)	.55	−0.29 to 0.54
MMST total score	24.76 (4.25)	24.47 (4.86)	0.30 (91)	.77	−0.35 to 0.48

[a]0 (no need for care) to 5 (most severe impairments).

[b]0=<€1000; 1=€1000-€1500; 2=€1500-€2000; >€2000.

[c]0 (healthy) to 5 (chronically ill).

[d]0 (not taking care of their health) to 5 (taking great care of their health).

[e]0 (“I rate my health as very bad”) to 5 (“I rate my health as very good”).

[f]MMST: Mini-Mental Status Test.

Outcome

There was no normal distribution in the sample, which can be neglected with a sample size of >30 [8]. In addition, the *F* tests carried out showed similar results. The sphericity was also not given (0.04). Owing to the violation of this requirement, a Greenhouse-Geisser correction of the degrees of freedom was carried out.

First, we analyzed the whole sample (N=111) to explore if there was a potential distinction between the intervention group and the control group. The mixed ANOVA showed no statistically significant interaction between time and group membership ($F_{2.710,295.379}$=1.942; *P*<.13; partial η²=0.018). There was also no significant main effect for time, which corresponded to no significant difference over time ($F_{2.710,295.379}$=0.383; *P*=.75; partial η²=0.04; Figure 3). There was also no significant main effect for group membership ($F_{1.109}$=2.405; *P*=.12; partial η²=0.022).

The clear tendencies of the MMST estimates of both participant groups, which are clearly shown in Figure 3 but were not found to be statistically significant, were another reason for a closer look at the intervention group and the decision to further divide participants based on the available MemoreBox data, that is, creating the second intervention group (n=38) based on the amount, duration, and regularity of game play.

On performing mixed ANOVA for group differences between the second intervention group and the control group, no statistically significant interaction between time and group membership was found ($F_{3.273}$=2.574; *P*<.054; partial η²=0.028; Figure 4).

However, the level of significance was only just exceeded, and at a level of significance of 10%, a clear main effect and thus an interaction of time and group affiliation could be identified. Hence, we could conclude that there were clear tendencies that playing with MemoreBox over a year improved the cognitive abilities of the participants who played regularly, whereas they deteriorated in the control group both in real numbers and in statistical comparison. The main effect of time ($F_{3.273}$=0.337; *P*=.78; partial η²=0.004) and group membership ($F_{1.91}$=2.701; *P*=.10; partial η²=0.029) for this ANOVA did not show statistical significance. The subsequent *t* tests (Table 3) showed a statistically significant difference between the groups in terms of their MMST values after 9 months of the intervention ($t_{88.1}$=−2.394; *P*=.02). This effect was already apparent after 6 months, but was easily canceled out after 12 months.

Figure 3. Results of the Mini-Mental Status Test (MMST) of the intervention group (all participants) and control group over time.

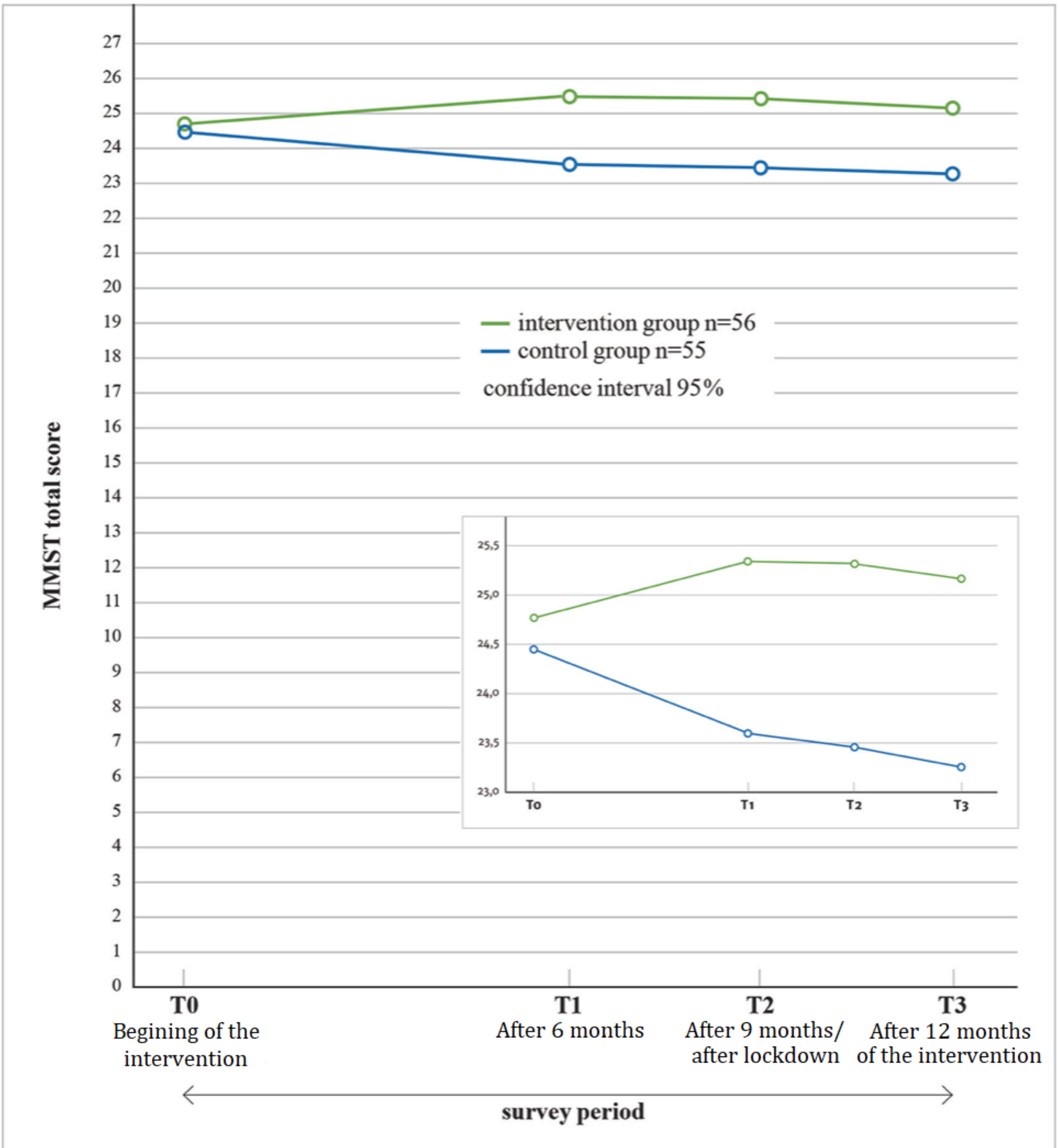


Figure 4. Results of the Mini-Mental Status Test (MMST) of the second intervention group playing regularly and the control group at all 4 measurement times.

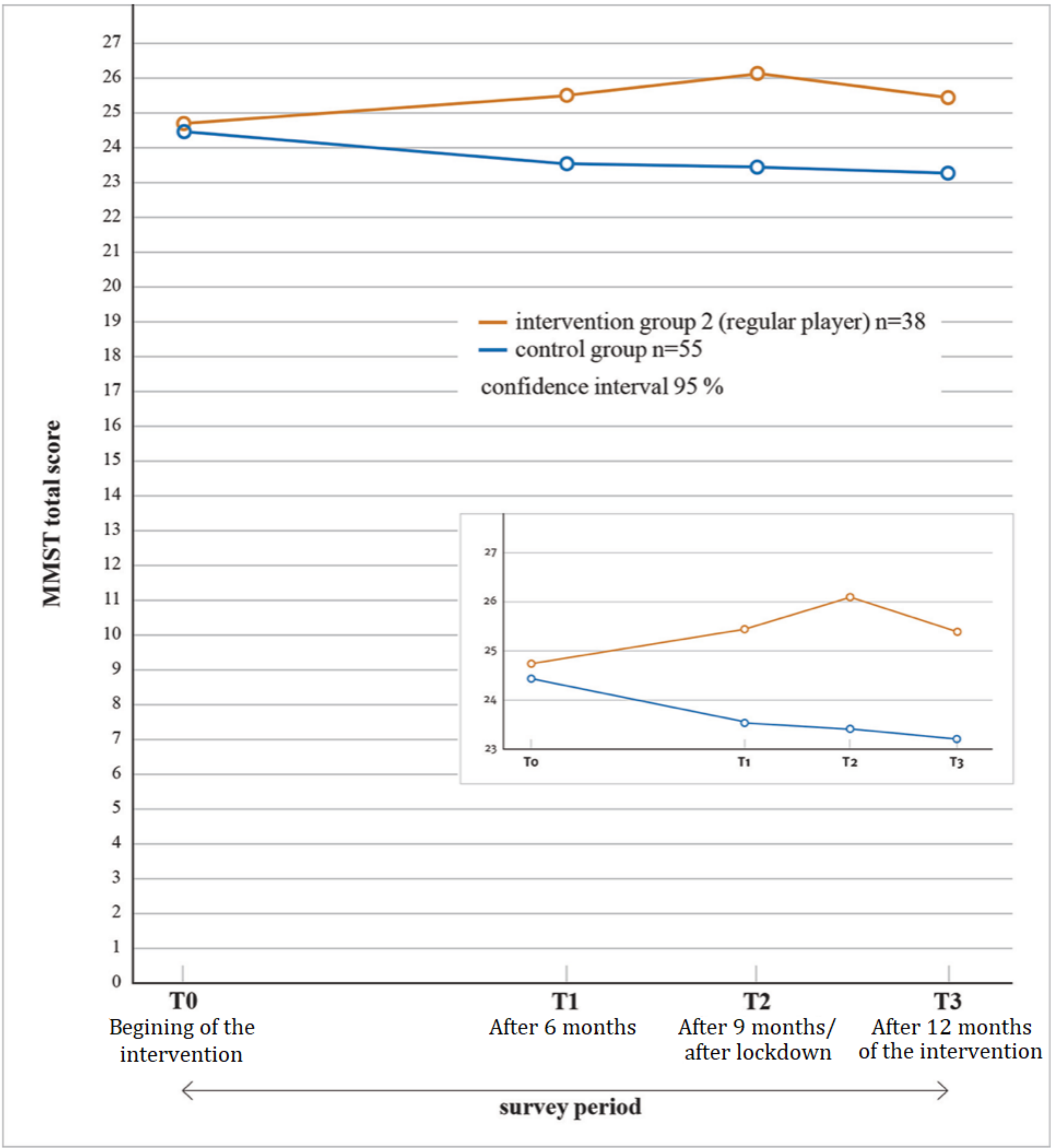


Table 3. Results of the *t* test for the intervention and control groups at 4 different measurement times.

Time point comparison[a]	Control group		Second intervention group		Control group vs second intervention group		
	t (df)	*P* value	*t (df)*	*P* value	Time point	*t (df)*	*P* value
T0–T1	1.433 (54)	.16	−1.341 (37)	.19	T0	−0.298 (91)	.77
T1–T2	0.379 (54)	.71	1.156 (37)	.26	T1	1.725 (89.7)	.09
T2–T3	−0.221 (54)	.83	1.357 (37)	.18	T2	2.349 (88.1)	.02[b]
T0–T3	1.681 (54)	.10	−0.893 (37)	.38	T3	1.597 (91)	.11

[a]T0: Q3 2019, T1: Q2 2020, T2: Q3 2020, T3: Q4 2020.

[b]Significant at 95% CI; N=93.

Discussion

Principal Findings

The baseline comparison showed that the intervention and control groups did not differ statistically in any variable at the beginning of the intervention and were therefore deemed comparable. The relevant sociodemographic variables roughly corresponded to the current findings about the need for long-term care in Germany [39].

The mixed ANOVA showed no significant impact on the cognitive abilities of those subjects who played regularly over the course of a year compared with the control group. However, owing to the size of the sample and the statistical significance, we hypothesize that the intervention had a potential impact on cognition.

This impact could not be detected among irregular players. An independent samples *t* test between the groups after 9 months of the intervention showed a statistically significant difference between the groups, which was not visible at earlier measurement times.

Furthermore, the figures and graphics indicate development tendencies in a way that the control group estimates decrease over the period of the investigation (1 year) from an initial average MMST value of 24.47 to an average MMST value of 23.25. However, this change was not statistically significant. Nevertheless, this development was understandable, given that a decline in cognitive abilities can be expected for this group of participants (ie, senior citizens) within 1 year [40-42]. In addition, within a year, the MMST values for the total sample fell below the frequently set limit value for normal cognitive function (24 points), which may indicate normal dementia development. However, by contrast, the values of the second intervention group showed an opposite trend (from 24.76 to 25.42).

Overall, despite the described limitations, the results clearly showed a tendency that supports the effectiveness of the intervention and suggests a positive impact on the cognitive abilities of seniors in nursing homes. Thus, the results of this study contribute considerably to our knowledge base in this novel and still underresearched area by providing insights into the challenges and complexities, as well as potential developments and implementations of serious games that can be further explored in future research. Moreover, the results indicate that serious games (ie, here MemoreBox) can have a positive impact on the cognitive abilities of seniors and should therefore be increasingly recognized and implemented to provide opportunities for engaging health promotion.

Limitations

The quasiexperimental design, the innovative character of the project, and the interruption by the global pandemic caused limitations. The study design involved nonrandomized assignment, no given double-blind procedure, and a small sample size.

To conduct a survey study over the long period of 12 months and avoid excessive drop-out rates, it is particularly helpful for research projects in the geriatric field to keep the motivation to participate as high as possible by means of voluntary instead of randomized allocation. The baseline comparison should represent the greatest possible compensation for this limitation.

Another major limitation is the small size of the sample that was analyzed at the end. The attrition from the initial 1000 participants to 111 usable data sets (approximately 10%) was very large. Due to the special circumstances of the participants (age, illnesses, care, and morbidity) and the high effort for the intervention group (1 year, 3 times a week commitment), the target group had a high drop-out rate from the start, which can be expected [43]. The sample size may also be a reason for the lack of statistically significant differences between the groups. A corresponding post-hoc power analysis with an alpha level of 5% and a target power of 80% (90%) showed a minimum number of 179 (231) cases to show a significant interaction effect with moderate effect sizes (ηÇ=0.25) with a mixed ANOVA in 2 groups and 4 measurements.

Comparison With Prior Work

Regarding the cognitive ability of the study participants, the results showed that the condition in the control group deteriorated slightly within 1 year, a finding that is in line with previous research [40-42]. However, it can be assumed that regular players at least maintained their cognitive abilities on average and that the progression of cognitive impairments can be slowed down by training. These effects and tendencies confirmed the findings from a previous pilot study by Trauzettel [35], who also tested the influence of MemoreBox on cognitive abilities in 2 nursing homes and found a significant improvement in the MMST values over a 6-month survey period, as well as a significant difference between the intervention group and control group at the last measurement time. The fact that therapeutic computer-based training had a positive effect on cognitive performance shows consistency with other research findings [44]. This is particularly important regarding the age-related decline in cognitive abilities [40]. Additionally, it means that success can lie not only in growth, but also in maintaining cognitive abilities [45]. Moreover, the results of this study coincide with other studies in which comparable interventions with therapeutic game consoles brought about multidimensional improvements in older people [46-49]. In addition, reviews reflect the high potential of serious games as an efficient and motivating component in prevention and health promotion [28,29,50-52].

Apart from the already reported contribution of the study to advancing the knowledge base in this new research area and the clear influence of the intervention in older people, it is also important to highlight the “nursing home” as a novel space and focus in research. Given that nursing homes require a considerable amount of social, financial, and health resources, this study focused on the implementation of serious games as a means of potentially aiding with constrains in these resources. It focused on the efficacies and points to positive effects relative to the 3 resources. Thus, our study clearly highlights the importance of this type of research as well as the attention this area of research should receive in the future.

In the following text, it will also be critically discussed why the results presented here do not show statistical significance, as the literature suggests. The initial values of all participants were in a high range of approximately 24 points, which is not considered to be dangerous. These high baseline values may be a reason for the ambiguous significance in the sense of an improvement, since older people with very low cognitive baseline values tend to benefit more from cognitive interventions than people who already show higher values before training [53]. It remains to be seen whether the second intervention group can maintain the generally stable MMST values and whether this would be reflected in a comparison with the control group over a significantly longer intervention period.

Another reason can be personal influences and associated other everyday activities. As previous research shows, cognitive activity throughout the life of a person is seen as an important factor influencing cognitive performance in old age [54,55]. This could mean that the influence of the intervention is less than assumed, since the level of activity in the previous life of the individual played a more important role. Since improvements in cognitive performance are associated, among other things, with increased health-related quality of life [56] and, on the contrary, a cognitive decline is associated with a lower health-related quality of life [57], there might have been other factors that impacted the results. The Hawthorne effect [58] could be one factor, that is, positive effects on the examined variables solely through conscious participation in the study can (unconsciously) motivate the control group to change their behavior. This could have resulted in stabilizing, compensatory, or diffusing effects that do not correctly reflect the characteristics in the constructs. Possible consequences are a lack of reliability of the implemented intervention and a limited construct validity, which could have resulted in an approximation of the measured values for the second intervention group and control group [59].

At the same time, the game concept of MemoreBox is based on stimulating interactivity and communication with other participants. Epidemiological studies have shown that social relationships are highly relevant to cognitive health and can even reduce the risk of death [60]. Being with other people involves cognitive stimulation through verbal and nonverbal communication. A lack of social relationships, on the other hand, can favor dementia [61]. The changing group size from 5 persons initially to, in some cases, only 1 person at the end of the 1-year intervention could therefore also have had an impact on the results.

Future Perspective

Follow-up studies are much needed to evaluate the continued effects of the changes and investigate the effectiveness beyond the current influence of the game. Future studies should take into account a new group, which receives a different multimodal offer (as similar as possible, but, for example, guided by therapists), in order to differentiate more precisely the specific areas of impact of the given serious game. Additionally, analyzing the exact impact on preventive and health-promoting changes could also be the focus of future research, which would contribute further to the literature that has outlined the relatively high benefits of serious games.

Establishing a high scientific standard requires not only a second intervention group, but also various more specialized instruments to document motor movements, which, in turn, can potentially provide insights into the effectiveness of serious games for different motor skill sets.

Furthermore, it is essential to carry out studies that have a larger sample size. Owing to the already mentioned challenges of recruitment and stability of the target group, there is a global lack of studies with a high number of participants. Future studies, like the present one, could obtain a larger sample size by establishing interventions directly in nursing homes.

In addition to focusing on a higher number of participants, future studies should also concentrate on follow-up examinations to document long-term effects and, in general, collect data over a fundamentally longer period of time. The potential reactivation/preventive effects relative to certain motor skills seem to take a long time, especially for the target group, as shown by the results reported in this paper, as well as in other relevant literature. Therefore, the duration of the studies must be adjusted in order to better assess the actual long-term benefits of serious games for the motor skills of seniors.

Conclusions

While confirming the current state of the research field, the results of this study showed that the intervention had an impact on the cognitive abilities of seniors, provided that they regularly played the serious game of MemoreBox. A particularly novel aspect of this study is that it was carried out in the actual care sector. The authors see the study as a continuation of a pilot I study [35]. Both studies have made it their task to consider the behaviors and circumstances of senior citizens in a resource-oriented and setting-related manner and thus to make a scientific contribution to the limited research in this area [35]. The contribution of the paper is that it shows the positive influence of serious games on the cognitive abilities of older people and can thus be seen as an important building block toward better understanding of preventive effects.

Implementing an easy-to-use serious game as an effective (prevention) tool and making it part of the standard care in nursing homes might contribute considerably to the weak health care system, in which there tends to be a lack of activating offers for senior citizens in partially inpatient care facilities [62].

Conflicts of Interest

None declared.

References

1. Felder S. [Health care expenditures and the aging population]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2012 May 21;55(5):614-621. [doi: 10.1007/s00103-012-1469-4] [Medline: 22526849]
2. Brown GC. Living too long: the current focus of medical research on increasing the quantity, rather than the quality, of life is damaging our health and harming the economy. EMBO Rep 2015 Feb 18;16(2):137-141 [FREE Full text] [doi: 10.15252/embr.201439518] [Medline: 25525070]
3. Ageing characterises the demographic perspectives of the European societies - Issue number 72/2008. European Commission. 2008. URL: https://ec.europa.eu/eurostat/web/products-statistics-in-focus/-/KS-SF-08-072 [accessed 2022-03-12]
4. Fisk A, Rogers W. Handbook of Human Factors and the Older Adult. Cambridge, MA: Academic Press; 1997.
5. Jaul E, Barron J. Age-Related Diseases and Clinical and Public Health Implications for the 85 Years Old and Over Population. Front Public Health 2017 Dec 11;5:335 [FREE Full text] [doi: 10.3389/fpubh.2017.00335] [Medline: 29312916]
6. Jacobs K, Kuhlmey A, Greß S, Klauber J, Schwinger A. Pflege-Report 2019 Mehr Personal in der Langzeitpflege - aber woher?. Berlin, Heidelberg: Springer; 2017.
7. Garms-Homolová V. Prävention bei Hochbetagten. In: Kuhlmey A, Schaeffer D, editors. Alter, Gesundheit und Krankheit. Bern: Huber; 2008:263-275.
8. Bortz J, Schuster C. Statistik für Human- und Sozialwissenschaftler. Berlin, Heidelberg: Springer; 2010:978-973.
9. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, et al. Mild cognitive impairment. The Lancet 2006 Apr;367(9518):1262-1270. [doi: 10.1016/S0140-6736(06)68542-5]
10. Hämmig O. Correction: Health risks associated with social isolation in general and in young, middle and old age. PLoS One 2019 Aug 29;14(8):e0222124 [FREE Full text] [doi: 10.1371/journal.pone.0222124] [Medline: 31465506]
11. Nicholson NR. A review of social isolation: an important but underassessed condition in older adults. J Prim Prev 2012 Jun 6;33(2-3):137-152. [doi: 10.1007/s10935-012-0271-2] [Medline: 22766606]
12. Landeiro F, Barrows P, Nuttall Musson E, Gray AM, Leal J. Reducing social isolation and loneliness in older people: a systematic review protocol. BMJ Open 2017 May 17;7(5):e013778 [FREE Full text] [doi: 10.1136/bmjopen-2016-013778] [Medline: 28515187]
13. Neely AS, Backman L. Maintenance of gains following multifactorial and unifactorial memory training in late adulthood. Educational Gerontology 2006 Aug 03;19(2):105-117. [doi: 10.1080/0360127930190202]
14. Schaie KW. Intellectual Development in Adulthood: The Seattle Longitudinal Study. Cambridge, United Kingdom: Cambridge University Press; 1997.
15. Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, Mahncke HW, et al. A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. J Am Geriatr Soc 2009 Apr;57(4):594-603 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.02167.x] [Medline: 19220558]
16. Edwards JD, Wadley VG, Vance DE, Wood K, Roenker DL, Ball KK. The impact of speed of processing training on cognitive and everyday performance. Aging Ment Health 2005 May 19;9(3):262-271. [doi: 10.1080/13607860412331336788] [Medline: 16019280]
17. Uchida S, Kawashima R. Reading and solving arithmetic problems improves cognitive functions of normal aged people: a randomized controlled study. Age (Dordr) 2008 Mar 23;30(1):21-29 [FREE Full text] [doi: 10.1007/s11357-007-9044-x] [Medline: 19424870]
18. Mozolic JL, Long AB, Morgan AR, Rawley-Payne M, Laurienti PJ. A cognitive training intervention improves modality-specific attention in a randomized controlled trial of healthy older adults. Neurobiol Aging 2011 Apr;32(4):655-668 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2009.04.013] [Medline: 19428142]
19. Kolassa I, Glöckner F, Leirer V, Diener C. Neuronale Plastizität bei gesundem und pathologischem Altern. In: Altern gestalten - Medizin, Technik, Umwelt. Berlin, Heidelberg: Springer; 2010:41-65.
20. World Health Organization. Ottawa charter for health promotion. Health Promot Int 1986;1(4):405-405. [doi: 10.1093/heapro/1.4.405]
21. Lippke S, Kuhlmann T. Gesundheitsförderungsmaßnahmen für ältere Menschen mittels neuer Medien. Zeitschrift für Gesundheitspsychologie 2013 Jan;21(1):34-44. [doi: 10.1026/0943-8149/a000082]
22. Schaeffer D, Büscher A. [Options for health care promotion in long-term care: empirical evidence and conceptual approaches]. Z Gerontol Geriatr 2009 Dec 30;42(6):441-451. [doi: 10.1007/s00391-009-0071-3] [Medline: 19787283]
23. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. Am J Public Health 1998 Sep;88(9):1337-1342. [doi: 10.2105/ajph.88.9.1337] [Medline: 9736873]
24. Nguyen T, Ishmatova D, Tapanainen T, Liukkonen T, Katajapuu N, Makila T, et al. Impact of Serious Games on Health and Well-being of Elderly: A Systematic Review. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: 50th Hawaii International Conference on System Sciences; January 4-7, 2017; Hilton Waikoloa Village, HI, USA. [doi: 10.24251/hicss.2017.447]
25. Jahresreport der deutschen Games-Branche. Game. 2020. URL: https://www.game.de/wp-content/uploads/2020/08/game-Jahresreport-2020.pdf [accessed 2022-03-12]

26. Saint-Maurice PF, Troiano RP, Matthews CE, Kraus WE. Moderate-to-Vigorous Physical Activity and All-Cause Mortality: Do Bouts Matter? J Am Heart Assoc 2018 Mar 22;7(6):e007678 [FREE Full text] [doi: 10.1161/JAHA.117.007678] [Medline: 29567764]
27. Wiemeyer J. Spielerische Förderung körperlicher Aktivität von Älteren. Präv Gesundheitsf 2018 Jul 5;13(4):285-291. [doi: 10.1007/s11553-018-0656-z]
28. Wiemeyer J, Kliem A. Serious games in prevention and rehabilitation—a new panacea for elderly people? Eur Rev Aging Phys Act 2011 Dec 08;9(1):41-50. [doi: 10.1007/s11556-011-0093-x]
29. Lau HM, Smit JH, Fleming TM, Riper H. Serious Games for Mental Health: Are They Accessible, Feasible, and Effective? A Systematic Review and Meta-analysis. Front Psychiatry 2016 Jan;7:209 [FREE Full text] [doi: 10.3389/fpsyt.2016.00209] [Medline: 28149281]
30. Smeddinck JD, Gerling KM, Malaka R. Anpassbare Computerspiele für Senioren. Informatik Spektrum 2014 Sep 13;37(6):575-579. [doi: 10.1007/s00287-014-0835-z]
31. Chao Y, Scherer YK, Montgomery CA. Effects of using Nintendo Wii™ exergames in older adults: a review of the literature. J Aging Health 2015 Apr;27(3):379-402. [doi: 10.1177/0898264314551171] [Medline: 25245519]
32. Yen H, Chiu H. Virtual Reality Exergames for Improving Older Adults' Cognition and Depression: A Systematic Review and Meta-Analysis of Randomized Control Trials. J Am Med Dir Assoc 2021 May;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]
33. Mura G, Carta MG, Sancassiani F, Machado S, Prosperini L. Active exergames to improve cognitive functioning in neurological disabilities: a systematic review and meta-analysis. Eur J Phys Rehabil Med 2018 Jun;54(3):450-462. [doi: 10.23736/s1973-9087.17.04680-9]
34. Sala G, Tatlidil KS, Gobet F. Still no evidence that exergames improve cognitive ability: A commentary on Stanmore et al. (2017). Neurosci Biobehav Rev 2021 Apr;123:352-353. [doi: 10.1016/j.neubiorev.2019.11.015] [Medline: 31760047]
35. Trauzettel F. Evaluation präventiver und gesundheitsförderlicher Aspekte von Serious Games im Alter. Humboldt-Universität zu Berlin. URL: https://edoc.hu-berlin.de/handle/18452/23023 [accessed 2022-03-12]
36. Kleschnitzki J, Großmann I, Arndt S, Beyer R, Beyer L. Analysis of a study design which evaluates a serious game for cognitive and motor activation of senior citizens. In: Empirische Evaluationsmethoden. Berlin, Germany: ZeE; 2020:24.
37. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. Journal of Psychiatric Research 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6]
38. Kessler J, Denzler PE, Markowitsch HJ. Mini-Mental-Status Test (MMST). Deutsche Fassung. Göttingen, Germany: Hogrefe; 1990.
39. Matzk S, Tsiasioti C, Behrendt S, Jürchott K, Schwinger A. Pflegebedürftigkeit in Deutschland. In: Jacobs K, Kuhlmey A, Greß S, Klauber J, Schwinger A, editors. Pflege-Report 2020. Berlin, Heidelberg: Springer; 2020:239-277.
40. Park DC, Smith AD, Lautenschlager G, Earles JL, Frieske D, Zwahr M, et al. Mediators of long-term memory performance across the life span. Psychol Aging 1996 Dec;11(4):621-637. [doi: 10.1037//0882-7974.11.4.621] [Medline: 9000294]
41. Oswald W, Gatterer G, Fleischmann U. Gerontopsychologie Grundlagen und klinische Aspekte zur Psychologie des Alterns. Vienna: Springer; 2008.
42. Dechamps A, Diolez P, Thiaudière E, Tulon A, Onifade C, Vuong T, et al. Effects of exercise programs to prevent decline in health-related quality of life in highly deconditioned institutionalized elderly persons: a randomized controlled trial. Arch Intern Med 2010 Jan 25;170(2):162-169. [doi: 10.1001/archinternmed.2009.489] [Medline: 20101011]
43. Chang CH, Yang H, Tang G, Ganguli M. Minimizing attrition bias: a longitudinal study of depressive symptoms in an elderly cohort. Int Psychogeriatr 2009 Oct;21(5):869-878 [FREE Full text] [doi: 10.1017/S104161020900876X] [Medline: 19288971]
44. Schmiedek F, Lövdén M, Lindenberger U. Hundred Days of Cognitive Training Enhance Broad Cognitive Abilities in Adulthood: Findings from the COGITO Study. Front Aging Neurosci 2010;2:27 [FREE Full text] [doi: 10.3389/fnagi.2010.00027] [Medline: 20725526]
45. Krupp LB, Charvet LE. Long-term Cognitive Consequences for Patients With Pediatric-Onset Multiple Sclerosis. JAMA Neurol 2019 Sep 01;76(9):1008-1009. [doi: 10.1001/jamaneurol.2019.0847] [Medline: 31206137]
46. 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]
47. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013 Sep 05;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]
48. Miller KJ, Dye RV, Kim J, Jennings JL, O'Toole E, Wong J, et al. Effect of a computerized brain exercise program on cognitive performance in older adults. Am J Geriatr Psychiatry 2013 Jul;21(7):655-663. [doi: 10.1016/j.jagp.2013.01.077] [Medline: 23602310]
49. Rand D, Kizony R, Weiss P. Virtual reality rehabilitation for all: Vivid GX versus Sony PlayStation II EyeToy. In: Proceedings of the 5th international conference on disability, virtual reality and associated technologies. 2004 Presented at: 5th international conference on disability, virtual reality and associated technologies; September 20-22, 2004; Oxford, UK p. 87-94.

50. Fleming TM, Cheek C, Merry SN, Thabrew H, Bridgman H, Stasiak K, et al. Juegos serios para el tratamiento o la prevención de la depresión: una revisión sistemática. RPPC 2015 Jan 15;19(3):227. [doi: 10.5944/rppc.vol.19.num.3.2014.13904]
51. Smeddinck J, Siegel S, Herrlich M. Adaptive difficulty in exergames for Parkinson's disease patients. In: GI '13: Proceedings of Graphics Interface 2013. 2013 Presented at: Graphics Interface 2013; May 29-31, 2013; Regina, Sascatchewan, Canada p. 141-148. [doi: 10.5555/2532129.2532154]
52. Wiemeyer J. Multimedia in sport – between illusion and realism. WIT Press. 2008. URL: https://www.witpress.com/Secure/elibrary/papers/9781845640644/9781845640644011FU1.pdf [accessed 2022-03-12]
53. Roheger M, Kessler J, Kalbe E. Structured Cognitive Training Yields Best Results in Healthy Older Adults, and Their ApoE4 State and Baseline Cognitive Level Predict Training Benefits. Cogn Behav Neurol 2019 Jun;32(2):76-86. [doi: 10.1097/WNN.0000000000000195] [Medline: 31205121]
54. Hertzog C, Kramer AF, Wilson RS, Lindenberger U. Enrichment Effects on Adult Cognitive Development: Can the Functional Capacity of Older Adults Be Preserved and Enhanced? Psychol Sci Public Interest 2008 Oct;9(1):1-65. [doi: 10.1111/j.1539-6053.2009.01034.x] [Medline: 26162004]
55. Lövdén M, Ghisletta P, Lindenberger U. Social participation attenuates decline in perceptual speed in old and very old age. Psychol Aging 2005 Sep;20(3):423-434. [doi: 10.1037/0882-7974.20.3.423] [Medline: 16248702]
56. Cohen RA, Moser DJ, Clark MM, Aloia MS, Cargill BR, Stefanik S, et al. Neurocognitive functioning and improvement in quality of life following participation in cardiac rehabilitation. Am J Cardiol 1999 May 01;83(9):1374-1378. [doi: 10.1016/s0002-9149(99)00103-4] [Medline: 10235098]
57. Carmelli D, Swan GE, LaRue A, Eslinger PJ. Correlates of change in cognitive function in survivors from the Western Collaborative Group Study. Neuroepidemiology 1997 Dec 3;16(6):285-295. [doi: 10.1159/000109699] [Medline: 9430128]
58. Roethlisberger FJ, Dickson WJ, Wright HA. Management and the Worker: An Account of a Research Program Conducted by the Western Electric Company, Hawthorne Works, Chicago. Cambridge, MA: Harvard University Press; 1993.
59. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin; 2002.
60. Zunzunegui M, Alvarado BE, Del Ser T, Otero A. Social networks, social integration, and social engagement determine cognitive decline in community-dwelling Spanish older adults. J Gerontol B Psychol Sci Soc Sci 2003 Mar 01;58(2):S93-S100 [FREE Full text] [doi: 10.1093/geronb/58.2.s93] [Medline: 12646598]
61. Ballesteros S, Kraft E, Santana S, Tziraki C. Maintaining older brain functionality: A targeted review. Neurosci Biobehav Rev 2015 Aug;55:453-477. [doi: 10.1016/j.neubiorev.2015.06.008] [Medline: 26054789]
62. Blüher S, Kuhlmey A. Demographischer Wandel, Altern und Gesundheit. In: Richter M, Hurrelmann K, editors. Soziologie von Gesundheit und Krankheit. Wiesbaden: Springer VS; 2016:313-324.

Abbreviations

ANOVA: analysis of variance
MMST: Mini-Mental Status Test

Edited by N Zary; submitted 26.08.21; peer-reviewed by M Aksoy, A Teles; comments to author 02.11.21; revised version received 25.12.21; accepted 17.02.22; published 01.04.22.

Please cite as:
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;10(2):e33169
URL: https://games.jmir.org/2022/2/e33169
doi:10.2196/33169
PMID:35172959

Original Paper

A Mobile-based Virtual Reality Speech Rehabilitation App for Patients With Aphasia After Stroke: Development and Pilot Usability Study

Xiaofan Bu[1], MSc; Peter HF Ng[2], PhD; Ying Tong[3], BSc; Peter Q Chen[2], MSc; Rongrong Fan[1], MSc; Qingping Tang[3], MD; Qinqin Cheng[4], PhD; Shuangshuang Li[5], MSc; Andy SK Cheng[6], PhD; Xiangyu Liu[7], PhD

[1]Nursing Teaching and Research Section, Hunan Cancer Hospital/The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China

[2]Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China

[3]Department of Rehabilitation, Brain Hospital of Hunan Province, Hunan University of Chinese Medicine, Changsha, China

[4]Faculty of Medicine, Nethersole School of Nursing, Chinese University of Hong Kong, Hong Kong, China

[5]Department of Nursing, Second Xiangya Hospital of Central South University, Changsha, China

[6]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China

[7]Department of Health Service Center, Hunan Cancer Hospital/The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China

Corresponding Author:
Xiangyu Liu, PhD
Department of Health Service Center
Hunan Cancer Hospital/The Affiliated Cancer Hospital of Xiangya School of Medicine
Central South University
No.283, Tongzipo Road, Yuelu District
Changsha, 410013
China
Phone: 86 18674869736
Email: 979596459@qq.com

Abstract

Background: Stroke has the highest disability-adjusted life-years lost in any disease, and approximately one-third of the patients get aphasia. Computers and tablets are innovative and aid in intensive treatments in speech rehabilitation for patients with aphasia. However, mechanical training limits the help to patients.

Objective: This study aims to provide a framework for an integrated virtual reality (VR) app to provide speech rehabilitation for patients with aphasia.

Methods: The content was generated through an in-depth literature review and discussion with experienced rehabilitation physicians and occupational therapists. We then conducted a 2-round Delphi study with 15 experts from hospitals and universities to rate the content using a 5-point Likert scale. The app was developed by an interdisciplinary team involving VR, medical science of rehabilitation, and therapeutic rehabilitation. Pilot usability testing of this novel app was conducted among 5 patients with aphasia, 5 healthy volunteers, 5 medical staff, and 2 VR experts.

Results: We designed 4 modules of speech rehabilitation: oral expression, auditory comprehension, cognition, and comprehensive application. Our VR-based interactive and intelligent app was developed to provide an alternative option for patients with aphasia. Pilot usability testing revealed user satisfaction with the app.

Conclusions: This study designed and tested a novel VR-based app for speech rehabilitation specifically adapted to patients with aphasia. This will guide other studies to develop a similar program or intelligent system in a clinical setting.

(JMIR Serious Games 2022;10(2):e30196) doi:10.2196/30196

KEYWORDS

virtual reality; speech rehabilitation; stroke; app; Delphi

Introduction

Background

Aphasia is an acquired language impairment following acquired brain injury (ABI) that affects some or all language modalities, including the expression and understanding of speech, reading, writing, and gestures [1]. ABI is a rapidly growing public health problem resulting from traumatic brain injury, stroke, hypoxic-ischemic encephalopathy after cardiac arrest, and brain tumors [2]. Stroke leads to the highest disability-adjusted life-year loss in any disease, with over 2 million new cases annually in China [3]. It is estimated that there are 1.1 million stroke-related deaths in China per year, and this number is increasing [3]. Globally, the number of stroke deaths is projected to rise to 7.8 million by 2030 [4].

Approximately one-third of stroke patients experience aphasia [1]. Patients with aphasia have a higher risk of not returning to work than those without aphasia [5]. It is likely that an individual's inability to reenter the workforce poststroke is due to the presence of aphasia [5]. The incidence of stroke in younger patients was considerably lower than that in the older cohorts; however, it remains on the rise [6], and rehabilitation needs are worthy of attention [5]. It has been reported that 94% of individuals are diagnosed with cognitive or communicative deficits. However, only 45% were referred for speech-language pathology services [7]. Computers and tablets have proposed innovative and intensive treatments for patients with aphasia in language rehabilitation [8-10]. Naming abilities are improved in patients who receive training, whereas no significant improvements have been shown in verbal communication skills [8,10]. A review concluded that many apps identified from the Google Play Store, Apple App Store, and web searches are available to adults with communication disorders for speech-language therapy. However, few have been designed to specifically meet this vulnerable population's engagement, functionality, aesthetics, and information quality [11].

Mechanical training limits the help to patients. Communication activity in real-world settings is more effective for patients in improving their communicative ability [1]. Neither smartphones nor computers can make a patient feel like they are in a real-life environment. The development of virtual reality (VR) technology has created interactive computer-generated worlds through visual, listening, and touch simulations. This makes patients more enthusiastic and more willing to speak. Patients will be much more easily accessible to dangerous, expensive, uncontrollable, and inaccessible environments in VR than in hospitals and will be able to try out new therapeutic strategies [12]. It has been identified that speech treatment based on VR is effective. The novel VR platform, EVA Park, is an online virtual island that contains various simulated locations, including houses, cafés, restaurants, health centers, hair salons, tropical bars, and discos [13]. It is designed to enable patients with aphasia to communicate successfully with 1 or more conversation partners via speech using a headset and microphone in real time [13]. Mirror neuron rehabilitation training software combined with a 4-channel VR panoramic helmet and Z-channel independent training equipment, including training contents of nouns, verbs, phrases, and sentence listening and reading, is effective for patients with Broca aphasia [14]. Additionally, due to the VR system almost completely shielding interference from the outside world, patients focus more on speech treatment [14]. However, these systems are semi-immersive VR environments that require partners and the help of a therapist.

Objectives

This study aims to provide a framework for an integrated VR app and perform a preliminary test of its usability and safety. Few apps are explicitly designed to meet the requirements of engagement, functionality, aesthetics, and information quality [11]. Therefore, we intended to add interactive and engaging elements by providing patients with opportunities to interact with a virtual environment and practice speaking in real-life scenarios.

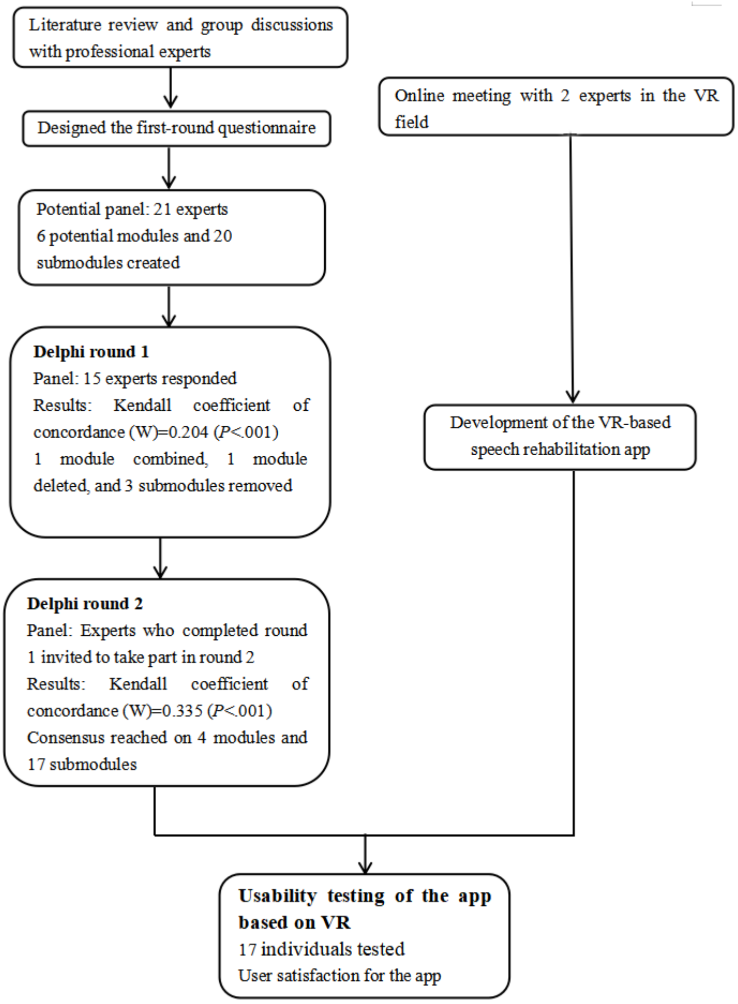
Methods

Phase I

Designing the Contents of VR-Based Language Rehabilitation: a Delphi Study

For the development of speech rehabilitation content based on VR for patients with aphasia after stroke, a 2-round Delphi study was conducted (Figure 1). The Delphi technique is used to obtain the most reliable consensus among a panel of experts using a series of questionnaires [15,16]. Experts can modify, add, or delete content, as appropriate [17]. We listed the module name, submodule name, specific contents, and a short description of the training modes during each Delphi round. Between 2 rounds, we revised the content based on expert feedback. In addition to written descriptions, the drafts included pictures and visual images.

Figure 1. A 2-round Delphi study and development of the app. VR: virtual reality.



Ethics Approval

The Ethics Review Committee of Nursing and Behavioral Medicine Research, School of Nursing, Central South University, approved the study (E202118; approval date: 21 April 2021).

Procedures and Participants

Before round 1 of the Delphi study, we conducted a literature review using the search terms "aphasia," "speech," "speech treatment," "language therapy," "communication," and "rehabilitation" via the Wan Fang Database, Chinese National Knowledge Infrastructure (CNKI), PubMed, and Web of Science in September and October 2020 to identify available content and modes related to language rehabilitation. We developed the content by learning from a model of cognitive communication competence [7]. It includes 7 functioning domains: individual, contextual or environmental, cognitive, communication, physical/sensory, emotional/psychosocial, and communication competence. This model provides evidence that spans the fields of speech-language pathology, psychology, neuroscience, rehabilitation, and education, and concerns the complex interplay between cognitive, communicative, emotional, and physical factors. Rehabilitation platforms, such as Constant Therapy [18] and Aphasia Therapy Online [19], were retrieved. In addition, we learned about existing computer- or tablet-based speech rehabilitation systems.

In November 2020, we held group discussions with experienced rehabilitation physicians and occupational therapists at the Brain Hospital of Hunan Province and the Hunan Cancer Hospital to determine the relationship between brain regions and language production, the rehabilitation process of speech, and their opinions about current speech treatment. In December 2020, we held an online meeting with 2 experts in the Department of Computing at the Hong Kong Polytechnic University to generate ideas about potential content based on VR. Based on the information collected, we developed an initial draft of 20 submodules in 6 modules. We designed a questionnaire for experts to review and use a 5-point Likert scale to evaluate which part should be included. Each piece was described with a name, specific content, and a short description of the training modes.

The panel consisted of experts with both theory-based and practice-based backgrounds to obtain a variety of insights from researchers. Experts with a theory-based background all published papers in the speech treatment or rehabilitation fields. Experts with a practice-based background were selected based on their practical experience and publications. Additionally, we approached our network and asked responding experts to provide the names of essential experts in this field, such as a snowball sampling technique. The experts who participated in the first Delphi round were also approached in the second round. In total, 21 experts were invited by email, WeChat, or an in-person meeting to participate in the 2 rounds of the Delphi study. It contained detailed information about the goal, study procedure, first-round questionnaire, experts' demographic information, and their judgment and familiarity with the corresponding field. Of the 21 experts, 15 (71.4% response rate) responded to the invitation.

First Round

Questionnaire

The first-round questionnaire was composed based on the contents identified by the literature review and group discussions. This resulted in a questionnaire consisting of 6 structured modules and 20 submodules. Experts were invited to rate these contents on a 5-point Likert scale ranging from 1 (not at all important) to 5 (extremely important). They were asked to modify, add, or delete content, as appropriate, and to provide suggestions they supposed were reasonable. The experts' demographic information, as well as their judgment and familiarity with the corresponding field, was required. We reminded those who did not complete the survey following the initial invitation at 2-3-week intervals by WeChat or email.

Data Analysis

Microsoft Excel 2003 and IBM SPSS Statistics (version 26.0) were used for data analysis and processing. The responses of the experts were independently analyzed by 2 researchers, and the data were double-entered to minimize typing errors. We listed and combined similar recommendations. The mean, SD, and coefficient of variation (CV) were calculated. The CV is defined as the SD divided by the mean used to describe the relative dispersion degree of the items' importance evaluation by experts [20,21]. A lower CV value represents a higher degree of coordination among the experts' opinions [21]. After calculating the mean (SD) and CV that the contents should include, we removed all capabilities for which the mean score was less than 4.0 and the CV was over 0.2. Consensus among the experts was evaluated using the Kendall coefficient W test. It refers to the level of intraexpert agreement for all indicators [22]. Statistical significance was considered at a 2-tailed *P* value of <.05 [21]. Cs represents the experts' familiarity with the research field, and Ca represents the judgment criteria the experts are based on [21]. The defined values are listed in Tables 1 and 2. Cr represents the authority coefficient of experts, which is the mean of the sum of familiar Cs and Ca [23]. We then produced the round 2 questionnaire. Complete data were collected from 15 (71.4%) of the 21 panel members, and the results were used to revise the contents and establish a second questionnaire for round 2 of the Delphi exercise.

Table 1. The value of judgment (relevance) criteria.

Criterion	Influence degree		
	Great	Medium	Small
Theory analysis	0.5	0.4	0.3
Working experience	0.3	0.2	0.1
Referring to literatures	0.15	0.1	0.05
Self-intuition	0.05	0.1	0.15

Table 2. The degree of familiarity with content.

Familiarity degree	Very familiar	Familiar	General	Unfamiliar	Very unfamiliar
Self-evaluation	0.9	0.7	0.5	0.3	0.1

Second Round

Questionnaire

In February 2021, we held an online meeting with 2 experts in the VR field to discuss the content based on VR before the second round. Experts who completed the first-round questionnaire were invited to participate in round 2. Again, we reminded those who did not complete the survey following the initial invitation at 2-3-week intervals by WeChat or email. The experts rated each revised section again. They were asked to rate how much they agreed that each element could be used and to comment on each part. Additional advice about the content that they assumed reasonable was required.

Data Analysis

Analysis following round 2 aimed to identify any consensus on the contents and determine whether an additional round was needed. Microsoft Excel 2003 and IBM SPSS Statistics (version 26.0) were used for data analysis and processing. We predefined a mean score of no less than 4, a CV no more than 0.2, and a 2-tailed *P* value of the Kendall coefficient W test of no more than .05 among experts that the contents should be included. The mean, SD, and CV values were calculated for each part. The Delphi survey was completed when all the items in the questionnaire met the aforementioned criteria.

Phase II

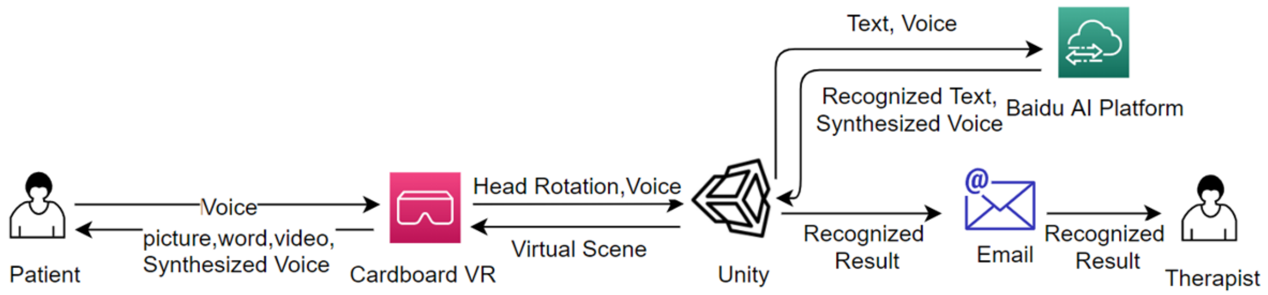
Development of the VR-Based Speech Rehabilitation App

In November 2020 and February 2021, we held online meetings with 2 experts in the VR field to discuss this project. A variety of devices and components can deliver a VR experience, with the main categories including smartphone VR headsets, tethered personal computer (PC)-based VR headsets (eg, HTC Vive), or stand-alone VR headsets (eg, Oculus Quest). HTC Vive and Oculus Quest create a more immersive experience and provide the user with a stronger sense of presence than smartphone VR owing to resolution, frame rate, and sufficient input mechanisms [24]. To provide an alternative option for patients under significant financial pressure, we chose to use smartphone VR and developed a mobile-based VR speech rehabilitation app. Smartphone VR headsets delivered a VR experience through a smartphone fitted on a headset, which was as simple as the original Google Cardboard. These types of VR apps are affordable and cost-effective [24]. Moreover, patients can also use existing smartphones.

The VR scene was deployed on mobile phones using the Google Cardboard VR plug-in. Combining Google Cardboard equipment and mobile phones makes it a portable and cheap VR equipment that most patients can afford. The software can be used to receive language rehabilitation whenever and wherever possible.

The entire VR scene was developed and constructed using the Unity game engine, and Adobe Photoshop was used for 2D assets. Unity is a bridge connecting artificial intelligence (AI) platforms, VR scenes, and back-end data storage. It uploads the designed text content to the Baidu AI platform and obtains a synthesized voice. The synthesized voice is used in the VR scene as a voice instruction to guide the user. Cardboard VR delivers pictures, words, videos, and synthesized voices to patients that substitute therapists in traditional treatment methods. The patients will see the designed questions and relevant materials in the VR scene and will record their voice answers using a mobile phone microphone or register their head rotation as an input for selecting answers in multiselection questions. Unity would then record their responses according to their head rotation or send the recorded voice to the Baidu AI platform to get the recognized text back. The results will be stored in an Excel sheet and sent to the therapist's email address with the patient’s name. Figure 2 elaborates on the detailed VR-based aphasia therapy process.

Figure 2. Detailed VR-based aphasia therapy process. AI: artificial intelligence; VR: virtual reality.



Pilot Application and Usability Testing of the App

We conducted feasibility exercises of the app on 5 patients with aphasia, 5 healthy volunteers, 5 medical staff, and 2 VR experts between 25 and 65 years of age (mean 39.53 years, SD 16.54 years). All 5 patients with aphasia were hospitalized for speech rehabilitation for 2 weeks, and all of them had ischemic stroke. The aphasia quotient (AQ) of the Chinese version of the Western Aphasia Battery ranged from 58 to 75 after 2 weeks of face-to-face training. This feasibility exercise allowed us to better understand and potentially reduce the likelihood of adverse effects in patients with aphasia who may have been more sensitive to sensory effects (eg, motion sickness and discomfort). Interviews were the measurements used in the experiments. The outline of interview questions after the participants’ experience with the mobile-based VR speech rehabilitation app is as follows: (1) What groups do you think this mobile-based VR speech rehabilitation app is suitable for? (2) How does the mobile-based VR speech rehabilitation app differ from face-to-face language rehabilitation? (3) What would encourage or hinder app adoption after using the mobile-based VR speech rehabilitation app? (4) What is your overall rating of the mobile-based VR speech rehabilitation app? All interviews were recorded using a digital voice recorder and transcribed verbatim within 24 hours. The researcher explored participants’ feelings about the app by asking questions in a private environment, and they were encouraged to put forward new questions and content.

Results

Results of the Delphi Study

Expert Panel

Table 3 lists the main characteristics of the 15 (71.4%) of 21 experts who participated in the first and second Delphi rounds. The 15 experts’ academic background included the medical science of rehabilitation or rehabilitation therapy. Of them, 14 (93.3%) were from 3A hospitals, 5 (33.3%) of them took a position in the university, and 2 (13.3%) of them were professors. Detailed demographic characteristics of the Delphi panel are presented in Table 3. The reliability of the experts is presented in Table 4. The mean value of the expert authority coefficient (Cr) was 0.84. The authority of the experts was high; thus, the results of this study are trustworthy.

Table 3. Characterization of experts participating in the Delphi panel (N=15).

Characteristic	Participants, n (%)
Gender	
Male	2 (13.3)
Female	13 (86.7)
Age (years)	
<35	6 (40.0)
≥35	9 (60.0)
Education	
University diploma	9 (60.0)
Master's degree	4 (26.7)
Doctoral degree	2 (13.3)
Academic background	
Medical science of rehabilitation	5 (33.3)
Rehabilitation Therapeutic	10 (66.7)
Working experience (years)	
5-9	5 (33.3)
10-14	5 (33.3)
≥15	5 (33.3)
Main area of the work role	
Clinical	10 (66.7)
Clinical and research	5 (33.3)
Professional title	
Intermediate	12 (80.0)
Deputy senior	1 (6.7)
Senior	2 (13.3)

Table 4. Reliability of experts (Ca[a], Cs[b], and Cr[c]).

Expert number	Criterion score						
	Theoretical analysis	Working experience	Referring to literature	Self-intuition	Ca	Cs	Cr
1	0.4	0.3	0.1	0.1	0.9	0.7	0.8
2	0.5	0.3	0.1	0.1	1	0.7	0.85
3	0.5	0.3	0.05	0.05	0.9	0.5	0.7
4	0.4	0.3	0.15	0.1	0.95	0.9	0.925
5	0.5	0.3	0.1	0.05	0.95	0.9	0.925
6	0.4	0.3	0.1	0.05	0.85	0.9	0.875
7	0.4	0.3	0.15	0.1	0.95	0.7	0.825
8	0.4	0.3	0.15	0.1	0.95	0.7	0.825
9	0.4	0.3	0.1	0.1	0.9	0.7	0.8
10	0.4	0.3	0.1	0.1	0.9	0.9	0.9
11	0.4	0.2	0.05	0.1	0.75	0.7	0.725
12	0.5	0.2	0.05	0.1	0.85	0.7	0.775
13	0.4	0.3	0.1	0.1	0.9	0.9	0.9
14	0.4	0.3	0.1	0.1	0.9	0.9	0.9
15	0.5	0.3	0.15	0.1	1.05	0.9	0.975

[a]Ca: judgment criteria the experts are based on.
[b]Cs: experts' familiarity with the research field.
[c]Cr: authority coefficient of experts.

First Round

The Kendall coefficient of concordance (W) was 0.204 (P<.001). Of the 6 modules and 20 submodules rated by the panel members in the first round (Table 5), the CV of 2 submodules (syllabic and naming with verbs, nouns, and adjectives) and the CV of the arithmetic therapy module that did not meet the criteria were removed (Table 6). In addition, 2 experts recommended to delete the arithmetic therapy module. Pictures and videos describing daily life were added based on the recommendations of 3 panel members. In addition, naming therapy was combined with oral expression therapy based on the advice of 2 panel members. Furthermore, 2 panel members suggested that single characters could be deleted because they are not as easy to express as phrases, so we deleted those. Some respondents suggested additional executive functioning, reading, and writing training. However, this app is meant to improve patients' oral expression abilities, so reading and writing are beyond our consideration, in particular, executive functioning and writing. Due to smartphone VR and technological limitations, we were unable to add writing training to VR. The revised content contained 4 modules and 17 submodules.

Table 5. Modules in round 1.

Module	Task
1	Oral expression therapy
2	Auditory comprehension therapy
3	Cognition therapy
4	Naming therapy
5	Arithmetic therapy
6	Comprehensive application

Table 6. Scores[a] in round 1.

Tasks	Mean (SD)	CV[b]
Oral expression therapy		
Read and repeat task: syllabic	4.13 (0.99)	0.24[c]
Read and repeat task: single characters, including the 150 most commonly used Chinese characters	4.60 (0.91)	0.20[d]
Read and repeat task: some phrases, including categories of digital, fruit, animals, vegetables, transportation, kitchen supplies, daily necessities, body parts, food, address, location, and sports	5 (0)	0.00
Read and repeat task: sentences, including proverbs and daily expressions	4.87 (0.35)	0.07
Answering some questions related to everyday life	4.80 (0.56)	0.12
Auditory comprehension therapy		
Listening and matching: included digital, color, pictures, and words	4.93 (0.26)	0.05
Yes-no questions	4.93 (0.26)	0.05
Listening to the passage and answering questions	4.73 (0.46)	0.10
Cognition therapy		
Attention training	4.27 (0.80)	0.19
Memory training	4.13 (0.74)	0.18
Reasoning and problem solving	4.47 (0.83)	0.19
Naming therapy		
Picture naming	4.93 (0.26)	0.05
Naming with verbs, nouns, and adjectives	4.33 (0.90)	0.21[c]
Naming a list of items with the exact nature	4.93 (0.26)	0.05
Arithmetic therapy		
Addition, subtraction, multiplication, and division	4.27 (1.16)	0.27[c]
Comprehensive application		
Supermarket task: A computer-generated virtual customer purchases goods. The patient has to pick items, count the total amount of money that the virtual customer has selected, and return the change.	4.80 (0.56)	0.12
Interview task: The patient plays the role of a job seeker. A computer-generated virtual interviewer asks questions about the patient's personal information (eg, name, gender, age, nationality, education level, birth date, height, weight, specialty, address, and family members) and assesses their language skills by reading text and describing pictures.	4.73 (0.59)	0.13
Bedroom task: The patient plays the role of a mother. By communicating with the computer-generated virtual daughter, the patient selects clothes for the virtual daughter according to the weather conditions and discusses breakfast and what to choose for dinner and the kind of transportation to use.	4.80 (0.56)	0.12
Ordering task: The patient acts like a customer to order in a virtual restaurant; communicates with a virtual server synthesized by the computer; requests a certain number of dishes, desserts, cakes, and drinks; and completes the payment.	4.80 (0.41)	0.09
Park task: By buying tickets, paying, and asking for directions in the virtual ticket office, together with a computer-generated virtual friend, the patient enters a virtual park. The patient and the virtual friend communicate with each other about the scene.	4.80 (0.41)	0.09

[a]Likert scale: 1, not important; 2, somewhat important; 3, moderately important; 4 = important; and 5, very important.

[b]CV: coefficient of variation.

[c]CV>0.2.

[d]Deleted on the recommendations of the 2 panel members.

Second Round

The results of the second round are presented in Tables 7 and 8 and Figure 3. All mean scores for importance were above 4.00, and the CVs were less than 0.20. The Kendall coefficient of concordance (W) was 0.335 (P<.001). Finally, a consensus was reached on 4 modules and 17 submodules. One expert suggested that patients should complete answering within the allotted time. Some respondents suggested additional executive functioning, reading, and writing training. However, this app is meant to improve patients' oral expression abilities, so reading and writing are beyond our consideration. The revised content contained 4 modules and 17 submodules.

Table 7. Modules in round 2.

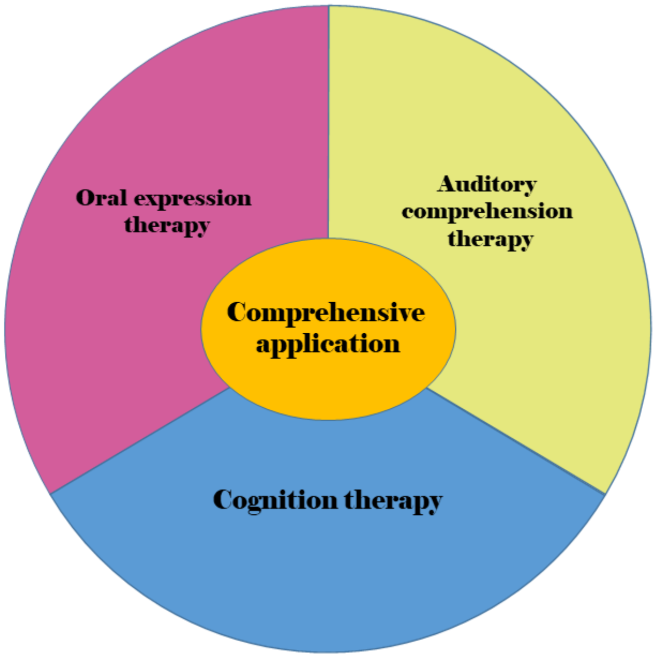
Module	Task
1	Oral expression therapy
2	Auditory comprehension therapy
3	Cognition therapy
4	Comprehensive application

Table 8. Scores in round 2.

Tasks	Mean (SD)	CV[a]
Oral expression therapy		
Read and repeat task: some phrases, including categories of digital, fruit, animals, vegetables, transportation, kitchen supplies, daily necessities, body parts, food, address, location, action, and sports	4.93 (0.26)	0.05
Read and repeat task: sentences, including proverbs and daily expressions	5.00 (0)	0
Answering some questions related to everyday life	4.87 (0.35)	0.07
Naming task: picture naming	5.00 (0)	0
Naming task: naming list of items with the exact nature	4.93 (0.26)	0.05
Describing pictures and videos pertaining to daily life	4.67 (0.49)	0.10
Auditory comprehension therapy		
Listening and matching: included digital, color, pictures, and words	5.00 (0)	0
Yes-no questions	4.93 (0.26)	0.05
Listening to the passage and answering questions	4.80 (0.41)	0.09
Cognition therapy		
Attention training	4.27 (0.70)	0.16
Memory training	4.13 (0.74)	0.18
Reasoning and problem solving	4.20 (0.77)	0.18
Comprehensive application		
Supermarket task: A computer-generated virtual customer purchases goods. The patient has to pick items, count the total amount of money that the virtual customer has selected, and return the change.	4.80 (0.56)	0.12
Interview task: The patient plays the role of a job seeker. A computer-generated virtual interviewer asks questions about the patient’s personal information (eg, name, gender, age, nationality, education level, birth date, height, weight, specialty, address, and family members) and assesses their language skills by reading text and describing pictures.	4.80 (0.56)	0.12
Bedroom task: The patient plays the role of a mother. By communicating with the computer-generated virtual daughter, the patient selects clothes for the virtual daughter according to the weather conditions and discusses breakfast and what to choose for dinner and the kind of transportation to use.	4.87 (0.52)	0.11
Ordering task: The patient acts like a customer to order in a virtual restaurant; communicates with a virtual server synthesized by the computer; requests a certain number of dishes, desserts, cakes, and drinks; and completes the payment.	4.87 (0.52)	0.11
Park task: By buying tickets, paying, and asking for directions in the virtual ticket office, together with a computer-generated virtual friend, the patient enters a virtual park. The patient and the virtual friend communicate with each other about the scene.	4.53 (0.64)	0.14

[a]CV: coefficient of variation.

Figure 3. Consensus on 4 modules.

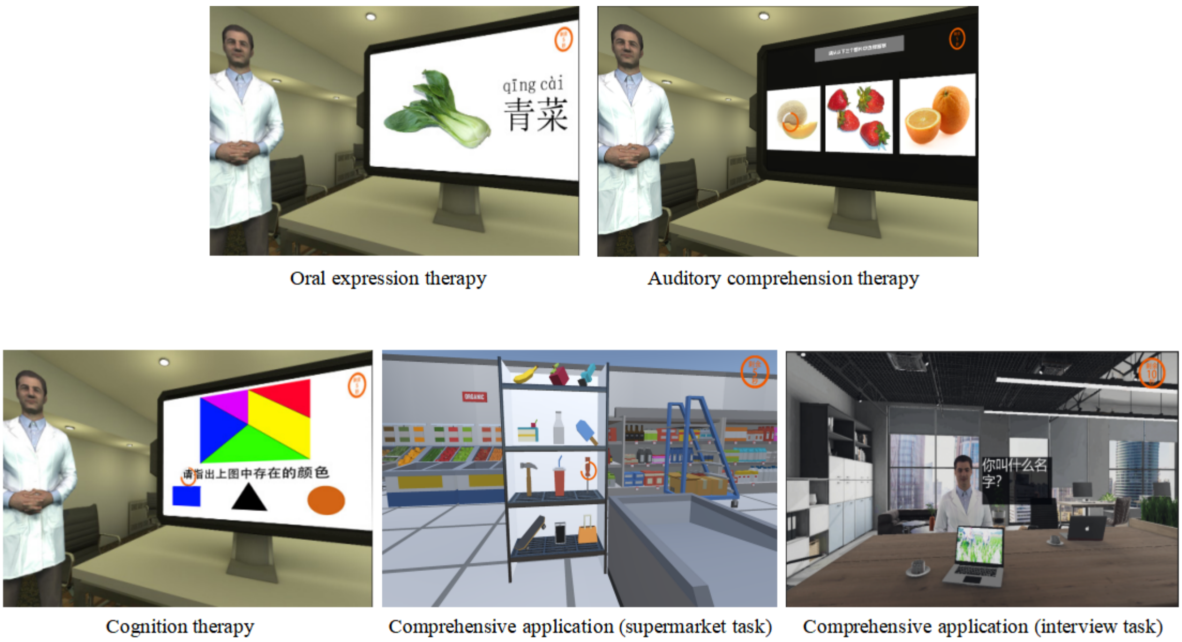


Usability Testing of the App

All 17 (100%) participants could open the app. A user-centered design philosophy was implemented to create additional interactive and customizable features that indicated a high degree of usability for patients with aphasia. Participants' comments were collected and descriptively analyzed based on the interview outline. The 4 responses were classified based on their feedback. Figure 4 shows the functionality of the VR-based aphasia therapy app.

In response to question 1, “What groups do you think this mobile-based VR speech rehabilitation app is suitable for?”, 12 (70.6%) of 17 participants suggested that the user group could include patients who were discharged or currently hospitalized. Participants stated that with this app, they could correct and refine their language skills, and 5 (29.4%) of 17 participants recommended that the user group should include people, especially those who are under significant financial pressure and do not have easy access to medical treatment but would prefer an opportunity to learn.

Figure 4. Functionality of the VR-based aphasia therapy app. VR: virtual reality.



In response to question 2, “How does the mobile-based VR speech rehabilitation app differ from face-to-face language training?”, all 17 (100%) participants mentioned the advantages of in-home apps. Participants stated that this app reduces the difficulty of going to the hospital and is more convenient as they do not have to worry about weather and traffic. In addition, 15 (88.2%) of 17 participants reported a high level of immersion and engagement and thought that patients could thoroughly enjoy themselves when using this app without judgment from others. Additionally, there is no need to worry about self-abasement because of the inability to speak fluently. Moreover, 4 (23.5%) of 17 participants critiqued the effectiveness of this app because patients could not receive immediate feedback, and the efficiency of using this app could not be guaranteed.

In response to question 3, “After using the mobile-based VR speech rehabilitation app, what would encourage or hinder your adoption of the app?”, 15 (88.2%) of 17 participants mentioned that the design and arrangement met patients’ daily communication needs; they felt more secure because the training content was designed by a professional rehabilitation team. In addition, 15 (88.2%) participants mentioned that the app was easy to use, and 3 (17.6%) reported low levels of motion sickness.

In response to question 4, “What is your overall rating of the mobile-based VR speech rehabilitation app?”, all 17 (100%) participants indicated that the app was novel, and 5 (29.4%) of 17 participants gave some recommendations, such as considering different training difficulties and increased functionality. All patients expressed interest in language rehabilitation using the new technology.

Discussion

Principal Findings

This study developed the first accessible and cost-effective mobile-based VR speech rehabilitation app in China. Through a 2-round Delphi study, panel members reached a consensus on 4 modules and 17 submodules that benefit patients in gaining essential daily communication language skills and various skills in everyday life. Patients get opportunities to interact with the virtual environment and can practice at any time. A review systematically identified and evaluated a series of mobile apps for speech-language therapy and found a lack of interactive and engaging elements in the apps, which failed to make patients self-manage [11]. The content we designed based on VR emphasizes user experience, engagement, and visual appeal to improve patient adherence. We developed real-life scenarios and gaming factors for the VR environment. Patients get opportunities to interact with a virtual environment to practice language in natural context communication settings in VR. We believe that the rehabilitation content will provide patients with a strong basis for returning to society and work.

We developed the content by learning from a model of cognitive communication competence [7]. Language processing is the core of cognition and requires the participation of other nonlinguistic cognitive functions [25,26]. Left hemispheric stroke leading to impairments in language processing often affects other cognitive functions, such as executive function, attention, visuospatial perception, logical thinking, and memory [27,28], which play an essential role in aphasia recovery and rehabilitation [29-31]. Patients' language and nonlinguistic cognitive functions cannot be separated owing to complex neurobiological networks [32]. Nonlinguistic cognitive function in patients with poststroke aphasia influences language rehabilitation outcomes [33].

Language rehabilitation combined with cognitive rehabilitation can improve patients' communication skills [34]. The cognitive domain includes processing, attention, working memory, social cognition, reasoning, and problem-solving [7]. Therefore, to improve the patients’ oral expression abilities, we designed cognitive training pertaining to reasoning, memory, attention, visuospatial perception, and logical thinking. A range of brain regions are related to cognition, and these nonlanguage domains play a role in the ability of patients with chronic aphasia [29].

To emphasize the communication demands of an individual's life, the contextual domain is placed at the top of the cognitive communication competence model [7]. A Cochrane review concluded that communication activity in real-world settings, or functional communication, could improve communicative ability [1]. Communication practices should occur in an ecological context. Therefore, we designed 5 virtual scenario tasks focusing on everyday communication activities and pragmatic conversational skills. Patients may receive language practice opportunities in natural context communication settings in VR. These tasks are a comprehensive practice for the patients, and different short communication dialogues of everyday activities and cognitive exercises were integrated, including oral expression, naming, calculation, attention, memory, and reasoning tasks. VR technology is a promising rehabilitation tool and may be a useful alternative to conventional training [35], but the scene in VR is inaccurate. However, we could create almost everything in VR, and patients experience it as an actual situation, which makes it easy to try new therapeutic strategies. All learning in VR can be transferred to the real world [12]. With clinician shortages and a higher prevalence of aphasia, VR is not meant to replace skilled therapists but to ease clinicians' burdens.

Limitations

This study had some limitations. First, a range of brain regions are related to attention and executive functioning, and these nonlanguage domains play a role in the abilities of patients with chronic aphasia [29]. Owing to the limitations of intelligent VR, we cannot design some daily instructions (eg, washing clothes, cutting vegetables, cooking). VR could be used to provide an enriched environment where patients could master skills related to daily life that cannot be finished in the hospital. Learning in VR can then be transferred to real-life situations. Second, dialects vary across different areas of China. However, we did not involve dialects from different regions of China to provide options for those who have difficulty understanding and expressing Mandarin. There are nonlanguage abilities that would be worthwhile to consider in aphasia rehabilitation, such as prosody and emotional tone in utterances. However, it is difficult to implement these in the app. In addition, the usability of the app was assessed by only 17 individuals, which may have biased the usability results. In the future, we will design a more significant multicenter trial with a longer-term follow-up. Additionally, this app is meant to improve patients' oral expression abilities, so reading and writing are beyond our consideration. In particular, owing to the current limitations of technology, we cannot plug writing training into VR.

Conclusion

We conducted a Delphi study and developed a mobile-based VR app. This study constitutes a step toward the development of a combination of health and VR. We believe there is significant potential to make our app an in-home app in the future, contributing to the automation of rehabilitation administration. However, further studies are needed to evaluate the feasibility and efficacy of our app. We hope that our research provides guidelines and references for others in the medical field.

Acknowledgments

We gratefully thank the experts for completing the study surveys. This research was supported by the Fundamental Research Funds for the Central Universities of Central South University (project no. 2020zzts847) and the Hunan Provincial Health Commission (project nos. 2020SK51104 and 202114021494).

Authors' Contributions

All authors listed in this publication conducted the research. XB was responsible for all stages of the project. XB, PHFN, QC, YT, RF, QT, QC, SL, ASKC, and XL participated in the design, analysis, and writing of the article. PHFN and QC provided expertise in the development of the app. PHFN, QC, and ASKC revised the article. All authors approved the final draft.

Conflicts of Interest

None declared.

References

1. Brady M, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. Cochrane DB Syst Rev 2020 Jun 01:2016. [doi: 10.1002/14651858.cd000425.pub4]
2. Menon DK, Bryant C. Time for change in acquired brain injury. Lancet Neurol 2019 Jan;18(1):28. [doi: 10.1016/s1474-4422(18)30463-0]
3. Wu S, Wu B, Liu M, Chen Z, Wang W, Anderson CS, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. Lancet Neurol 2019 Apr;18(4):394-405. [doi: 10.1016/s1474-4422(18)30500-3]
4. Strong K, Mathers C, Bonita R. Preventing stroke: saving lives around the world. Lancet Neurol 2007 Feb;6(2):182-187. [doi: 10.1016/s1474-4422(07)70031-5]
5. Graham JR, Pereira S, Teasell R. Aphasia and return to work in younger stroke survivors. Aphasiology 2011 May 25;25(8):952-960. [doi: 10.1080/02687038.2011.563861]
6. Béjot Y, Bailly H, Durier J, Giroud M. Epidemiology of stroke in Europe and trends for the 21st century. Presse Med 2016 Dec;45(12 Pt 2):e391-e398. [doi: 10.1016/j.lpm.2016.10.003] [Medline: 27816343]
7. MacDonald S. Introducing the model of cognitive-communication competence: a model to guide evidence-based communication interventions after brain injury. Brain Inj 2017 Oct 24;31(13-14):1760-1780. [doi: 10.1080/02699052.2017.1379613] [Medline: 29064304]
8. Lavoie M, Macoir J, Bier N. Effectiveness of technologies in the treatment of post-stroke anomia: a systematic review. J Commun Disord 2017 Jan;65:43-53. [doi: 10.1016/j.jcomdis.2017.01.001] [Medline: 28171741]
9. Kurland J, Liu A, Stokes P. Effects of a tablet-based home practice program with telepractice on treatment outcomes in chronic aphasia. J Speech Lang Hear Res 2018 May 17;61(5):1140-1156. [doi: 10.1044/2018_jslhr-l-17-0277]
10. Lavoie M, Routhier S, Légaré A, Macoir J. Treatment of verb anomia in aphasia: efficacy of self-administered therapy using a smart tablet. Neurocase 2016 May 26;22(1):109-118. [doi: 10.1080/13554794.2015.1051055] [Medline: 26007615]
11. Vaezipour A, Campbell J, Theodoros D, Russell T. Mobile apps for speech-language therapy in adults with communication disorders: review of content and quality. JMIR Mhealth Uhealth 2020 Oct 29;8(10):e18858. [doi: 10.2196/18858]
12. 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 Mar 22;47(14):2393-2400. [doi: 10.1017/s003329171700040x]
13. Marshall J, Booth T, Devane N, Galliers J, Greenwood H, Hilari K, et al. Evaluating the benefits of aphasia intervention delivered in virtual reality: results of a quasi-randomised study. PLoS One 2016 Aug 12;11(8):e0160381 [FREE Full text] [doi: 10.1371/journal.pone.0160381] [Medline: 27518188]
14. Zhang Y, Chen P, Li X. Clinical research on therapeutic effect of virtual reality technology on Broca aphasia patients. 2009 Presented at: Proceedings of the International Conference on Information Technology; 2017; Ioannina, Epirus, Greece p. 134. [doi: 10.1109/incit.2017.8257880]
15. Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. Int J Nurs Stud 2001 Apr;38(2):195-200. [doi: 10.1016/s0020-7489(00)00044-4]
16. Powell C. The Delphi technique: myths and realities. J Adv Nurs 2003 Feb;41(4):376-382. [doi: 10.1046/j.1365-2648.2003.02537.x] [Medline: 12581103]

17. Mullen PM. Delphi: myths and reality. J Health Organ Manag 2003 Feb;17(1):37-52. [doi: 10.1108/14777260310469319] [Medline: 12800279]
18. Brain Health, for Life. URL: https://constanttherapyhealth.com/ [accessed 2022-04-01]
19. Therapy Menu. URL: http://www.aphasiatherapyonline.com/main.html [accessed 2022-04-01]
20. Reed GF, Lynn F, Meade BD. Use of coefficient of variation in assessing variability of quantitative assays. Clin Diagn Lab Immunol 2002 Nov;9(6):1235-1239 [FREE Full text] [doi: 10.1128/cdli.9.6.1235-1239.2002] [Medline: 12414755]
21. Chen Y, Li Z, Peng J, Shen L, Shi J. Development of a care bundle for stroke survivors with psychological symptoms: evidence summary and Delphi study. Evid Based Complement Alternat Med 2020 Jun 30;2020:7836024 [FREE Full text] [doi: 10.1155/2020/7836024] [Medline: 32714421]
22. Xing Q, Zhang M, Zhao F, Zhou Y, Mo Y, Yuan L. The development of a standardized framework for primary nurse specialists in diabetes care in China: a Delphi study. J Nurs Res 2019 Dec;27(6):e53. [doi: 10.1097/jnr.0000000000000330] [Medline: 31162295]
23. Ren H, Liu C, Wang R, Zhang M, Ma F, Li R, et al. Core competencies required for gastroenterology nursing specialists in China. Gastroenterol Nurs 2019;42(2):169-178. [doi: 10.1097/sga.0000000000000392]
24. Shen J, Xiang H, Luna J, Grishchenko A, Patterson J, Strouse RV, et al. Virtual reality-based executive function rehabilitation system for children with traumatic brain injury: design and usability study. JMIR Serious Games 2020 Aug 25;8(3):e16947 [FREE Full text] [doi: 10.2196/16947] [Medline: 32447275]
25. Ardila A, Rubio-Bruno S. Aphasia from the inside: the cognitive world of the aphasic patient. Appl Neuropsychol Adult 2018 May 23;25(5):434-440. [doi: 10.1080/23279095.2017.1323753] [Medline: 28535071]
26. Peach R. Cognitive approaches to aphasia treatment: application of the cognition of language to aphasia intervention. Semin Speech Lang 2017 Feb 15;38(1):3-4. [doi: 10.1055/s-0036-1597259] [Medline: 28201832]
27. Murray LL. Attention and other cognitive deficits in aphasia: presence and relation to language and communication measures. Am J Speech Lang Pathol 2012 May;21(2):S51-S64. [doi: 10.1044/1058-0360(2012/11-0067)]
28. Yao J, Liu X, Liu Q, Wang J, Ye N, Lu X, et al. Characteristics of non-linguistic cognitive impairment in post-stroke aphasia patients. Front Neurol 2020 Sep 30;11:1038 [FREE Full text] [doi: 10.3389/fneur.2020.01038] [Medline: 33117251]
29. Schumacher R, Halai AD, Lambon Ralph MA. Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. Brain 2019 Oct 01;142(10):3202-3216 [FREE Full text] [doi: 10.1093/brain/awz258] [Medline: 31504247]
30. El Hachioui H, Visch-Brink EG, Lingsma HF, van de Sandt-Koenderman MWME, Dippel DWJ, Koudstaal PJ, et al. Nonlinguistic cognitive impairment in poststroke aphasia: a prospective study. Neurorehabil Neural Repair 2014 Nov 08;28(3):273-281. [doi: 10.1177/1545968313508467] [Medline: 24213963]
31. Brownsett SLE, Warren JE, Geranmayeh F, Woodhead Z, Leech R, Wise RJS. Cognitive control and its impact on recovery from aphasic stroke. Brain 2014 Jan;137(Pt 1):242-254 [FREE Full text] [doi: 10.1093/brain/awt289] [Medline: 24163248]
32. Hagoort P. The core and beyond in the language-ready brain. Neurosci Biobehav Rev 2017 Oct;81(Pt B):194-204. [doi: 10.1016/j.neubiorev.2017.01.048] [Medline: 28193452]
33. Gilmore N, Meier EL, Johnson JP, Kiran S. Nonlinguistic cognitive factors predict treatment-induced recovery in chronic poststroke aphasia. Arch Phys Med Rehabil 2019 Jul;100(7):1251-1258 [FREE Full text] [doi: 10.1016/j.apmr.2018.12.024] [Medline: 30639272]
34. Zakariás L, Keresztes A, Marton K, Wartenburger I. Positive effects of a computerised working memory and executive function training on sentence comprehension in aphasia. Neuropsychol Rehabil 2018 Apr 21;28(3):369-386. [doi: 10.1080/09602011.2016.1159579] [Medline: 26999324]
35. Dockx K, Bekkers E, Van DBV. Virtual reality for rehabilitation in Parkinson's disease. Cochrane DB Syst Rev 2016:CD010760. [doi: 10.1002/14651858.cd010760]

Abbreviations

ABI: acquired brain injury
AI: artificial intelligence
CV: coefficient of variation
VR: virtual reality

Edited by N Zary; submitted 23.05.21; peer-reviewed by T Ong, G Cárdenas-López; comments to author 03.12.21; revised version received 28.02.22; accepted 05.03.22; published 07.04.22.

Please cite as:
Bu X, Ng PHF, Tong Y, Chen PQ, Fan R, Tang Q, Cheng Q, Li S, Cheng ASK, Liu X
A Mobile-based Virtual Reality Speech Rehabilitation App for Patients With Aphasia After Stroke: Development and Pilot Usability Study
JMIR Serious Games 2022;10(2):e30196
URL: https://games.jmir.org/2022/2/e30196
doi:10.2196/30196
PMID:35389349

Original Paper

Augmented Reality-Based Surgery on the Human Cadaver Using a New Generation of Optical Head-Mounted Displays: Development and Feasibility Study

Behrus Puladi[1,2], MD; Mark Ooms[1], MD; Martin Bellgardt[3], MSc; Mark Cesov[1,3], MSc; Myriam Lipprandt[2], PhD; Stefan Raith[1], PhD; Florian Peters[1], DMD; Stephan Christian Möhlhenrich[1,4], DMD, PhD; Andreas Prescher[5], MD, PhD; Frank Hölzle[1], MD, DMD, PhD; Torsten Wolfgang Kuhlen[3], PhD; Ali Modabber[1], MD, DMD, PhD

[1]Department of Oral and Maxillofacial Surgery, University Hospital RWTH Aachen, Aachen, Germany
[2]Institute of Medical Informatics, University Hospital RWTH Aachen, Aachen, Germany
[3]Visual Computing Institute, RWTH Aachen University, Aachen, Germany
[4]Department of Orthodontics, Private University of Witten/Herdecke, Witten, Germany
[5]Institute of Molecular and Cellular Anatomy, University Hospital RWTH Aachen, Aachen, Germany

Corresponding Author:
Behrus Puladi, MD
Department of Oral and Maxillofacial Surgery
University Hospital RWTH Aachen
Pauwelsstraße 30
Aachen, 52074
Germany
Phone: 49 241 80 88231
Fax: 49 241 82430
Email: bpuladi@ukaachen.de

Abstract

Background: Although nearly one-third of the world's disease burden requires surgical care, only a small proportion of digital health applications are directly used in the surgical field. In the coming decades, the application of augmented reality (AR) with a new generation of optical-see-through head-mounted displays (OST-HMDs) like the HoloLens (Microsoft Corp) has the potential to bring digital health into the surgical field. However, for the application to be performed on a living person, proof of performance must first be provided due to regulatory requirements. In this regard, cadaver studies could provide initial evidence.

Objective: The goal of the research was to develop an open-source system for AR-based surgery on human cadavers using freely available technologies.

Methods: We tested our system using an easy-to-understand scenario in which fractured zygomatic arches of the face had to be repositioned with visual and auditory feedback to the investigators using a HoloLens. Results were verified with postoperative imaging and assessed in a blinded fashion by 2 investigators. The developed system and scenario were qualitatively evaluated by consensus interview and individual questionnaires.

Results: The development and implementation of our system was feasible and could be realized in the course of a cadaver study. The AR system was found helpful by the investigators for spatial perception in addition to the combination of visual as well as auditory feedback. The surgical end point could be determined metrically as well as by assessment.

Conclusions: The development and application of an AR-based surgical system using freely available technologies to perform OST-HMD–guided surgical procedures in cadavers is feasible. Cadaver studies are suitable for OST-HMD–guided interventions to measure a surgical end point and provide an initial data foundation for future clinical trials. The availability of free systems for researchers could be helpful for a possible translation process from digital health to AR-based surgery using OST-HMDs in the operating theater via cadaver studies.

(JMIR Serious Games 2022;10(2):e34781) doi:10.2196/34781

KEYWORDS

digital health in surgery; surgical technique; surgical training; computer-assisted surgery; optical see-through head-mounted display; HoloLens; surgical navigation; medical regulation; open-source; AR; augmented reality; surgery; surgeon; cadaver; serious game; head-mounted display

Introduction

Health care is increasingly supported by digital technologies [1]. Almost one-third of the world's disease burden requires surgical intervention [2], yet only a small fraction of the potential applications of digital health is used in the surgical domain [1]. Current digital health applications such as artificial intelligence (AI)-based predictive models, the use of telemedicine, and wearables do not touch the core of surgical activity in the operating theater [3]. Assistance systems based on augmented reality (AR) or robotics use, on the other hand, would allow the surgeon's core activities to benefit from digital health in the coming decades [4]. However, the use of autonomous robots in surgery is ambitious considering surgeons take many years to become trained and surgical interventions can often be very situation-specific. In this respect, unlike robotics, AR as an assistance system for the surgeon supports spatial perception and simultaneously incorporates the surgeon's experience. Thereby, AR-based surgery could be a near future and feasible step toward digital health in the operating theater [5].

Further technical development has made optical-see through head-mounted displays (OST-HMDs) such as the HoloLens (Microsoft Corp) commercially viable with broad use in the industry [6]. Health care and surgery, in particular, are not primarily affected by this development, among other things due to the high regulatory requirements for medical devices. Any researcher in medicine can quickly develop AI-based models with a few lines of script code based on public data and provide proof of performance. However, in surgery with next-generation technologies like AR with OST-HMDs, this development is not yet foreseeable [7].

By feeding back relevant information to the surgeon during surgical tasks based on preoperative or intraoperative medical imaging data or AI-based guided models, AR with or without image-guided surgery (IGS) could overcome one of the main problems of surgical procedures, which is that they mainly rely on the surgeon's spatial awareness or haptic perception in the surgical field [4,5]. AR itself augments the otherwise real environment with virtual objects, is located in a reality-virtuality continuum, and includes a wide range of technologies [8]. Beside visual perception, AR can also refer to one or multiple combined modalities of perception, such as auditory or haptic [9].

AR applications have been used since the mid-1990s, mainly for surgical procedures on rigid tissue in the head and neck region. Examples of applications in the operating theater are orthognathic surgery, oncology including parotid surgery, and traumatology. Anatomical and pathological structures, drilling and implant position, resection margins, and reconstructive planning are visualized using different AR technologies [10,11]. Similar examples can also be found on cadavers [12-14].

Due to the underlying technology with external monitors, however, many of these deployed systems result in a dissociation between the perceptual site and the operational field [5,10]. HMDs, on the other hand, enable an egocentric view [15] with virtual objects directly displayed in the surgical field of view [10]. Even though HMDs were first described in the 1960s [16], the capabilities of the various HMDs used intraoperatively still vary widely [17]. Basically, 2 classes of AR HMDs can be distinguished [15], optical see-through and video see-through HMDs, the former having the advantage of an unobstructed view of the surgical field [17].

On the road to widespread use of this rapidly developing technology, proof of performance is essential, especially in the regulatory context. Cadaveric studies have long provided a contribution to demonstrating the performance of new medical technologies and are considered a prestudy proof of performance prior to clinical trials [18]. However, cadaveric studies are rare when using OST-HMDs [19-22], and it is still unclear whether cadavers are generally suitable for testing surgical applications with OST-HMDs.

In order to enhance the development of digital health in surgery, we aimed to develop an AR- and OST-HMD–based system for a cadaveric study using free technologies to make it available and adoptable for research in various experimental surgical scenarios as proof of performance. Furthermore, we wanted to investigate if cadaver studies using this system would be suitable for testing system feasibility, applicability in a surgical task that relies primarily on spatial and haptic perception, and evaluability of its surgical end point.

We chose a simple and understandable AR scenario on fresh cadaver heads using a HoloLens, where the surgeon had to reduce a fractured zygomatic arch, a common injury of the human face.

Methods

System Development

Concept and Requirements

The purpose of our study was to develop an AR-based system for IGS to be used in a surgical environment with human cadavers. The aim was to augment the surgeon's spatial perception with 3D models based on previous medical imaging by overlaying them on the surgical field using AR. This overlay is intended to be adaptive by adjusting to the current position of the cadaver and surgical instruments and to allow interaction between both. Three essential feedback functions should be provided here: feedback of the proximity of the surgical instrument to surgical target structures by means of a visual signal, an auditory signal, and a visual representation of the movement of the surgical instrument. The graphical user interface should allow intuitive selection of the different cases with specific models and different functions by gestures via

AR-based buttons. Furthermore, all described functionalities should also be usable with voice commands to enable hands-free working.

Overall, the system should be easily adaptable to different surgical scenarios, cost-efficient, and easily replicable by third parties, allowing it to be universally applicable as proof of performance for OST-HMDs in surgery on cadavers.

Implementation

We aimed to achieve our requirements by using the commercially available HoloLens 1 as one of the state-of-the-art and most broadly used OST-HMDs combined with the camera-based tracking system Vuforia (version 8.5.8, PTC Inc). Our software prototype was developed using the C# programming language with the popular game engine Unity 2018.4.13f LTS (Unity Technologies) and the Mixed Reality Toolkit (version 2.3.0, Microsoft Corp) for rapid prototyping. Our software prototype was then developed into a prerelease candidate of an open-source software as part of a master's thesis in computer science [23].

Based on medical imaging data, 3D models were created for the cadavers and surgical instruments (Figure 1a). In order to attach the mounts for image tracking, the cadavers were prepared beforehand to obtain a definite reference point (Figure 1b and Figure 2a). Subsequently, mounts were designed in Autodesk Inventor Professional 2020 (Autodesk Inc), 3D printed using a Fortus 450mc (Stratasys), and attached to the cadaver heads and surgical instruments (Figure 1b and Figure 2b). The image target was used for tracking the cadaver heads (Figure 1c) and for the half-cube for holographic verification described below (Figure 2c). The Vuforia multitarget (corresponds to a combination of image targets so that the surgical instrument can be tracked from both sides) was used for tracking the surgical instrument; in our scenario, a Stromeyer hook (Figure 1c and Figure 2d).

Figure 1. System development: (a) Creation of virtual 3D representations (in purple) of the cadaver and surgical instruments based on computed tomography images. (b) 3D printing of mounts with image targets for attachment to the cadaver and the surgical instruments for camera-based tracking. (c) Superimposition of the virtual 3D models (purple) and real-world object (gray) resulting in an augmented reality (AR) object (cyan). (d) Performing AR-based surgery with an optical see-through head-mounted display. Possibility of interaction between surgical target structures and instruments by means of visual and auditory feedback. Software can be controlled via gestures using an AR-based graphical user interface. DICOM: Digital Imaging and Communications in Medicine; CT: computed tomography; OST-HMD: optical see-through head-mounted display; AR: augmented reality.

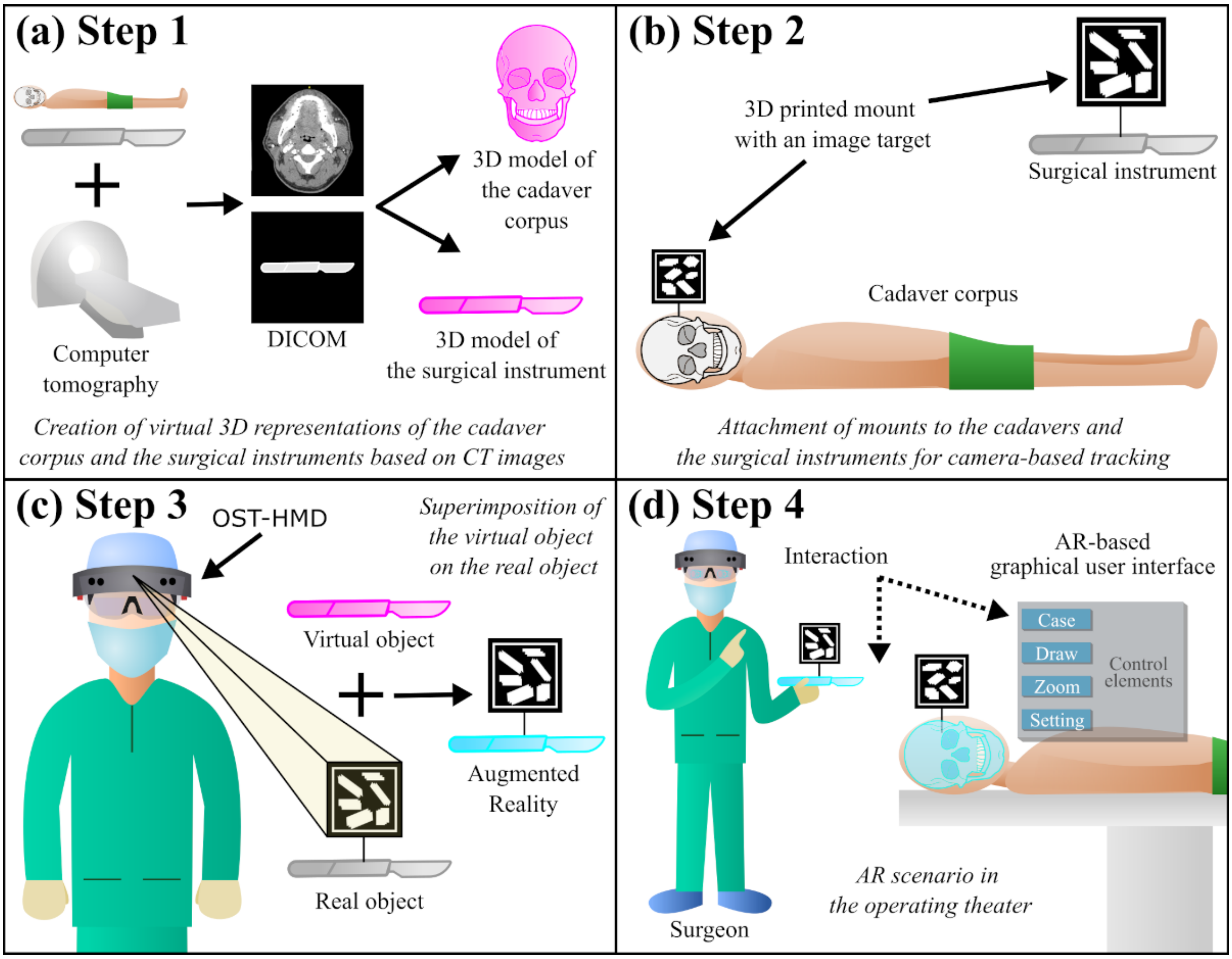
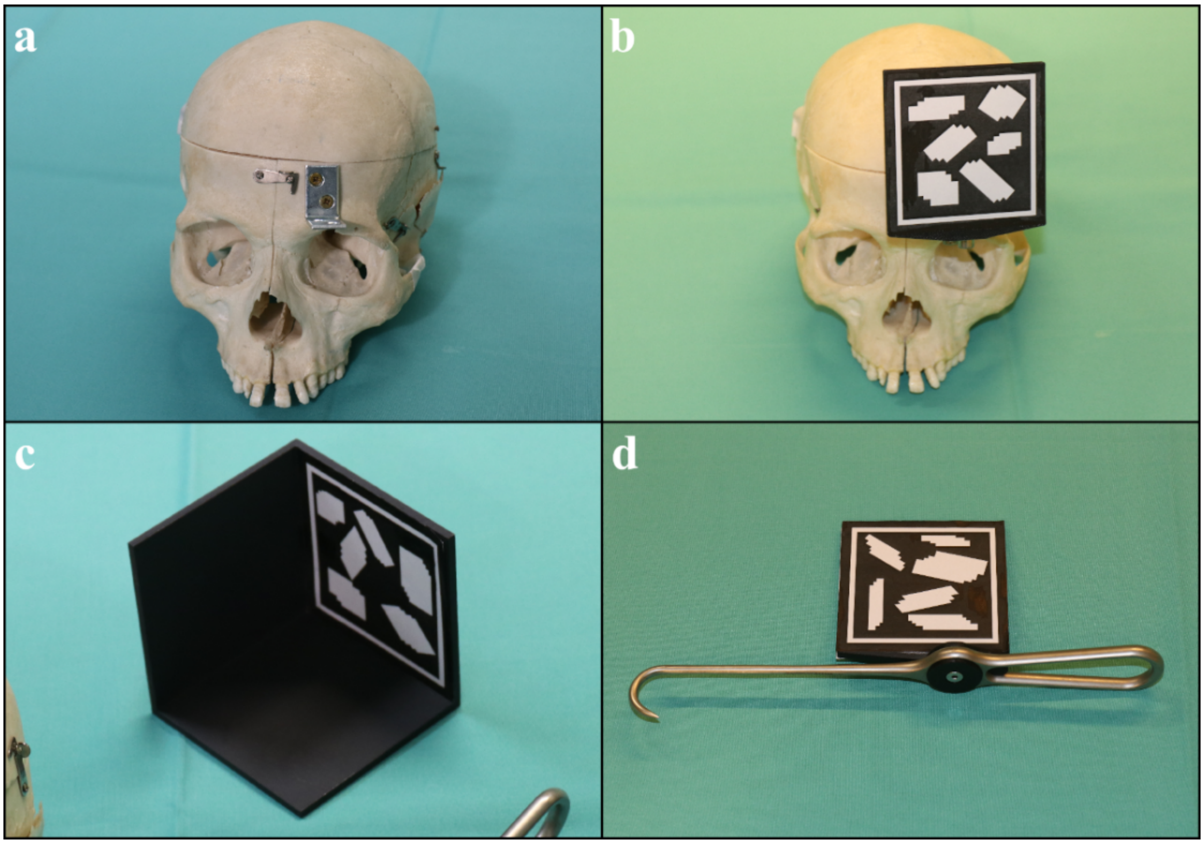


Figure 2. Technical setting: (a) To ensure a consistent method for tracking, a metal angle was attached to the bone of the forehead of each cadaver to attach the tracking mount as not every cadaver head had proper dentition for a stable splint-based tracking. (b) An image target on the mount connected to the forehead via the metal angle. (c) A half-cube for holographic verification can be used for testing the superimposition between real and virtual objects to represent possible errors in the fit of an optical see-through head-mounted display (OST-HMD) or errors in tracking by user-verifiable reference surfaces. (d) The surgical instrument (Stromeyer hook, in our scenario) with an attached tracking mount.



A graphical user interface was developed for the AR software to make the virtual anatomical models of the respective cadavers selectable via AR-based buttons and to adjust the software (Figure 1d and Figure 3). It had 3 main functionalities: visual feedback in the region of interest by a color transition of the models from green to red when the tip of the virtual surgical instrument touches the virtual cadaver model (Figure 4c and 4d), auditory feedback from an acoustic tone whose pitch was modulated depending on the distance between the tip of the virtual instrument and the virtual model (Figure 3a and 3b), and visual feedback through virtual drawing (Figure 3d). In our scenario, it was possible to trace the inner contour of the zygomatic arch with the tip of the Stromeyer hook and then visualize it within a bounding box at different sizes and from different directions to evaluate the shape of the inner zygomatic arch contour. All functionalities described were also usable with voice.

When a virtual model was selected, it was superimposed on the real cadaver head by continuous tracking (Figure 1d, Figure 4c, and Figure 4d). The surgical instrument was tracked throughout. To evaluate the perceived superimposition between virtual and real surgical instruments, a half-cube was printed with distinct reference surfaces.

Figure 3. Graphical user interface: (a) Display of the graphical user interface in the Unity development environment. Not all functions mentioned were used in our scenario. The “Recalibrate” button can be used to align the virtual representation of the surgical instrument with the virtual representation of the half-cube. For this, the real surgical instrument (Stromeyer hook) must be exactly aligned with the surfaces of the real half-cube for the holographic verification. Acoustic feedback can be deactivated via the “Sound: Off” button. Additionally, the slider provides an adjustment of the sound functionality depending on the distance [d] and the factor [a] with the formula [d^a]. To visualize the movement of the surgical instrument, the “Draw” function can be used to display the trajectory by a 3D line. With the button “Enlarge drawing,” the drawing can be zoomed in and with the button “Delete drawing,” the drawing function can be reset. A selection of radio buttons to choose the appropriate cadaver case. (b) Additionally, the virtual and real cadaver head could be adjusted by hand movement, and the sensitivity of the adjustment could be controlled by a slider. This function was not used. (c) Representation of the calibration function with the half-cube in blue, the surgical instrument in white, and the holder for the image target in green. (d) Illustration of the visualization of the instrument trajectory.

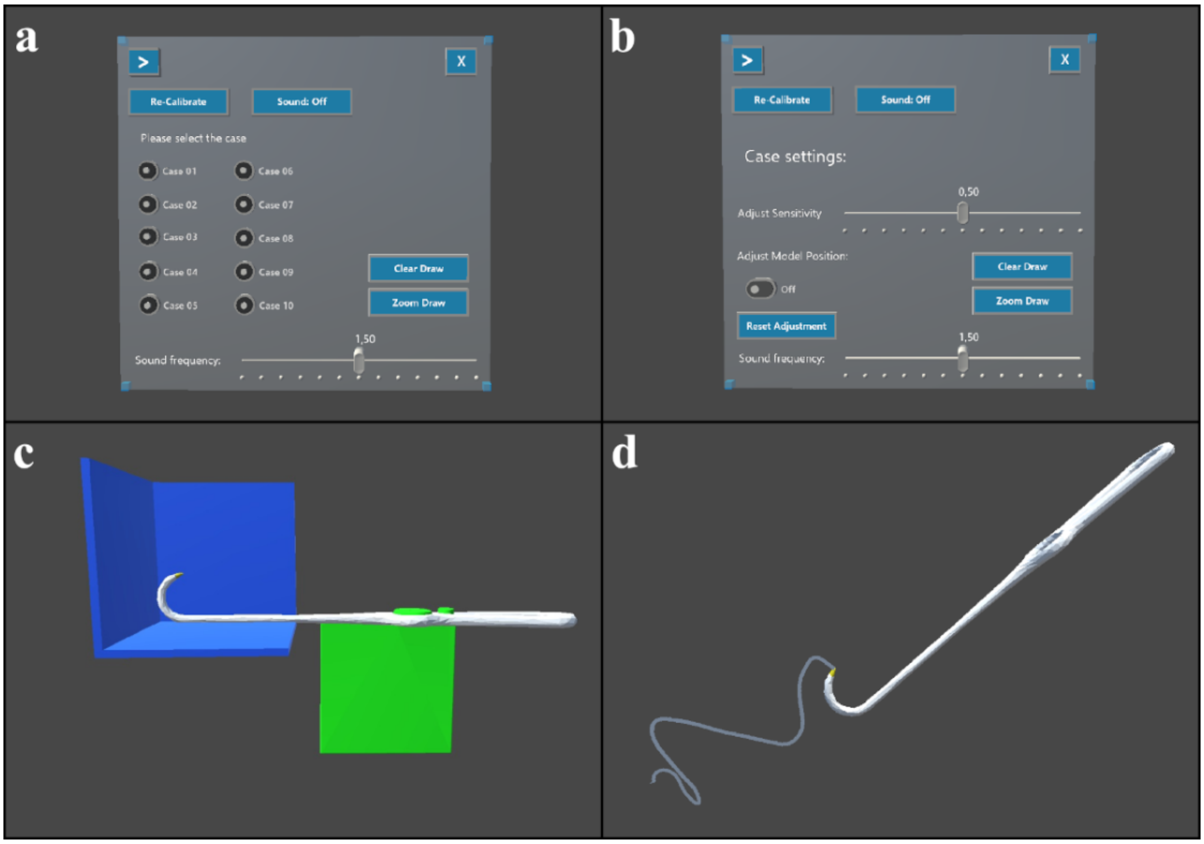
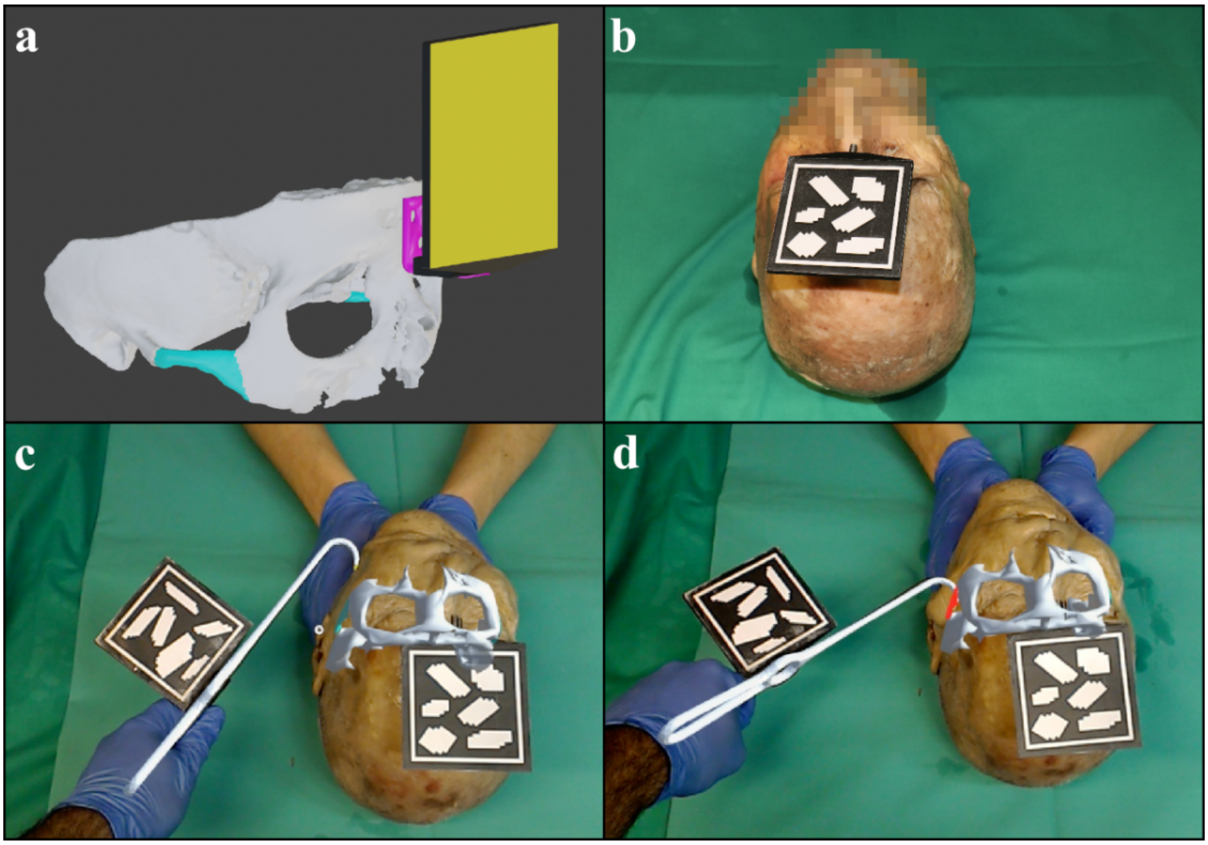


Figure 4. Cadaver trial: (a) A 3D model of the facial skull (white) of one of the cadaver cases with a color representation of the unfractured zygomatic arch (cyan), fixed metal angle (purple), mounting (black), and image target (yellow). (b) A fresh cadaver head shows the placed mount and image target for navigation. (c-d) Photograph taken through HoloLens as one of the investigators performs the cadaver trial. (c) The cadaver head is overlaid with the virtual bone model. The zygomatic arch is shown in green because the tip of the virtual Stromeyer hook has not yet collided with the intended position of the nonfractured zygomatic arch model. The Stromeyer hook is superimposed with an accurate virtual model of itself. (d) The tip of the virtual Stromeyer hook now touches the model of the nonfractured zygomatic arch, resulting in a color change of the zygomatic arch model to red.



Ethics Approval

This article does not include studies with live human participants or animals. Approval by the ethics committee of the University Hospital RWTH Aachen (approval number EK 348/21) has been granted. The investigators agreed to participate in the study.

Cadaveric Trial

Preparation

Ten fresh cadaver heads were randomly selected. The initial condition of the facial skeleton was first scanned with cone beam computed tomography (CBCT, Dentsply Sirona). To ensure a consistent method, a metal angle was attached to the bone of the forehead of each cadaver to attach the tracking mount (Figure 2a and 2b) as not every cadaver head had proper dentition for stable splint-based tracking. Subsequently, all zygomatic arches were randomly fractured by a direct blow with a surgical hammer. The fractured state was then scanned again with CBCT. Thereafter, all cadaver heads were frozen until study examination. The resulting fractured zygomatic arches had 1 to 5 fragments. A total of 16 zygomatic arch fractures were classified as type II, 3 as type III, and 1 as type IV, according to Yamomoto et al [24] (Multimedia Appendix 1).

Based on the acquired medical imaging data, 3D models were created for all cadaver heads with the initial situation and fractured zygomatic arches. Both models were registered using the best fit alignment feature of Geomagic Studio 2013 (3D Systems Inc) and loaded into our software. In addition, a Stromeyer hook, a surgical instrument routinely used to reduce zygomatic arch fractures, was digitized and loaded into our software. For tracking, mounts were then 3D printed for fixation on the cadaver heads and on the Stromeyer hook (Figure 2d).

Trial

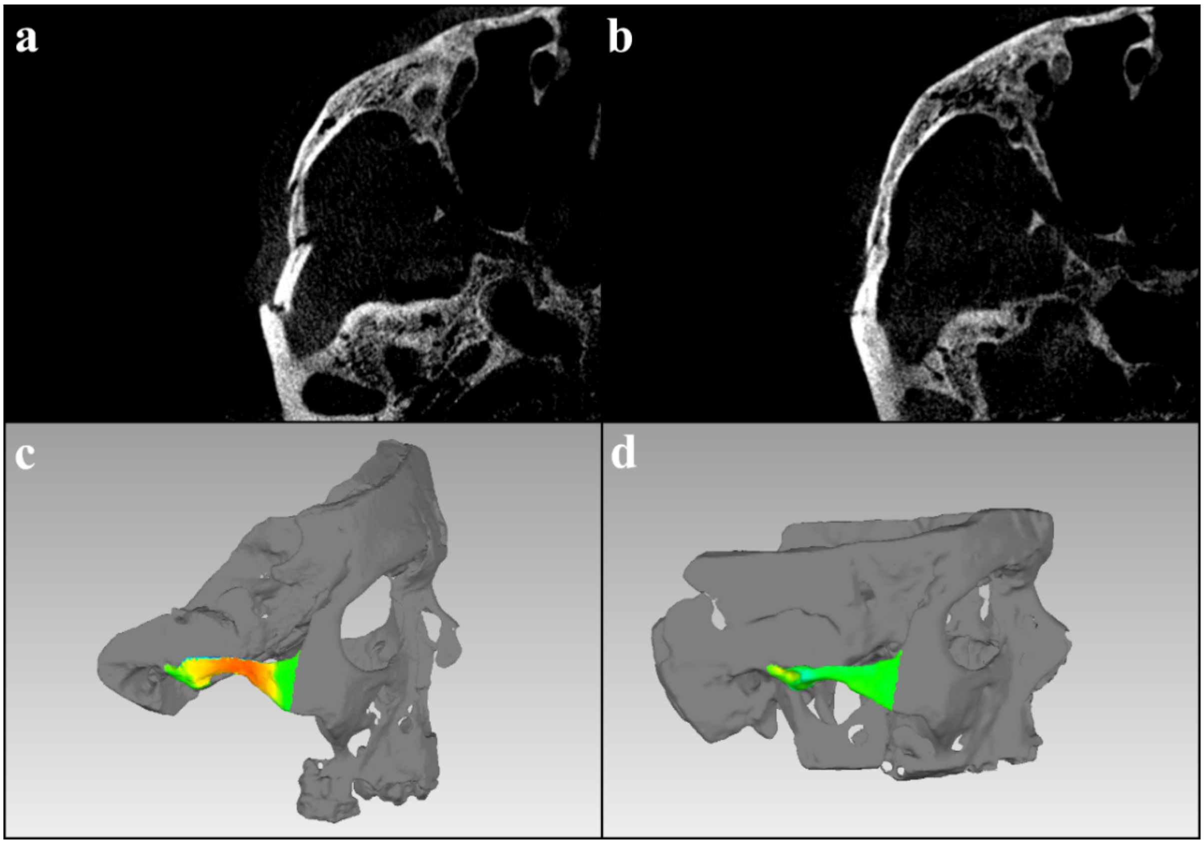
The cadaver heads were randomly assigned to the investigators (a resident and a senior surgeon). One zygomatic arch side of each head was randomly selected (based on a random number generator) for reduction by the conventional method and the opposite side by the AR-based method. Before reduction, the investigators were able to view the CBCT imaging data with the fractured situation on a computer. The conventional reduction was performed with the Stromeyer hook through a percutaneous incision and was based only on haptic perception. The AR-based reduction was performed identically, with the addition of a registered virtual model of the Stromeyer hook and a registered virtual model of the intact zygomatic arch of the corresponding cadaver displayed on the HoloLens. The aforementioned functionalities of feedback through color transition, virtual drawing, and audio signals provided the investigators with additional visual and auditory perception

(Figure 4c and 4d). For both methods, the time between the percutaneous incision and performed reduction was measured. After complete reduction, the corresponding cadaver head was scanned with CBCT.

Evaluation

Based on postoperative imaging, 3D models of the cadaver heads were created and registered with the corresponding preoperative and initial situation using Geomagic Studio 2013. The zygomatic arch was defined as the region from the temporal origin of the zygomatic process to a straight vertical extension line at the posterior margin of the frontosphenoid process of the zygoma and converted to separate models. The deviation of the different models was then compared (settings: maximum deviation 10 mm, critical angle 45.0°; display resolution set to fine). The initial nonfractured model was used as a reference and compared to the fractured model and subsequently to the reduced model (Figure 5). The results obtained were exported for statistical analysis.

Figure 5. Evaluation: (a) A fractured zygomatic arch visualized before reduction and (b) after reduction in axial cone beam computed tomography slices. (c) The deviation of a fractured zygomatic arch is displayed in color in Geomagic Studio 2013 (3D Systems Inc). Red is for severe deviation (≥1 mm) and green for minor deviation (<1 mm). (d) The same case in Geomagic Studio 2013 after reduction.



Finally, reduction quality was classified into 4 levels based on postoperative imaging by 2 noninvestigators in a consensus and blinded fashion according to Yakomoto et al [24]: poor for reduction without improvement in bone fragment shape and continuity, fair for incomplete restoration but an improvement in bone fragment shape and continuity, good for near-complete restoration of shape with and without continuity of bone fragments, and excellent for complete restoration of shape with continuity of bone fragments.

The AR software was assessed using the System Usability Scale (SUS) [25]. Afterward, a consensus interview with open-ended questions was conducted with both investigators, and the AR-based scenario was qualitatively assessed using an individual questionnaire (Multimedia Appendix 2) on a 5-point Likert scale (1=strongly disagree; 5=strongly agree).

Statistical Analysis

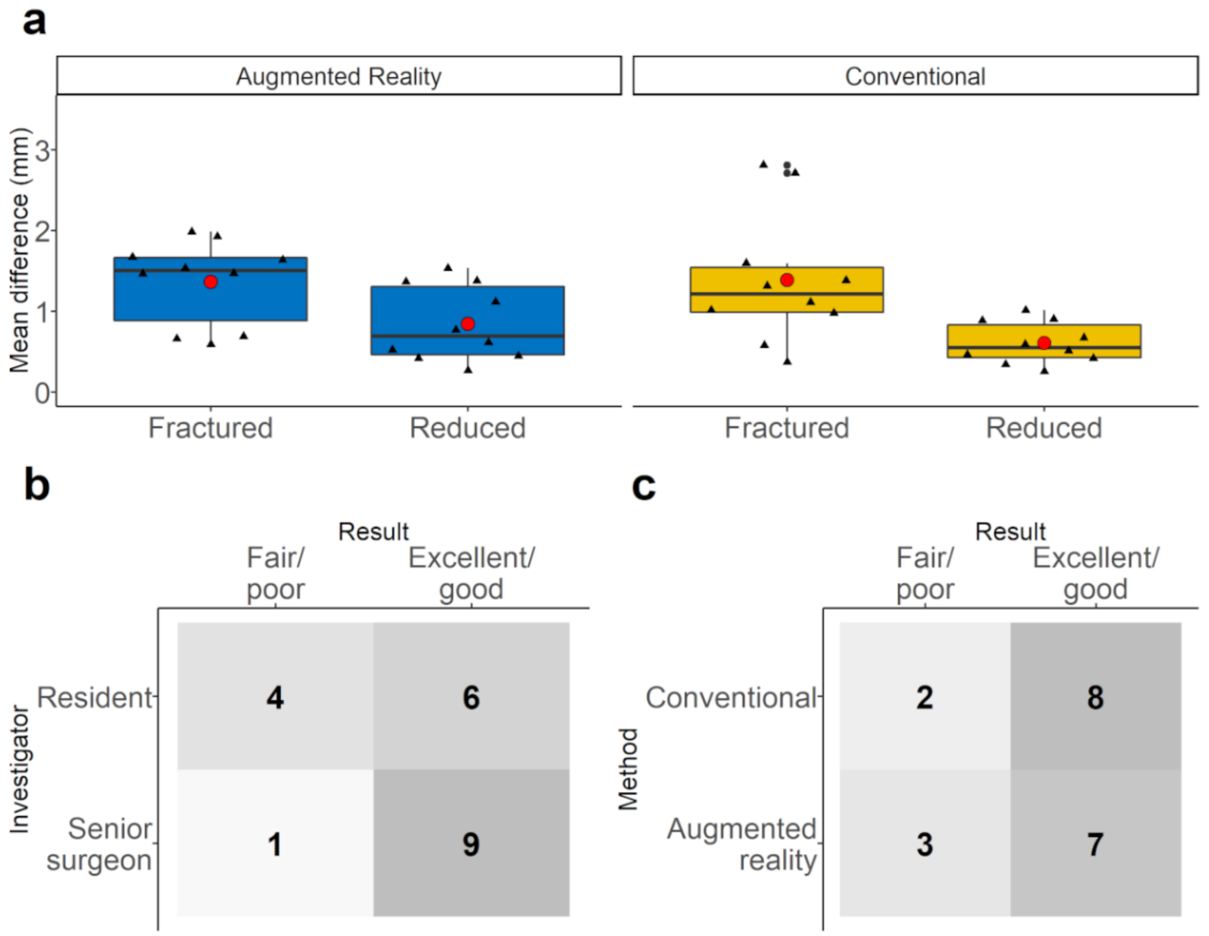
The R programming language (R Foundation for Statistical Computing) was used for statistical analysis. Results were expressed as mean and standard deviation. The 95% confidence intervals were calculated by bootstrapping with 1000 replications [26].

Results

Surgical Outcome

Within the quantitative reduction measurement between fractured and reduced zygomatic arches, our test scenario showed a mean reduction of 0.78 mm (95% CI 0.37-1.29 mm) for the conventional method and 0.52 mm (95% CI 0.23-0.77 mm) for the AR-based method (Figure 6a). The mean time to perform zygomatic arch reduction using the conventional method was 84 seconds (95% CI 52-116 s) and for the AR-based method was 115 seconds (95% CI 54-198 s). A distinct difference in zygomatic arch reduction was observed between the resident and the senior surgeon. Of the 10 zygomatic arch reductions performed by the senior surgeon, 9 were rated good or excellent, while 6 of 10 performed by the resident were rated good or excellent (Figure 6b). This distinct difference was not present between the conventional and AR method, with 8 of 10 zygomatic arch repositions rated good or better for the conventional method and 7 of 10 for the AR method (Figure 6c).

Figure 6. Results: (a) The absolute surface deviation of the fractured and reduced model was calculated in comparison to the nonfractured model and presented as a boxplot before and after reduction for the augmented reality–based method (in blue) and the conventional method (in yellow). Black triangles represent individual measured values. The large red dot represents the mean value and black dots represent outliers. (b,c) Results of zygomatic arch repositioning were determined by 2 investigators in a blinded fashion (for the method) and by consensus. Displayed as a 4-panel chart. Excellent/good was rated as an adequate and fair/poor as an inadequate surgical outcome. (b) Comparison of the resident with the senior surgeon. (c) Comparison based on the method used.



Evaluation by Investigators

Based on consensus interviews with both investigators, the vertical field of view (FOV) was considered small and tracking mounts could interfere in certain surgical scenarios. When using the HoloLens 1, it was noted that an incorrect fit on the head could also lead to an error in superimposition between virtual and real objects. In this context, the simple half-cube for holographic verification was perceived as helpful for evaluation. Visualization of the fractured condition was preferred over the nonfractured one for navigation.

In addition, an individual Likert questionnaire was performed (Table 1). Both investigators agreed that the holographic visualization of the skeleton by means of an OST-HMD was helpful for spatial perception (mean 4.5) and that it appeared as an integrated part of the fresh cadaver head (mean 4.0). They disagreed that the attached mount for tracking the surgical instrument was perceived as disturbing in that scenario (mean 2.5). Both disagreed with the statement that they felt insecure using the AR-based method (mean 1.5) and agreed that they felt confident using the AR-based method when reducing the zygomatic arch (mean 4.5). They also expressed a preference to use the AR-based method on real patients (mean 4.5) and strongly agreed that they found the AR-based method helpful in the field of haptic surgery (mean 5.0). The average SUS for the AR application was 90 and can thus be rated as best imaginable.

Table 1. Questionnaire results[a].

Item	Resident	Senior surgeon	Mean
1. I found the holographic visualization of the zygomatic arch by means of OST-HMD[b] helpful for my spatial perception.	4	5	4.5
2. I felt the holographic representation of the zygomatic arch was an integrated part of the cadaver head.	4	4	4.0
3. I found the visual feedback from the color change during the zygomatic arch reduction helpful.	5	4	4.5
4. I found the auditory feedback by changing the tone amplitude during the zygomatic arch reduction helpful.	4	5	4.5
5. I found the drawing function helpful for the visual representation of bone contours.	4	4	4.5
6. I found the navigation holder for the surgical instrument disturbing.	2	3	2.5
7. I think the AR[c]-based method is helpful in haptic surgery.	5	5	5.0
8. I felt more confident in the zygomatic arch reduction using the AR-based method.	5	4	4.5
9. I have felt insecure about the zygomatic arch reduction due to the AR-based method.	1	2	1.5
10. I would like to use the AR-based method on real patients.	5	4	4.5

[a]1=strongly disagree; 5=strongly agree.

[b]OST-HMD: optical see-through head-mounted display.

[c]AR: augmented reality.

Discussion

Principal Findings

In our study, we demonstrated that it is possible to develop an adaptable and usable AR-based system with OST-HMDs and image-guided capacities for surgical interventions by combining freely available technologies and evaluating them in a test scenario on human cadavers. This system can be adopted by researchers worldwide and adapted to their own surgical scenarios. The implementation will require a HoloLens or Unity-compatible OST-HMD, the ability to capture 3D medical imaging data, and a 3D printer to produce suitable mounts for the cadaver and surgical equipment. The software and models of the 3D mounts are freely available under an open-source license.

However, the presented system still has shortcomings. When using HoloLens, we have found that an incorrect fit on the head leads to a positioning error between the eyes and the semitransparent display, causing perceptual errors. Only with the correct fit were both real and virtual objects correctly superimposed. For this reason, the simple half-cube we have developed for holographic verification can be used for evaluating the superimposition between real and virtual objects to represent possible errors in the fit of an OST-HMD or, additionally, errors in tracking by user-verifiable reference surfaces (Figure 2c and Figure 3c). The FOV of the HoloLens 1 with 34° was perceived as relatively low [27], although it was judged to be sufficient for our procedure. Depending on the surgical procedure, it could also lead to poor ergonomics and potentially affect the success of the surgical intervention. An enlarged FOV like on the HoloLens 2 [27] could possibly alleviate this.

Although the tracking mount was not found to be a disturbance by the investigators, it could become a potential concern during surgical procedures where space is limited or could lead to tracking errors because the tracking mount is obscured. One solution would be a mount-free or electronic tracking method to avoid disturbing surgeons in such situations [28].

Previous studies that evaluated image-based tracking using Vuforia and a HoloLens indicated a position error of 1.74 to 1.94 mm [29,30]. Our visual evaluation using a printed half-cube conformed to the range of the aforementioned studies (~2 mm). We did not perform a reexamination because we used the same system as the studies mentioned [29,30]. Overall, we considered this sufficient for a majority of surgical scenarios for the first proof of performance. However, the tracking was slow with quick instrument or head movements. This could be improved by increased hardware performance provided by the HoloLens 2 or by holographic remoting, where the main computational load is carried out on an external computer [31]. During the development of our system, we noticed that tracking with Vuforia is faster when the image targets have a black background, which further addressed performance limits [23]. Another alternative would be markerless registration, which has shown an average positioning error between 3.3 to 9.3 mm, depending on the spatial direction. In the future, this error might be reduced with more powerful hardware and could be a serious alternative, especially since markerless registration has no potentially disturbing markers in the surgical field [32]. Manual registration, which showed a mean error of alignment of 12.4 mm, would be another option. After appropriate training or assistance by fiducial markers, the error was reduced to 10 mm [11,33]. Consequently, image-based registration with Vuforia, which is much more accurate, is still the method of choice for most applications [29,30].

In our scenario, it was possible to visualize both the fractured and nonfractured situation, as the healthy bone condition can often be reconstructed with little effort by mirroring the nonfractured side, especially for the face [34]. Since individual bone fragments could not be tracked with our method, we presented the nonfractured situation as a guide for bone reposition. During the consensus interview, this was perceived as a disadvantage for conducting the reduction since the ideal situation can be easily imagined by the investigators themselves. Consequently, the fractured situation and, if necessary, an additional nonfractured situation should be offered for visualization in future examinations. The visualization of the bone with shaded 3D models was perceived as an integral part of the cadaver head. However, further research should focus on whether the use of different display methods such as points, lines, contours, planes, surfaces, wireframes, meshes, and volumes offer advantages in surgical procedures with AR [5].

Our system provided visual and auditory feedback depending on the distance of the surgical instrument working area and intended reduction situation of the bone. This was realized visually via a color change of the virtual zygomatic arch model as well as via the possibility of graphical representation of the movement path of the surgical instrument tip. The drawing function can be used to show the internal contour of bones—in our case, the contour of the fractured bone. If a fracture offset is present, it would be represented by an offset of the drawing line (Figure 3d). To our knowledge, this is the first time that an intraoperative drawing function has been applied in surgery with an OST-HMD. The acoustic feedback operated by increasing the amplitude as a function of the distance between the working area of the surgical instrument and surgical target structure. The combination of visual and auditory feedback was found to be helpful by the investigators. This is consistent with the observation in an AR-based model scenario for a needle biopsy performed by surgeons, where the combination of visual and auditory feedback significantly reduced localization error and increased the success rate [35].

It has already been shown that studies with OST-HMDs on cadavers are suitable to measure the difference between the planning of drill holes, placement of screws, or performance of osteotomies and the actual performance [19-22]. Our study was also able to demonstrate that cadaveric studies with OST-HMDs are suitable to determine fracture reductions quantitatively and qualitatively and thus in one of the veritable surgical end points. In this regard, expected differences between a resident and a senior surgeon were observed. The advantage of using fresh cadavers is the presence of realistic and complex anatomical conditions and thus a situation analogous to the living patient without taking possible surgical risks.

Studies on fresh cadavers, however, cannot determine clinical outcomes such as pain, patient-guided range of motion, dysfunction, or other clinical parameters. Nevertheless, cadaveric studies can be used to provide a data basis for subsequent clinical study planning. The technical system can be evaluated and tested. The developed AR system did not result in a large temporal difference from the conventional method in our scenario. It is important to measure duration as an end point, as surgery time is an important quality indicator. Prolonged surgery durations lead to a greater number of complications for the patient [36] and increased costs for the health care system [37]. Furthermore, the quantitative (reduction in mm) and qualitative (assessed reduction quality) data obtained can be used to plan the sample size for larger cadaveric studies or clinical trials. Procedures where the number of subjects to be treated according to sample size planning is already very large and thus the effect is at the same time very weak may therefore not add much value and could be avoided in this way.

Overall, the number of studies with application of AR-based surgery with OST-HMDs in cadaveric studies is small [19-22]. In contrast, AI-based models can be developed by any researcher today with public data and a few lines of scripting. By this method, breakthrough results in diagnostics and nonsurgical therapy were achieved. However, a similar development for surgery that digitizes the operation field is missing. For this to happen, AR-based applications must become mass-market ready and proof of performance must be provided. We hope that other researchers will feel motivated to develop their cadaver test scenarios with this prototype system.

Conclusion

The development and application of an AR-based surgical system using freely available technologies to perform OST-HMD–guided surgical procedures in cadavers is feasible, but our presented open-source prototype should be further developed. Cadaver studies are suitable for OST-HMD–guided interventions to measure a surgical end point and provide an initial data foundation for future clinical trials. In this regard, it has been shown in our scenario that the effect of the AR-based approach could be more likely to make a difference in residents. This should be considered when planning future trials. The availability of free systems for researchers could be helpful for a possible translation process from digital health to AR-based surgery using OST-HMDs in the operating theater via cadaver studies.

Acknowledgments

The authors express their posthumous thanks to the body donors who made this study possible. We thank Rainer Röhrig for his valuable feedback and suggestions. Thanks to Axel Honné for helping to 3D print the mountings and half-cube. Thanks to Max Schulze-Hagen for scanning the Stromeyer hook with a computer tomograph. This research received no external funding. The data presented in this study are available from the corresponding author upon reasonable request. The source code of the software will be available online with publication.

Authors' Contributions

BP was responsible for conceptualization, methodology, software, investigation, resources, original draft preparation, visualization, and project administration. AM was responsible for conceptualization, investigation, supervision, and project administration. MB was responsible for methodology. MO was responsible for methodology and investigation. TWK was responsible for methodology and supervision. MC was responsible for the software. FP was responsible for investigation. FH was responsible for investigation, resources, and supervision. SCM and AP were responsible for resources. All authors were responsible for formal analysis and reviewing and editing the manuscript. All authors had full access to all the data, have read and agreed to the published version of the manuscript and agree to be accountable for all aspects of work ensuring integrity and accuracy.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of zygomatic arch fractures.
[DOCX File , 24 KB - games_v10i2e34781_app1.docx]

Multimedia Appendix 2
Questionnaire.
[DOCX File , 20 KB - games_v10i2e34781_app2.docx]

References

1. Cohen AB, Dorsey ER, Mathews SC, Bates DW, Safavi K. A digital health industry cohort across the health continuum. NPJ Digit Med 2020;3:68 [FREE Full text] [doi: 10.1038/s41746-020-0276-9] [Medline: 32411829]
2. Shrime MG, Bickler SW, Alkire BC, Mock C. Global burden of surgical disease: an estimation from the provider perspective. Lancet Glob Health 2015 Apr 27;3 Suppl 2:S8-S9 [FREE Full text] [doi: 10.1016/S2214-109X(14)70384-5] [Medline: 25926322]
3. Jayaraman PP, Forkan ARM, Morshed A, Haghighi PD, Kang Y. Healthcare 4.0: A review of frontiers in digital health. WIREs Data Mining Knowl Discov 2019 Dec 25;10(2):1. [doi: 10.1002/widm.1350]
4. Tarassoli SP. Artificial intelligence, regenerative surgery, robotics? What is realistic for the future of surgery? Ann Med Surg (Lond) 2019 May;41:53-55 [FREE Full text] [doi: 10.1016/j.amsu.2019.04.001] [Medline: 31049197]
5. Kersten-Oertel M, Jannin P, Collins DL. DVV: a taxonomy for mixed reality visualization in image guided surgery. IEEE Trans Vis Comput Graph 2012 Feb;18(2):332-352. [doi: 10.1109/TVCG.2011.50] [Medline: 21383411]
6. Santi GM, Ceruti A, Liverani A, Osti F. Augmented reality in industry 4.0 and future innovation programs. Technologies 2021 Apr 29;9(2):33. [doi: 10.3390/technologies9020033]
7. Atallah S. Digital Surgery. Cham: Springer International Publishing; 2021.
8. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Transactions Inf Syst 1994;77(12):1321-1329 [FREE Full text]
9. Sigrist R, Rauter G, Riener R, Wolf P. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. Psychon Bull Rev 2013 Feb;20(1):21-53. [doi: 10.3758/s13423-012-0333-8] [Medline: 23132605]
10. Badiali G, Cercenelli L, Battaglia S, Marcelli E, Marchetti C, Ferrari V, et al. Review on augmented reality in oral and cranio-maxillofacial surgery: toward “surgery-specific” head-up displays. IEEE Access 2020;8:59015-59028. [doi: 10.1109/access.2020.2973298]
11. Scherl C, Stratemeier J, Rotter N, Hesser J, Schönberg SO, Servais JJ, et al. Augmented reality with HoloLens® in parotid tumor surgery: a prospective feasibility study. ORL J Otorhinolaryngol Relat Spec 2021;83(6):439-448 [FREE Full text] [doi: 10.1159/000514640] [Medline: 33784686]
12. McJunkin JL, Jiramongkolchai P, Chung W, Southworth M, Durakovic N, Buchman CA, et al. Development of a mixed reality platform for lateral skull base anatomy. Otol Neurotol 2018 Dec;39(10):e1137-e1142 [FREE Full text] [doi: 10.1097/MAO.0000000000001995] [Medline: 30239435]
13. Watzinger F, Birkfellner W, Wanschitz F, Ziya F, Wagner A, Kremser J, et al. Placement of endosteal implants in the zygoma after maxillectomy: a Cadaver study using surgical navigation. Plast Reconstr Surg 2001 Mar;107(3):659-667. [doi: 10.1097/00006534-200103000-00003] [Medline: 11304589]
14. Citardi MJ, Agbetoba A, Bigcas J, Luong A. Augmented reality for endoscopic sinus surgery with surgical navigation: a cadaver study. Int Forum Allergy Rhinol 2016 May;6(5):523-528 [FREE Full text] [doi: 10.1002/alr.21702] [Medline: 26718626]
15. Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: a class of displays on the reality-virtuality continuum. Telemanipulator Telepresence Technol 1995:282-292. [doi: 10.1117/12.197321]
16. Sutherland I. A head-mounted three dimensional display. AFIPS '68 Proc Fall Joint Comput Conf, Part I 1968:757-764. [doi: 10.1145/1476589.1476686]
17. Rolland JP, Fuchs H. Optical versus video see-through head-mounted displays in medical visualization. Presence Teleoperators Virtual Envir 2000 Jun;9(3):287-309. [doi: 10.1162/105474600566808]

18. Recommended content and format of non-clinical bench performance testing information in premarket submissions: guidance for industry and Food and Drug Administration staff. FDA. 2019. URL: https://www.fda.gov/media/113230/download [accessed 2021-09-21]
19. Hoch A, Liebmann F, Carrillo F, Farshad M, Rahm S, Zingg P, et al. Augmented reality based surgical navigation of the periacetabular osteotomy of ganz: a pilot cadaveric study. In: Rauter G, Cattin P, Zaman A, Riener R, editors. New Trends in Medical and Service Robotics. Cham: Springer; 2021:192-201.
20. Kriechling P, Loucas R, Loucas M, Casari F, Fürnstahl P, Wieser K. Augmented reality through head-mounted display for navigation of baseplate component placement in reverse total shoulder arthroplasty: a cadaveric study. Arch Orthop Trauma Surg 2021 Jul 02:1. [doi: 10.1007/s00402-021-04025-5] [Medline: 34213578]
21. Spirig JM, Roner S, Liebmann F, Fürnstahl P, Farshad M. Augmented reality-navigated pedicle screw placement: a cadaveric pilot study. Eur Spine J 2021 Aug 04;30:3731-3737. [doi: 10.1007/s00586-021-06950-w] [Medline: 34350487]
22. Molina CA, Phillips FM, Colman MW, Ray WZ, Khan M, Orru' E, et al. A cadaveric precision and accuracy analysis of augmented reality-mediated percutaneous pedicle implant insertion. J Neurosurg Spine 2020 Oct 30:1-9. [doi: 10.3171/2020.6.SPINE20370] [Medline: 33126206]
23. Cesov M. Intra-Operative Guide in Augmented Reality With Head Mounted Devices in Oral and Maxillofacial Surgery. Aachen: RWTH Aachen University; Sep 07, 2020.
24. Yamamoto K, Murakami K, Sugiura T, Fujimoto M, Inoue M, Kawakami M, et al. Clinical analysis of isolated zygomatic arch fractures. J Oral Maxillofac Surg 2007 Mar;65(3):457-461. [doi: 10.1016/j.joms.2006.06.276] [Medline: 17307593]
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Human Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
26. Canty A, Ripley BD. boot: Bootstrap R (S-Plus) functions. 2021. URL: https://cran.r-project.org/web/packages/bootstrap/ [accessed 2022-03-22]
27. Microsoft: HoloLens documentation. 2021. URL: https://docs.microsoft.com/en-us/hololens/ [accessed 2021-10-28]
28. Mautz R. Indoor positioning technologies. ZürichTH. Zurich: ETH Zurich, Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry; 2012. URL: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/54888/eth-5659-01.pdf?sequence=1&isAllowed=y [accessed 2022-03-22]
29. Frantz T, Jansen B, Duerinck J, Vandemeulebroucke J. Augmenting Microsoft's HoloLens with vuforia tracking for neuronavigation. Healthc Technol Lett 2018 Oct;5(5):221-225 [FREE Full text] [doi: 10.1049/htl.2018.5079] [Medline: 30464854]
30. Seifabadi R, Li M, Long D, Xu S, Wood B. Accuracy study of smartglasses/smartphone AR systems for percutaneous needle interventions. Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling February 2020, Houston, Texas, United States. Bellingham, WashingtonPIE; 2020 Presented at: Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling; 2020; Houston p. 30. [doi: 10.1117/12.2549278]
31. Lee J, Kim S, Kim L, Kang J, Lee S, Kwon S. A study on virtual studio application using Microsoft Hololens. Int J Adv Smart Converg 2017;6(4):80-87. [doi: 10.7236/IJASC.2017.6.4.12]
32. Pepe A, Trotta GF, Mohr-Ziak P, Gsaxner C, Wallner J, Bevilacqua V, et al. A marker-less registration approach for mixed reality-aided maxillofacial surgery: a pilot evaluation. J Digit Imaging 2019 Dec;32(6):1008-1018 [FREE Full text] [doi: 10.1007/s10278-019-00272-6] [Medline: 31485953]
33. Scherl C, Stratemeier J, Karle C, Rotter N, Hesser J, Huber L, et al. Augmented reality with HoloLens in parotid surgery: how to assess and to improve accuracy. Eur Arch Otorhinolaryngol 2021 Jul;278(7):2473-2483. [doi: 10.1007/s00405-020-06351-7] [Medline: 32910225]
34. Gibelli D, Cellina M, Gibelli S, Oliva AG, Termine G, Pucciarelli V, et al. Assessing symmetry of zygomatic bone through three-dimensional segmentation on computed tomography scan and "mirroring" procedure: a contribution for reconstructive maxillofacial surgery. J Craniomaxillofac Surg 2018 Apr;46(4):600-604. [doi: 10.1016/j.jcms.2018.02.012] [Medline: 29548881]
35. Bork F, Fuers B, Schneider A, Pinto F, Graumann C, Navab N. Auditory and visio-temporal distance coding for 3-dimensional perception in medical augmented reality. 2015. URL: http://campar.in.tum.de/pub/bork2015ismar/bork2015ismar.pdf [accessed 2022-03-22]
36. Cheng H, Clymer JW, Po-Han Chen B, Sadeghirad B, Ferko NC, Cameron CG, et al. Prolonged operative duration is associated with complications: a systematic review and meta-analysis. J Surg Res 2018 Sep;229:134-144 [FREE Full text] [doi: 10.1016/j.jss.2018.03.022] [Medline: 29936980]
37. Mischkowski RA, Zinser MJ, Kübler AC, Krug B, Seifert U, Zöller JE. Application of an augmented reality tool for maxillary positioning in orthognathic surgery: a feasibility study. J Craniomaxillofac Surg 2006 Dec;34(8):478-483. [doi: 10.1016/j.jcms.2006.07.862] [Medline: 17157519]

Abbreviations

AI: artificial intelligence
AR: augmented reality

CBCT: cone beam computed tomography
FOV: field of view
IGS: image-guided surgery
OST-HMD: optical see-through head-mounted display
SUS: System Usability Scale

Edited by N Zary; submitted 08.11.21; peer-reviewed by J Egger, C Scherl, A Cohen; comments to author 30.12.21; revised version received 04.01.22; accepted 05.03.22; published 25.04.22.

Please cite as:
Puladi B, Ooms M, Bellgardt M, Cesov M, Lipprandt M, Raith S, Peters F, Möhlhenrich SC, Prescher A, Hölzle F, Kuhlen TW, Modabber A
Augmented Reality-Based Surgery on the Human Cadaver Using a New Generation of Optical Head-Mounted Displays: Development and Feasibility Study
JMIR Serious Games 2022;10(2):e34781
URL: https://games.jmir.org/2022/2/e34781
doi:10.2196/34781
PMID:35468090

Original Paper

HelperFriend, a Serious Game for Promoting Healthy Lifestyle Behaviors in Children: Design and Pilot Study

Ismael Edrein Espinosa-Curiel[1], PhD; Edgar Efrén Pozas-Bogarin[1], BSc; Maryleidi Hernández-Arvizu[1], BSc; Maria Elena Navarro-Jiménez[1], BSc; Edwin Emeth Delgado-Pérez[2], MSc; Juan Martínez-Miranda[1], PhD; Humberto Pérez-Espinosa[1], PhD

[1]Centro de Investigación Científica y de Educación Superior de Ensenada, Unidad de Transferencia Tecnológica Tepic, Tepic, Nayarit, Mexico
[2]Centro de Estudios e Investigaciones en Comportamiento, Universidad de Guadalajara, Guadalajara, Jalisco, Mexico

Corresponding Author:
Ismael Edrein Espinosa-Curiel, PhD
Centro de Investigación Científica y de Educación Superior de Ensenada
Unidad de Transferencia Tecnológica Tepic
Andador 10, entre calles 3 y 4, Ciudad del Conocimiento
Tepic, Nayarit, 63173
Mexico
Phone: 52 3111295930 ext 28607
Email: ecuriel@cicese.mx

Abstract

Background: The use of health games is a promising strategy for educating and promoting healthy lifestyle behaviors among children.

Objective: We aimed to describe the design and development of a serious game, called HelperFriend, and evaluate its feasibility, acceptability, and preliminary effects in children in a pilot study. HelperFriend is a vicarious experiential video game designed to promote 3 lifestyle behaviors among young children: physical activity, healthy eating, and socioemotional wellness.

Methods: Participants aged 8 to 11 years were recruited from an elementary school and randomized to receive a healthy lifestyle behavior educational talk (control) or play six 30-minute sessions with HelperFriend (intervention). Assessments were conducted at baseline (T0) and after the intervention (ie, 4 weeks) (T1). The primary outcome was gain in knowledge. The secondary outcomes were intention to conduct healthy behaviors, dietary intake, and player satisfaction.

Results: Knowledge scores of intervention group participants increased from T0 to T1 for physical activity (t_{14}=2.01, P=.03), healthy eating (t_{14}=3.14, P=.003), and socioemotional wellness (t_{14}=2.75, P=.008). In addition, from T0 to T1, the intervention group improved their intention to perform physical activity (t_{14}=2.82, P=.006), healthy eating (t_{14}=3.44, P=.002), and socioemotional wellness (t_{14}=2.65, P=.009); and there was a reduction in their intake of 13 unhealthy foods. HelperFriend was well received by intervention group.

Conclusions: HelperFriend appears to be feasible and acceptable for young children. In addition, this game seems to be a viable tool to help improve the knowledge, the intention to conduct healthy behaviors, and the dietary intake of children; however, a well-powered randomized controlled trial is needed to prove the efficacy of HelperFriend.

(JMIR Serious Games 2022;10(2):e33412) doi:10.2196/33412

KEYWORDS

serious game; children; education and behavior change; healthy lifestyle behaviors; physical activity; healthy eating; socioemotional wellness

Introduction

Healthy Lifestyle Behaviors

Unhealthy lifestyle behaviors (eg, physical inactivity, unhealthy diet, and sedentary time) put individuals at high risk of developing several health conditions (eg, dental caries, hypertension, diabetes, cardiopathy, and cancer) and are key drivers of obesity and being overweight [1-3]. In contrast, healthy lifestyle behaviors (eg, physical activity and healthy diet) can provide a general feeling of well-being and are the foundation of disease prevention [4]. Individuals who embrace healthy lifestyle behaviors can withstand health risks linked to disability and illness in later life [4]. Healthy lifestyle behaviors can benefit people of all ages; however, it is crucial to encourage these behaviors from early childhood—when habits are formed—because they are likely to be maintained during adulthood [3,5,6].

Two critical healthy lifestyle behaviors for children are healthy eating and physical activity. Children need to have a correct diet [7], increase the intake of healthy food, decrease the intake of unhealthy food [8], and engage in a minimum of 60 minutes of moderate to vigorous physical activity daily [9]. These behaviors help prevent weight gain in children [10,11] and are included in the strategies to face the dramatic increase in childhood obesity, given that childhood obesity is associated with unhealthy eating and physical inactivity [12,13]. Despite the importance of these behaviors, many children do not perform them. For example, in Mexico, only 43.5% of children meet the recommended intake of fruits [14], 22% of children meet the recommended intake of vegetables [14], and 17.3% of children engage in at least 60 minutes of daily physical activity [15]. The promotion of these behaviors should start when children are young (between 8 and 11 years), because these are the years during which rates of being overweight and obesity increase significantly [15].

Lifestyle interventions should focus on healthy eating and physical activity to have a more significant effect on health [10]. In addition, social and emotional factors should be taken into account when developing lifestyle interventions because these factors affect healthy eating and physical activity [16,17]. Therefore, it is essential that children also learn how to recognize and manage emotions, establish and maintain constructive and healthy relationships, make responsible decisions, and avoid unhealthy social and emotional behaviors associated with eating and physical activity [18,19].

Serious Games for Health

In recent years, the development of serious games as innovative methods to support health education and treatment initiatives and programs has increased [20]. Serious games integrate engagement and fun elements (eg, stories, levels, rewards, and feedback) with educational and psychological resources and techniques to achieve health outcomes [21]. Serious games for health can include the simulation of real-life situations, collection of information that supports the identification of behaviors, and provision of information and suggestions to guide the process to improve attitudes and behaviors of players [20]. Thus, they offer the possibility to support initiatives to deliver education and health services to populations that currently cannot or do not obtain necessary access to these services due to costs, logistical issues, stigma, or convenience [22]. There are currently serious health games for a wide variety of purposes, such as health education, physical and psychological therapy, and disease self-management [20,23].

Serious games are increasingly being used to encourage children to adopt healthy lifestyle behaviors, leveraging the fact that most children enjoy playing video games [24]. However, there are mixed opinions about this strategy. One concern is related to screen time—screen time is considered to be a risk factor for several health, emotional, and psychological problems in children [25,26]. However, it appears that screen time for playing video games does not represent as high a risk when compared with that for watching television [27]. Another concern is the effectiveness of these games because, while most studies have reported positive effects on obesity-related outcomes (improvement of weight-related parameters, physical activity, or dietary behavior and knowledge), these effects were small [28,29]. In addition, while many games focus on improving health knowledge, this does not necessarily result in behavioral change [28]. Conversely, several studies [30-32] have shown that serious games offer an enormous advantage for health promotion interventions in children.

Video games for promoting healthy lifestyle behaviors in children are aimed to improve knowledge about nutrition, eating habits, and exercise; increase physical activity while playing (exergames); change eating behaviors; or combine several approaches [28,31,33,34]. Nutrition and eating habits–related games focus on the concepts of energy balance [35,36], MyPlate guidelines [37], the 5 macronutrients of foods [38], Mediterranean diet and behavioral moderation [39], healthy and unhealthy nutrition [40,41], and dietary energy density [42]. Despite the importance of psychosocial or psychological aspects of nutrition and eating habits, only one game considered these aspects through the integration of the coping of stress technique [42]. Although positive results were obtained in these studies [35-42], there is still a need to understand the application and limitations of such games as well as how to improve their effectiveness, such as the inclusion of the underlying mechanisms for behavioral change of video games [31] or the integration of psychosocial aspects in video games [28]. In addition, none of these games implemented a vicarious experiential environment that includes behavior change techniques to promote physical activity, promote healthy eating, and address social and emotional issues related to these behaviors in young children.

Objective and Hypotheses

We aimed to design and develop a motion-controlled serious game for young children (HelperFriend) and evaluate its feasibility, acceptability, and preliminary effects. We hypothesized that children who played the game would demonstrate (1) better knowledge, (2) greater intention to carry out healthy lifestyle behaviors, and (3) improvements in dietary intake and that (4) children would enjoy playing the game.

Methods

HelperFriend Video Game

Design and Development

HelperFriend was developed by a multidisciplinary team that included nutritionists, psychologists, physical activity experts, human–computer interaction experts, and software engineers based on published design methodology [43]. The methodology included activities from game implementation to evaluation based on 4 essential principles: a procedure-centric approach, expert collaboration, agile development, and low-cost modeling.

The knowledge domains of HelperFriend are physical activity, healthy eating, and socioemotional wellness.

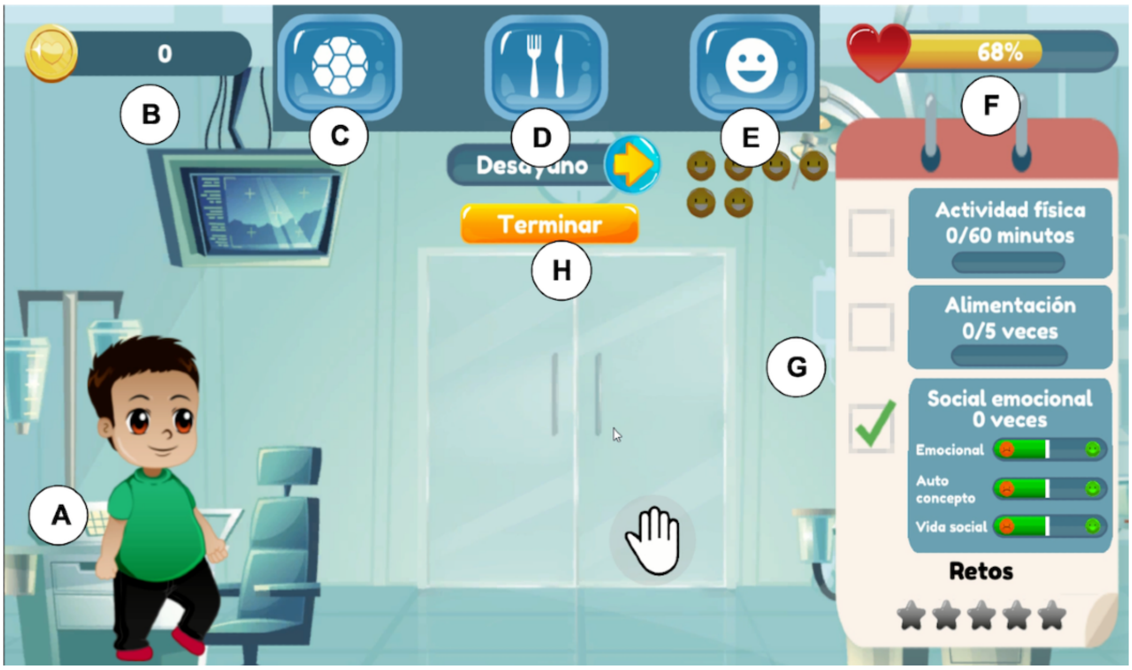
HelperFriend integrates experiential and vicarious learning. In experiential learning environment, learners engage in direct experiences to enhance their knowledge, skills, and values through human–environment interaction in a cycle of doing, reflecting, concluding, and trying the learned experience [44]. In vicarious learning, individuals learn from the experiences of others (eg, by observing the choices another person makes and the consequences they have on their health). By observing the behavior of others, individuals can identify difficulties and expectations associated with behaviors and acquire the information and competencies to perform the behavior successfully [45].

In addition, several behavior change techniques [46] were integrated into the game elements to generate an attractive and stimulating environment in which knowledge and healthy lifestyle behaviors are encouraged and reinforced: instruction on how to perform the behavior, providing information about health consequences, behavioral practice, behavioral substitution, incentives and rewards, goal setting, reviewing behaviors goal, monitoring behaviors, providing feedback on behavior, discrepancies between current behaviors and goals, monitoring emotion consequences, and prompts or cues. These behavior change techniques are based on behavioral, cognitive, and social cognitive theories that have ample empirical evidence to demonstrate their usefulness in adopting healthy lifestyles [47-50].

Description

In HelperFriend, the players are secret agents who need to care for a group of children who forgot healthy lifestyle behaviors because a villain chef erased their memory. In each match, the player needs to ensure that one of these children engages in physical activity, eats well, and performs socioemotional activities to improve their health (Figure 1). The children characters continuously interact with the player, expressing their necessities or stating situations for which they need help. Player actions that improve the children's lifestyle behaviors add points. The game session is finished when the player presses the finish button. At the end, the player can earn extra points if their decisions helped the child to meet healthy lifestyle recommendations. The player has full-body interaction to encourage physical activity and improve satisfaction and fun [51]. A video of the game is provided as Multimedia Appendix 1.

Figure 1. The main screen of HelperFriend: (A) child being cared for, (B) coins score, (C) button to carry out physical activity, (D) button for feeding the child, (E) button to carry out socioemotional activities, (F) health bar of the child, (G) game indicators section, and (H) finish button.



Modules

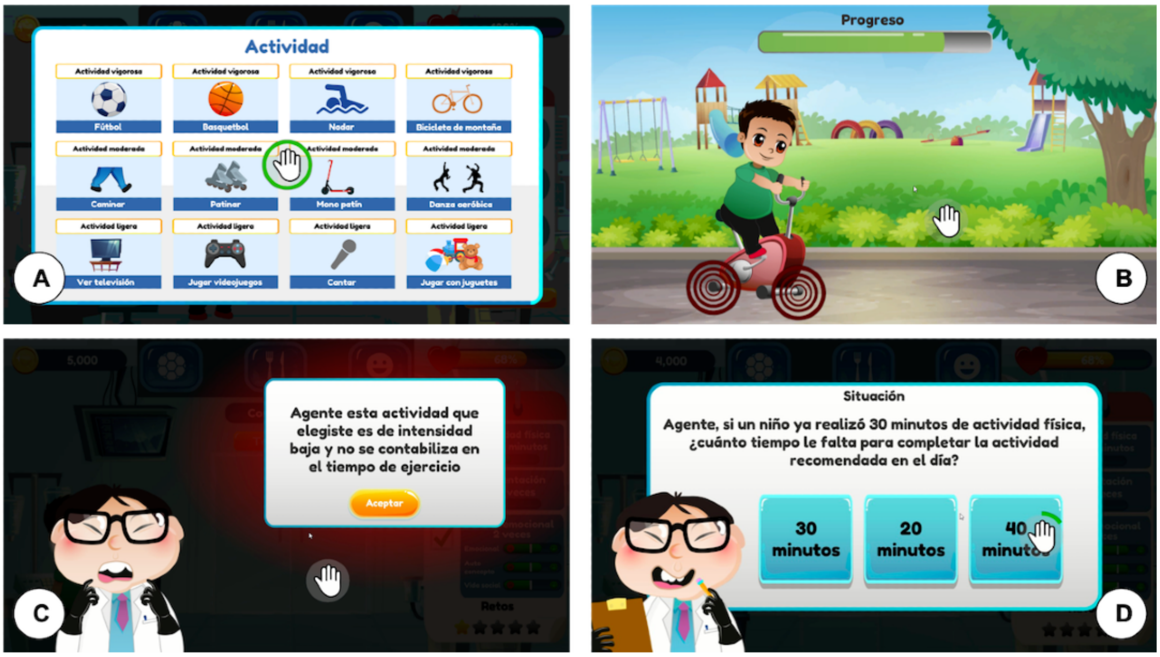
Overview

Each module increases in difficulty to keep players engaged and having fun until the end of the video game. Modules have 3 components. The education component teaches basic health knowledge. The training component encourages players to practice healthy lifestyle behaviors. The challenge component presents challenging situations in which players have to help the children.

Physical Activity

This module (Figure 2) addresses World Health Organization physical activity recommendations for children and adolescents (aged 5 to 17 years). Children and adolescents should engage in a minimum of 60 minutes of moderate to vigorous physical activity on a daily basis, most of which should consist of aerobic exercise [9]. Engaging in more than 60 minutes of physical activity provides additional health benefits. In addition, vigorous physical activities and muscle and bone strengthening activities each should be incorporated at least 3 days per week.

Figure 2. Physical activity screens: (A) screen for selecting a physical activity, (B) child doing physical activity, (C) alert feedback screen because the player selected a sedentary activity, and (D) physical activity situation in which the child needs the player's help.



Healthy Eating

This module (Figure 3) addresses diet. According to the Mexican Official Standard [7], a correct diet for children is one that is complete, balanced, innocuous, suitable, and varied. In addition, this module addresses portion intake and recommendations that children should eat approximately 5 times a day [52], increase water and healthy food (including fruits and vegetables) consumption, and decrease unhealthy food (eg, candies, sweetened cereals, and sugary drinks) consumption [8].

Figure 3. Healthy eating screens: (A) screen teaching player about complete diet, (B) screen for selecting food, (C) feeding information screen, and (D) feeding situation in which the child needs the player's help.

Socioemotional

This module (Figure 4) addresses social and emotional behaviors related to physical activity and healthy eating. Children need to acquire skills to recognize and manage emotions, establish and maintain constructive and healthy relationships, take an interest in the well-being of others, and make responsible decisions [18,19]. Examples of the social and emotional issues included are low motivation to improve exercise and eating habits, emotions associated with eating junk food, and the influence of parents and friends on eating habits and physical activity.

Figure 4. Socioemotional wellness screens: (A) screen where the child shows a socioemotional situation to the player, (B) screen for selecting a socioemotional activity, (C) alert feedback message because the player made an inadequate socioemotional activity choice, and (D) socioemotional situation in which the child needs the player's help.

Intervention

Overview

We conducted a parallel randomized controlled pilot trial over 4 weeks between May and June 2019 in an elementary school in Mexico.

Ethics

School administrators and teachers gave written permission for the trial to be performed at school facilities. All study procedures were approved by the institutional review board of the *Centro de Investigación Científica y de Educación Superior de Ensenada* (2S.3.1 HUM 2019). No changes occurred to the methods after the beginning of the trial.

Participants

Students (n=40) from 3 school groups was considered for this study. Inclusion criteria were being aged 8 to 11 years and not receiving pharmacological treatment. Exclusion criteria were having been diagnosed with or having an ongoing neuropsychiatric disorder, a physical problem (because the game required children to interact through whole-body movements), and obesity treatment in the past 6 months. Written informed consent was obtained from parents of children who expressed interest in participating in the study.

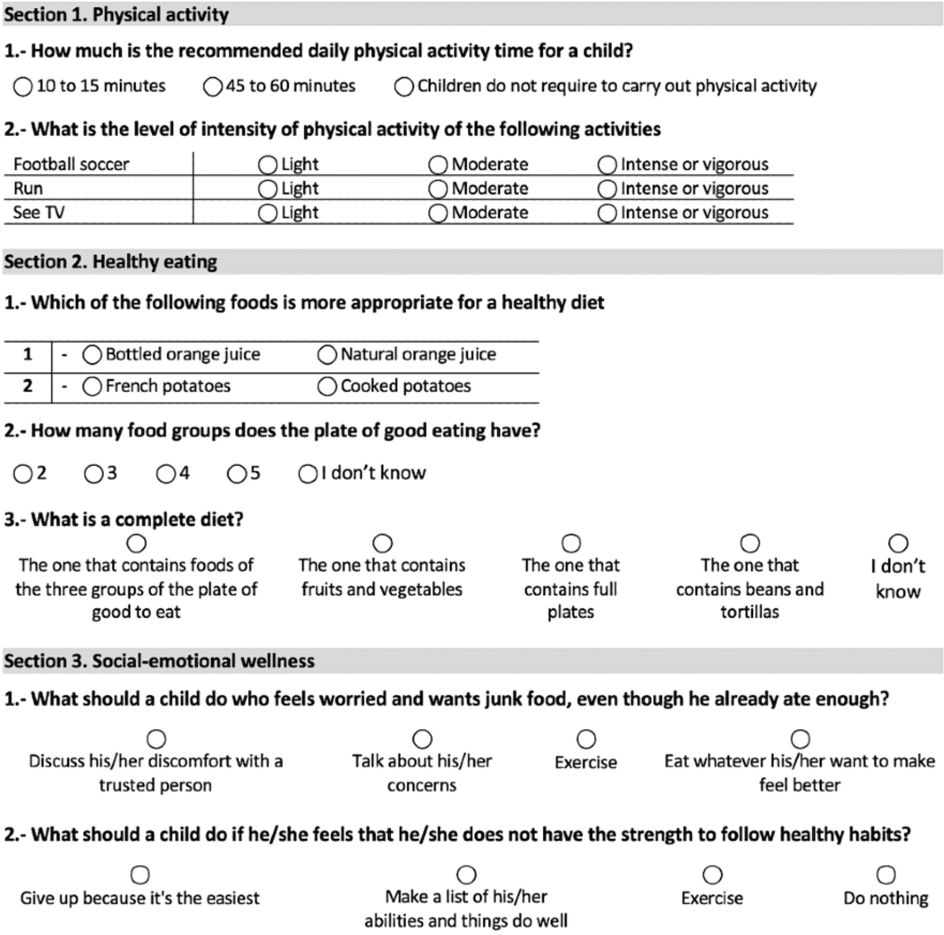
Design

Children were randomly allocated to either the control group or the intervention group. The children in the intervention group played HelperFriend during six 30-minute game sessions. All playing sessions were conducted over 21 days. We set up 3 gaming stations in a room; each station contained a PC, a 50-inch screen, a Kinect sensor V2, and the HelperFriend video game. Participants in the control group received only a 45-minute talk about the importance of healthy behaviors, such as engaging in physical activity, eating healthy, and maintaining socioemotional health; no further intervention was applied.

Outcome Measures

Outcomes were assessed in both groups the week after being assigned to the groups (T0) and 4 weeks after baseline (T1). The primary outcome was the gain in knowledge measured using a questionnaire (developed by the research group and designed specifically for the serious game). The questionnaire was evaluated in a pilot with 5 children and adapted. The final questionnaire consisted of 82 questions in 3 sections: physical activity (13 questions, each with 3 response options), healthy eating (64 questions, each with 3 to 5 response options in food groups, food portions equivalence, correct diet, and healthy/unhealthy food subsections), and socioemotional wellness (5 questions, each with 4 response options). Figure 5 provides some examples of the questions. Children completed the questionnaire by themselves. The sum of questions that had been appropriately answered for each section was calculated.

Figure 5. Healthy behaviors knowledge questionnaire example questions.



Section 1. Physical activity

1.- How much is the recommended daily physical activity time for a child?

◯ 10 to 15 minutes ◯ 45 to 60 minutes ◯ Children do not require to carry out physical activity

2.- What is the level of intensity of physical activity of the following activities

Football soccer	◯ Light	◯ Moderate	◯ Intense or vigorous
Run	◯ Light	◯ Moderate	◯ Intense or vigorous
See TV	◯ Light	◯ Moderate	◯ Intense or vigorous

Section 2. Healthy eating

1.- Which of the following foods is more appropriate for a healthy diet

1	-	◯ Bottled orange juice	◯ Natural orange juice
2	-	◯ French potatoes	◯ Cooked potatoes

2.- How many food groups does the plate of good eating have?

◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ I don't know

3.- What is a complete diet?

◯ The one that contains foods of the three groups of the plate of good to eat ◯ The one that contains fruits and vegetables ◯ The one that contains full plates ◯ The one that contains beans and tortillas ◯ I don't know

Section 3. Social-emotional wellness

1.- What should a child do who feels worried and wants junk food, even though he already ate enough?

◯ Discuss his/her discomfort with a trusted person ◯ Talk about his/her concerns ◯ Exercise ◯ Eat whatever his/her want to make feel better

2.- What should a child do if he/she feels that he/she does not have the strength to follow healthy habits?

◯ Give up because it's the easiest ◯ Make a list of his/her abilities and things do well ◯ Exercise ◯ Do nothing

Secondary outcomes were intention to conduct healthy behaviors, dietary intake, and player experience satisfaction.

Children's intention to conduct healthy behaviors was measured using a questionnaire tailored specifically for the serious game. The questionnaire was pilot-tested with 5 children and adapted. The final questionnaire (Multimedia Appendix 2) consisted of 33 questions in 3 sections: physical activity (4 questions), healthy eating and correct diet (24 questions), and socioemotional wellness (5 questions). These questions state everyday situations that children experiment in their daily lives, and they have to decide what action to take to solve the case. Physical activity, healthy eating, and socioemotional wellness questions have 3, 3, and 4 response options, respectively. Questions related to correct diet have 3 response options and use a graphical representation to facilitate children's answers (a previous study [53] used similar graphical questionnaires with children). For each question, there was only 1 appropriate answer, and children completed the questionnaire by themselves. The sum of questions that had been appropriately answered for each section was calculated.

Dietary behavior was measured with a food frequency intake questionnaire [41]. This questionnaire was explicitly designed for the diet of school-age Mexican children and included features of the highest validated food frequency questionnaires. It consists of 78 food items considered to be indicators for healthy and unhealthy eating behaviors. For each item, a 7-point scale, from 0 (never) to 6 (two or more times per day), is used to indicate the frequency that the food is consumed. Each question was scored individually. In order to facilitate the completion of the questionnaire, a facilitator read the questions to the children, who only had to answer with the number of times that they had eaten the food in the past month.

Player experience satisfaction was only assessed in the intervention group (at T1). An adapted version of the Game User Experience Satisfaction Scale [54] was used. The scale consists of 23 questions in 7 domains: playability, narratives, enjoyment, personal gratification, creative freedom, audio aesthetics, and visual aesthetics. We asked participants to specify their agreement level using a 5-point Likert scale, from 1 (totally disagree) to 5 (totally agree). A score was calculated for each domain.

Data Analysis

Data were analyzed using SPSS software (version 26; IBM Corp). The statistical significance for all analyses was $P<.05$. Variables are reported as means (with standard deviations) or medians (with interquartile ranges). The normality distribution of interval variables was tested using the Shapiro-Wilk test. For metric data, differences between pretest and posttest were analyzed using 1-tailed paired *t* tests. For nonmetric data, the differences between pretest and posttest were analyzed using Wilcoxon signed-rank sum tests. Differences between groups were analyzed using 1-tailed independent *t* tests. The relationships between subscale items were tested using Cronbach α. No power estimation was performed since this was a pilot study.

Results

Participant Characteristics

Of 40 children approached for the trial, 27 (68%) children agreed to participate (age: mean 9.9 years, SD 0.9 years; girls: 16/27, 59%; boys: 11/27, 41%). The control group had 12 participants (age: mean 9.8 years, SD 0.62; girls: 7/12, 58%; boys: 4/12, 33%), and the intervention group had 15 participants (age: mean 9.9 years, SD 0.94; girls: 9/15, 60%; boys: 6/15, 40%). We created a game environment where children felt comfortable during game sessions; however, 1 child missed 1 session, and 6 children missed 2 sessions. Participants in the intervention group played an average of 3.1 hours.

Primary Outcome: Healthy Behaviors Knowledge

Knowledge of intervention group participants increased significantly from T0 to T1 for physical activity (t_{14}=2.01, *P*=.03), healthy eating (t_{14}=3.14, *P*=.003), and socioemotional wellness (t_{14}=2.75, *P*=.008). There were no significant changes in the knowledge of the control group participants from T0 to T1 for physical activity (t_{11}=0.29, *P*=.39), healthy eating (t_{11}=–0.64, *P*=.27), and socioemotional wellness (t_{11}=0.01, *P*=.50). At T1, between-group differences were statistically significant for physical activity (t_{25}=1.98, *P*=.03), healthy eating (t_{25}=1.85, *P*=.04), and socioemotional wellness (t_{25}=1.97, *P*=.03); the intervention group scored higher than the control group for all 3 (Table 1).

Table 1. Outcomes of healthy behaviors knowledge and intention to conduct healthy behaviors.

Measure	Items, n	Control group			Intervention group			Between-group postintervention comparison
		Baseline (T0), mean (SD)	Postintervention (T1), mean (SD)	*P* value	Baseline (T0), mean (SD)	Postintervention (T1), mean (SD)	*P* value	*P* value
Knowledge								
Physical activity	13	4 (1.5)	4.08 (1.6)	.39	4.5 (2)	5.7 (2.4)	.03	.03
Healthy eating	64	38.9 (11.9)	38.3 (12.9)	.27	40 (9.2)	45.5 (7.1)	.003	.04
Socioemotional wellness	5	1.83 (1.2)	1.83 (1.2)	.50	2.1 (1.2)	2.9 (1.4)	.008	.03
Intention to conduct healthy behaviors								
Physical activity	4	2.4 (1.4)	2.3 (1.4)	.22	2.13 (1.6)	3.07 (0.8)	.006	.03
Healthy eating	24	13.9 (4.3)	14.2 (3.5)	.35	13.5 (3.5)	16.3 (2.2)	.002	.03
Socioemotional wellness	5	4.4 (1.2)	4.5 (0.8)	.41	4.6 (0.5)	4.9 (0.25)	.009	.01

Secondary Outcomes

Intention to Conduct Healthy Behaviors

Intention to perform healthy behaviors in the intervention group increased from T0 to T1 for physical activity (t_{14}=2.82, *P*=.006), healthy eating (t_{14}=3.44, *P*=.002), and socioemotional wellness (t_{14}=2.65, *P*=.009). There were no significant differences in intention to conduct healthy lifestyle behaviors for control group participants for physical activity (t_{11}=–0.80, *P*=.22), healthy eating (t_{11}=0.40, *P*=.40), and socioemotional wellness (t_{11}=0.23, *P*=.41). In addition, differences in T1 intention scores between the intervention group and control group were statistically significant for physical activity (t_{25}=1.95, *P*=.03), healthy eating (t_{25}=1.91, *P*=.03), and socioemotional wellness (t_{25}=2.43, *P*=.01); the intervention group scored higher than the control group for all 3 (Table 1).

Food Frequency Intake

Participants in the intervention group reported reduced consumption frequency of ham, sausage, soft drinks, wheat burritos, hamburgers, breaded chicken, sopes, tamales, salt peanuts, sweet cookies, potatoes chips, cake, and sweet soft cakes. Participants in the control group indicated reduced self-reported frequency intake of 5 healthy foods (cantaloupe, carrot, fish soup, fish ceviche, and fresh fruit juice) and 1 unhealthy food (bottled fruit juice) (Table 2).

Table 2. Outcomes of food frequency intake.

	Classification	Baseline (T0) score[a], median (IQR)	Postintervention (T1) score[a], median (IQR)	*P* value
Intervention group				
Ham	Unhealthy	2 (1.3)	1 (2)	.02
Sausage	Unhealthy	2 (1.3)	1 (2)	.02
Soft drinks	Unhealthy	1 (2)	1 (2)	.02
Wheat burritos	Unhealthy	1 (2)	0 (1)	.03
Hamburgers	Unhealthy	1 (1)	1 (1)	.01
Breaded chicken	Unhealthy	0 (1)	0 (0)	.048
Sopes	Unhealthy	2 (1.3)	0 (1)	.02
Tamales	Unhealthy	1 (1)	0 (1)	.02
Salt peanuts	Unhealthy	1 (1.3)	0 (1)	.04
Sweet cookies	Unhealthy	1 (1.3)	0 (1.25)	.01
Potato chips	Unhealthy	1 (1)	0 (1)	.03
Cake	Unhealthy	1 (1.3)	0 (1)	.04
Sweet soft cakes	Unhealthy	1 (0.3)	0 (1)	.02
Control group				
Cantaloupe	Healthy	1 (1.25)	0 (1)	.047
Carrot	Healthy	1 (2)	0 (1)	.02
Fish soup	Healthy	0.5 (2)	0 (0)	.03
Fish ceviche	Healthy	2 (1.15)	0.5 (1)	.04
Fresh fruit juice	Healthy	1.5 (2.25)	0 (1)	.02
Bottle fruit juice	Unhealthy	1.5 (1.25)	0.5 (1)	.04

[a]0 indicated never, 1 indicated one to three times per month, 2 indicated once per week, 3 indicated two to four times per week, 4 indicated five to six times per week, 5 indicated daily, and 6 indicated two or more times per day.

Player Video Game Satisfaction

Satisfaction ratings were significantly higher than the neutral value for all domains: playability (t_{14}=7.04, *P*<.001), narrative (t_{14}=4.00, *P*<.001), enjoyment (t_{14}=4.77, *P*<.001), creative freedom (t_{14}=7.69, *P*<.001), audio aesthetics (t_{14}=4.33, *P*<.001), personal gratification (t_{14}=5.99, *P*<.001), and visual aesthetics (t_{14}=5.12, *P*<.001). Most participants agreed that the game was easy to learn to play and use (14/14, 100%), has a clear history (12/15, 80%), is fun and original (11/15, 73%), has good music (12/15, 80%), has good graphics (12/15, 80%), and made them feel successful when they overcame the game's challenges (13/15, 87%). In addition, most participants wanted to play HelperFriend again (12/15, 80%). All measures obtained Cronbach α values ≥.73, except for narrative (Cronbach α=.56) (Table 3).

Table 3. Player satisfaction questionnaire results.

Measure	Items, n	Cronbach α (n=15)	Mean (SD)	Neutral value	*P* value	Lower	Upper
Playability	9	.76	4 (0.5)	3	<.001	2.8	5
Narrative	2	.56	3.9 (0.9)	3	<.001	1.5	5
Enjoyment	3	.73	4 (0.8)	3	<.001	2.7	5
Creative freedom	2	.73	4.1 (0.6)	3	<.001	3	5
Audio aesthetics	3	.90	4.1 (1)	3	<.001	1.7	5
Personal gratification	4	.75	4.2 (0.7)	3	<.001	2.5	5
Visual aesthetics	2	.85	4 (0.8)	3	<.001	2.5	5

Discussion

Principal Findings

Knowledge

Children in the intervention group significantly improved their knowledge about physical activity (t_{14}=2.01, P=.03), healthy eating (t_{14}=3.14, P=.003), and socioemotional wellness (t_{14}=2.75, P=.008) after gameplay; thus, the first hypothesis was verified. Previous studies have also shown that video games for health can help children improve their understanding of physical activity [55] and healthy eating [35,37-39,41,55]. Other than in a recent study [42], in which stress and stress-coping strategies were included, emotional and social issues related to adopting healthy lifestyle behaviors have not been previously taken into account in interventions with serious games that have targeted healthy eating. We also identified that it is possible to teach this type of issue with serious games. Our results support those of a previous study [55], showing that it is feasible to improve physical activity and healthy eating knowledge together. These promising results may be explained by the experiential and vicarious learning environment of the game in which players observe the behavior of others [45] and engage in direct experiences through doing, reflecting, concluding, and trying the learned experience [44]. Conversely, because improving health knowledge through serious games does not necessarily result in behavioral change [28], we also plan to adjust content and learning strategies in a future version.

Intention to Conduct Healthy Behaviors

Second, we hypothesized that the intention to conduct physical activity, healthy eating, and healthy socioemotional behaviors would be higher after the intervention; we also verified this hypothesis. A previous study [40] also showed that video games for health could help children improve dietary and exercise attitudes, but we did not identify any studies on intention to conduct healthy socioemotional behaviors. Improved intention to engage in healthy behaviors may be as a result of integrating multiple behavior change techniques—a cornerstone for efficacy in behavior change interventions [56]. Because a change in intention leads to a small to medium change in behavior [57], these findings help assess the potential impact of video games on children's lifestyle behaviors. The smallest increase in intention was for socioemotional wellness, possibly because the intervention group already scored higher at baseline, which may be associated with the fact that the questionnaire contained simple questions that make it easy to obtain a high score.

Food Frequency Intake

The third hypothesis stated that, after the intervention, children's diets would improve. A lower intake frequency was found for 13 unhealthy foods (such as soft drinks, hamburgers, sweet cookies, potatoes chips, and sweet soft cakes). These changes are relevant because Mexican children commonly consume these foods in schools and at home [58] and changes in children's diets and eating habits can promote changes in the whole family [59]. Our findings confirm that games to improve healthy food consumption are beneficial, which has been demonstrated in some earlier studies (eg, [36,39,41]), but not others (eg, [42]). Unlike some previous studies (eg, [35,39]), we only achieved a reduction in the consumption of unhealthy foods. One possible explanation is that providing information or visual images of foods alone is insufficient to increase children's preferences for the intake of healthy foods [60,61]. Instead, repeated exposure to healthy foods is more effective for improving children's preferences [61]—even more so than strategies based on rewards [62]. A future version of HelperFriend should include behavior change techniques (eg, self-monitoring, setting and examining goals, and action planning) that directly support and encourage healthy food intake. Surprisingly, the participants in the control group indicated reduced intake of 5 healthy foods (cantaloupe, carrot, fish soup, fish ceviche, fresh fruit juice) and 1 unhealthy food (bottled fruit juice); however, we did not collect any other information from the control group that could facilitate the interpretation of this result.

Game Acceptance and Satisfaction

The fourth hypothesis was also verified; children felt good during gameplay, and game acceptance was high. HelperFriend obtained very positive results on personal gratification, playability, creative freedom, enjoyment, narrative, and visual and audio aesthetics—factors which have been shown to be correlated with and predictors of learning [63]. However, for four specific aspects, there is room for improvement: (1) improving the fit of the difficulty curve of the game to the capacities of the children, (2) simplifying the game flow to foster player autonomy, (3) increasing socioemotional elements, and (4) implementing a daily activity tracking system in the game to make it easier for children to understand the daily activities they have to carry out to have healthy lifestyles. These characteristics could be essential aspects that positively influence the general perception, acceptability, and effectiveness of the game.

Limitations

First, the results should be cautiously interpreted because a small group of children participated in the study. However, given that we aimed to evaluate the feasibility, acceptability, and preliminary effects of HelperFriend, our findings can offer valuable information in designing health games for children to improve lifestyle behaviors and that consider socioemotional issues. Second, medium- and long-term effects were not examined. Medium- and long-term studies could provide interesting findings since video games, especially those involving physical activity, can become boring quickly [64]. Third, we developed the intention questionnaire because we did not find any available for young children; however, a detailed review would be necessary prior to its use in a full randomized controlled trial. Finally, the frequency of food intake was self-reported. The results could be limited by the known constraints of food frequency questionnaires, such as trouble recalling experiences and over- or understatement of food intake [65]. Nevertheless, food frequency questionnaires are the most frequently used approach because they are easy to use, reliable, and valid. In addition, there is previous evidence that children's self-reported food intake is more accurate than that reported by parents [66].

We plan to conduct a randomized controlled clinical trial with sample size calculation to address some of these limitations. Moreover, we plan to extend the exposure period and conduct repeated exposure to account for medium- and long-term effects. Finally, we plan to improve the intention questionnaire and include another behavioral test (eg, physical activity).

Conclusions

HelperFriend, a vicarious experiential health game for promoting physical activity, healthy eating, and socioemotional wellness, appears to be feasible and acceptable for young children. Preliminary results suggest that this game improves knowledge about and the intention to conduct healthy lifestyle behaviors and improves dietary intake in children. In future versions of HelperFriend, some game elements should be improved and other behavior change techniques that promote children's intake of healthy foods should be integrated. Given that this was a pilot study with a limited sample size, a well-powered randomized controlled trial is needed to determine the efficacy of HelperFriend.

Acknowledgments

This work was supported by Mexican National Council for Science and Technology (grant PDCPN-2015-824). We are thankful to the manager of *Colegio Real de San Juan* for enabling the evaluation of HelperFriend in their institution and to all educators, scholars, and parents who participated in this study. Finally, we are thankful to Amara Janeth Aguilar-Partida and Mitzi Josue Martínez Rosas, who elaborated the graphical design of HelperFriend.

Authors' Contributions

IEEC designed the game and experiment, performed all statistical analyses, and drafted the manuscript. EEPB developed the game and participated in running the experiment. MHA conducted the experiment and performed data collection and processing. EEDP and MENJ designed the game also provided the game's psychological and nutritional foundations. JMM and HPE analyzed and interpreted data and drafted the manuscript. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Video of the game.
[MP4 File (MP4 Video), 135568 KB - games_v10i2e33412_app1.mp4]

Multimedia Appendix 2
Intention to conduct healthy behaviors questionnaire.
[PDF File (Adobe PDF File), 900 KB - games_v10i2e33412_app2.pdf]

References

1. Chung A, Peeters A, Gearon E, Backholer K. Contribution of discretionary food and drink consumption to socio-economic inequalities in children's weight: prospective study of Australian children. Int J Epidemiol 2018 Jun 01;47(3):820-828. [doi: 10.1093/ije/dyy020] [Medline: 29514246]
2. Cureau FV, Sparrenberger K, Bloch KV, Ekelund U, Schaan BD. Associations of multiple unhealthy lifestyle behaviors with overweight/obesity and abdominal obesity among Brazilian adolescents: a country-wide survey. Nutr Metab Cardiovasc Dis 2018 Jul;28(7):765-774. [doi: 10.1016/j.numecd.2018.04.012] [Medline: 29843935]
3. Kelsey MM, Zaepfel A, Bjornstad P, Nadeau KJ. Age-related consequences of childhood obesity. Gerontology 2014;60(3):222-228. [doi: 10.1159/000356023] [Medline: 24434909]
4. Tamanal JM, Kim CH. Promoting healthy lifestyle in high school students: determination of the lifestyle status through the healthy lifestyle screen (HLS) assessment. J Lifestyle Med 2020 Jan 31;10(1):30-43. [doi: 10.15280/jlm.2020.10.1.30] [Medline: 32328446]
5. Birch L, Savage JS, Ventura A. Influences on the development of children's eating behaviours: from infancy to adolescence. Can J Diet Pract Res 2007;68(1):s1-s56 [FREE Full text] [Medline: 19430591]
6. Issanchou S, Habeat consortium. Determining factors and critical periods in the formation of eating habits: results from the Habeat project. Ann Nutr Metab 2017;70(3):251-256 [FREE Full text] [doi: 10.1159/000471514] [Medline: 28407627]
7. Norma Oficial Mexicana NOM-043-SSA2-2012, Servicios básicos de salud. Promoción y educación para la salud en materia alimentaria. Diario Oficial de la Federeación. 2013 Jan 22. URL: http://dof.gob.mx/nota_detalle.php?codigo=5285372&fecha=22/01/2013 [accessed 2018-04-20]
8. Nishtar S, Gluckman P, Armstrong T. Ending childhood obesity: a time for action. Lancet 2016 Mar 27;387(10021):825-827. [doi: 10.1016/S0140-6736(16)00140-9] [Medline: 26823046]

9. Recomendaciones mundiales sobre actividad física para la salud. Organización Mundial de la Salud. 2010. URL: https://apps.who.int/iris/bitstream/handle/10665/44441/9789243599977_spa.pdf [accessed 2018-03-11]
10. Brown T, Moore TH, Hooper L, Gao Y, Zayegh A, Ijaz S, et al. Interventions for preventing obesity in children. Cochrane Database Syst Rev 2019 Jul 23;7:CD001871 [FREE Full text] [doi: 10.1002/14651858.CD001871.pub4] [Medline: 31332776]
11. Westenhoefer J. Establishing dietary habits during childhood for long-term weight control. Ann Nutr Metab 2002;46 Suppl 1:18-23 [FREE Full text] [doi: 10.1159/000066396] [Medline: 12428077]
12. Hesketh K, Waters E, Green J, Salmon L, Williams J. Healthy eating, activity and obesity prevention: a qualitative study of parent and child perceptions in Australia. Health Promot Int 2005 Mar;20(1):19-26. [doi: 10.1093/heapro/dah503] [Medline: 15668217]
13. Lobstein T, Jackson-Leach R, Moodie ML, Hall KD, Gortmaker SL, Swinburn BA, et al. Child and adolescent obesity: part of a bigger picture. Lancet 2015 Jun 20;385(9986):2510-2520 [FREE Full text] [doi: 10.1016/S0140-6736(14)61746-3] [Medline: 25703114]
14. Gaona-Pineda EB, Martínez-Tapia B, Arango-Angarita A, Valenzuela-Bravo D, Gómez-Acosta LM, Shamah-Levy T, et al. [Food groups consumption and sociodemographic characteristics in Mexican population]. Salud Publica Mex 2018;60(3):272-282. [doi: 10.21149/8803] [Medline: 29746744]
15. 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. In: Instituto Nacional de Salud Pública. Cuernavaca, México: Instituto Nacional de Salud Pública; 2020.
16. Aparicio E, Canals J, Arija V, De Henauw S, Michels N. The role of emotion regulation in childhood obesity: implications for prevention and treatment. Nutr Res Rev 2016 Jun;29(1):17-29. [doi: 10.1017/S0954422415000153] [Medline: 27045966]
17. Salvy S, de LHK, Bowker JC, Hermans RCJ. Influence of peers and friends on children's and adolescents' eating and activity behaviors. Physiol Behav 2012 Jun 06;106(3):369-378 [FREE Full text] [doi: 10.1016/j.physbeh.2012.03.022] [Medline: 22480733]
18. Elias JM, Zins JE, Weissberg RP, Frey KS, Greenberg MT, Haynes NM, et al. Promoting Social and Emotional Learning: Guidelines for Educators. Alexandria, Virginia: ASCD; 1997.
19. Zins JE, Elias MJ. Social and emotional learning: promoting the development of all students. J Educ Psychol Consult 2007 Jul 30;17(2-3):233-255. [doi: 10.1080/10474410701413152]
20. 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]
21. 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 Mar;5(1):1-12 [FREE Full text] [doi: 10.1089/g4h.2015.0026] [Medline: 26262772]
22. Ricciardi F, De Paolis LT. A comprehensive review of serious games in health professions. Int J Comput Games Technol 2014 Aug;2014(787968):1-11. [doi: 10.1155/2014/787968]
23. 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]
24. Holzmann SL, Dischl F, Schäfer H, Groh G, Hauner H, Holzapfel C. Digital gaming for nutritional education: a survey on preferences, motives, and needs of children and adolescents. JMIR Form Res 2019 Mar 13;3(1):e10284 [FREE Full text] [doi: 10.2196/10284] [Medline: 30758290]
25. Liu W, Wu X, Huang K, Yan S, Ma L, Cao H, et al. Early childhood screen time as a predictor of emotional and behavioral problems in children at 4 years: a birth cohort study in China. Environ Health Prev Med 2021 Jan 07;26(1):3 [FREE Full text] [doi: 10.1186/s12199-020-00926-w] [Medline: 33413099]
26. Lissak G. Adverse physiological and psychological effects of screen time on children and adolescents: literature review and case study. Environ Res 2018 Jul;164:149-157. [doi: 10.1016/j.envres.2018.01.015] [Medline: 29499467]
27. Rey-López JP, Vicente-Rodríguez G, Biosca M, Moreno LA. Sedentary behaviour and obesity development in children and adolescents. Nutr Metab Cardiovasc Dis 2008 Mar;18(3):242-251. [doi: 10.1016/j.numecd.2007.07.008] [Medline: 18083016]
28. Mack I, Bayer C, Schäffeler N, Reiband N, Brölz E, Zurstiege G, et al. Chances and limitations of video games in the fight against childhood obesity-a systematic review. Eur Eat Disord Rev 2017 Jul;25(4):237-267. [doi: 10.1002/erv.2514] [Medline: 28467004]
29. Ameryoun A, Sanaeinasab H, Saffari M, Koenig HG. Impact of game-based health promotion programs on body mass index in overweight/obese children and adolescents: a systematic review and meta-analysis of randomized controlled trials. Child Obes 2018;14(2):67-80. [doi: 10.1089/chi.2017.0250] [Medline: 29185787]
30. 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]
31. Chow CY, Riantiningtyas RR, Kanstrup MB, Papavasileiou M, Liem GD, Olsen A. Can games change children’s eating behaviour? a review of gamification and serious games. Food Qual Prefer 2020 Mar;80:103823. [doi: 10.1016/j.foodqual.2019.103823]

32. Dias JD, Domingues AN, Tibes CM, Zem-Mascarenhas SH, Fonseca LMM. Serious games as an educational strategy to control childhood obesity: a systematic literature review. Rev Lat Am Enfermagem 2018 Sep 03;26:e3036 [FREE Full text] [doi: 10.1590/1518-8345.2509.3036] [Medline: 30183872]
33. Guy S, Ratzki-Leewing A, Gwadry-Sridhar F. Moving beyond the stigma: systematic review of video games and their potential to combat obesity. Int J Hypertens 2011;2011:179124 [FREE Full text] [doi: 10.4061/2011/179124] [Medline: 21629863]
34. Villani D, Carissoli C, Triberti S, Marchetti A, Gilli G, Riva G. Videogames for emotion regulation: a systematic review. Games Health J 2018 Apr;7(2):85-99. [doi: 10.1089/g4h.2017.0108] [Medline: 29424555]
35. Baranowski T, Baranowski J, Thompson D, Buday R, Jago R, Griffith MJ, et al. Video game play, child diet, and physical activity behavior change a randomized clinical trial. Am J Prev Med 2011 Jan;40(1):33-38 [FREE Full text] [doi: 10.1016/j.amepre.2010.09.029] [Medline: 21146765]
36. Majumdar D, Koch PA, Lee H, Contento IR, Islas-Ramos ADL, Fu D. "Creature-101": a serious game to promote energy balance-related behaviors among middle school adolescents. Games Health J 2013 Oct;2(5):280-290 [FREE Full text] [doi: 10.1089/g4h.2013.0045] [Medline: 24761326]
37. Johnson-Glenberg MC, Hekler EB. "Alien Health Game": an embodied exergame to instruct in nutrition and MyPlate. Games Health J 2013 Dec;2(6):354-361. [doi: 10.1089/g4h.2013.0057] [Medline: 26197077]
38. 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]
39. Marchetti D, Fraticelli F, Polcini F, Lato R, Pintaudi B, Nicolucci A, et al. Preventing adolescents' diabesity: design, development, and first evaluation of "Gustavo in Gnam's Planet". Games Health J 2015 Oct;4(5):344-351 [FREE Full text] [doi: 10.1089/g4h.2014.0107] [Medline: 26287925]
40. Sharma SV, Shegog R, Chow J, Finley C, Pomeroy M, Smith C, et al. Effects of the Quest to Lava Mountain computer game on dietary and physical activity behaviors of elementary school children: a pilot group-randomized controlled trial. J Acad Nutr Diet 2015 Aug;115(8):1260-1271. [doi: 10.1016/j.jand.2015.02.022] [Medline: 25912520]
41. Espinosa-Curiel IE, Pozas-Bogarin EE, Lozano-Salas JL, Martínez-Miranda J, Delgado-Pérez EE, Estrada-Zamarron LS. Nutritional education and promotion of healthy eating behaviors among Mexican children through video games: design and pilot test of FoodRateMaster. JMIR Serious Games 2020 Apr 13;8(2):e16431 [FREE Full text] [doi: 10.2196/16431] [Medline: 32281539]
42. Mack I, Reiband N, Etges C, Eichhorn S, Schaeffeler N, Zurstiege G, et al. The Kids Obesity Prevention Program: cluster randomized controlled trial to evaluate a serious game for the prevention and treatment of childhood obesity. J Med Internet Res 2020 Apr 24;22(4):e15725 [FREE Full text] [doi: 10.2196/15725] [Medline: 32329742]
43. Torrente J, Borro-Escribano B, Freire M, del Blanco A, Marchiori EJ, Martinez-Ortiz I, et al. Development of game-like simulations for procedural knowledge in healthcare education. IEEE Trans Learning Technol 2014 Jan;7(1):69-82. [doi: 10.1109/tlt.2013.35]
44. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development. New Jersey, USA: Pearson Education; 2015.
45. Lu AS, Baranowski T, Thompson D, Buday R. Story immersion of videogames for youth health promotion: a review of literature. Games Health J 2012 Jun;1(3):199-204 [FREE Full text] [doi: 10.1089/g4h.2011.0012] [Medline: 24416639]
46. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
47. Danielsson P, Kowalski J, Ekblom Ö, Marcus C. Response of severely obese children and adolescents to behavioral treatment. Arch Pediatr Adolesc Med 2012 Dec;166(12):1103-1108. [doi: 10.1001/2013.jamapediatrics.319] [Medline: 23108856]
48. Kelishadi R, Azizi-Soleiman F. Controlling childhood obesity: a systematic review on strategies and challenges. J Res Med Sci 2014 Oct;19(10):993-1008. [Medline: 25538786]
49. Canavera M, Sharma M, Murnan J. Development and pilot testing a social cognitive theory-based intervention to prevent childhood obesity among elementary students in rural Kentucky. Int Q Community Health Educ 2008;29(1):57-70. [doi: 10.2190/IQ.29.1.e] [Medline: 19342357]
50. Sharma M. Dietary education in school-based childhood obesity prevention programs. Adv Nutr 2011 Mar;2(2):207S-216S [FREE Full text] [doi: 10.3945/an.111.000315] [Medline: 22332053]
51. Kim SY, Prestopnik N, Biocca FA. Body in the interactive game: how interface embodiment affects physical activity and health behavior change. Comput Hum Behav 2014 Jul;36:376-384. [doi: 10.1016/j.chb.2014.03.067]
52. Kuźbicka K, Rachoń D. Bad eating habits as the main cause of obesity among children. Pediatr Endocrinol Diabetes Metab 2013;19(3):106-110. [Medline: 25577898]
53. Nissinen K, Korkalo L, Vepsäläinen H, Mäkiranta P, Koivusilta L, Roos E, et al. Accuracy in the estimation of children's food portion sizes against a food picture book by parents and early educators. J Nutr Sci 2018;7:e35 [FREE Full text] [doi: 10.1017/jns.2018.26] [Medline: 30627432]

54. Phan MH, Keebler JR, Chaparro BS. The development and validation of the game user experience satisfaction scale (GUESS). Hum Factors 2016 Dec;58(8):1217-1247. [doi: 10.1177/0018720816669646] [Medline: 27647156]
55. Holzmann SL, Schäfer H, Groh G, Plecher DA, Klinker G, Schauberger G, et al. Short-term effects of the serious game "Fit, Food, Fun" on nutritional knowledge: a pilot study among children and adolescents. Nutrients 2019 Aug 30;11(9):2031 [FREE Full text] [doi: 10.3390/nu11092031] [Medline: 31480257]
56. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Transl Behav Med 2018 Mar 01;8(2):212-224 [FREE Full text] [doi: 10.1093/tbm/ibx019] [Medline: 29381786]
57. 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]
58. López-Olmedo N, Carriquiry AL, Rodríguez-Ramírez S, Ramírez-Silva I, Espinosa-Montero J, Hernández-Barrera L, et al. Usual intake of added sugars and saturated fats is high while dietary fiber is low in the Mexican population. J Nutr 2016 Sep;146(9):1856S-1865S. [doi: 10.3945/jn.115.218214] [Medline: 27511932]
59. Basdevant A, Boute D, Borys JM. Who should be educated? education strategies: could children educate their parents? Int J Obes Relat Metab Disord 1999 May;23 Suppl 4:S10-2-discussion S12-3. [doi: 10.1038/sj.ijo.0800914] [Medline: 10385274]
60. Coulthard H, Ahmed S. Non taste exposure techniques to increase fruit and vegetable acceptance in children: effects of task and stimulus type. Food Quality and Preference 2017 Oct;61:50-54. [doi: 10.1016/j.foodqual.2017.04.012]
61. Johnson SL. Developmental and environmental influences on young children's vegetable preferences and consumption. Adv Nutr 2016 Jan;7(1):220S-231S [FREE Full text] [doi: 10.3945/an.115.008706] [Medline: 26773030]
62. Wardle J, Herrera M, Cooke L, Gibson EL. Modifying children's food preferences: the effects of exposure and reward on acceptance of an unfamiliar vegetable. Eur J Clin Nutr 2003 Feb;57(2):341-348. [doi: 10.1038/sj.ejcn.1601541] [Medline: 12571670]
63. 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 [FREE Full text] [doi: 10.2196/21813] [Medline: 32940609]
64. Simons M, de Vet E, Hoornstra S, Brug J, Seidell J, Chinapaw M. Adolescents' views on active and non-active videogames: a focus group study. Games Health J 2012 Jun;1(3):211-218. [doi: 10.1089/g4h.2011.0032] [Medline: 26193439]
65. Thompson FE, Subar AF, Coulston AM, Boushey CJ, Ferruzzi MG, Delahanty LM. Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi MG, Delahanty LM, editors. Nutrition in the Prevention and Treatment of Disease (Fourth Edition). London, United Kingdom: Academic Press; 2017:5-48.
66. Kolodziejczyk JK, Merchant G, Norman GJ. Reliability and validity of child/adolescent food frequency questionnaires that assess foods and/or food groups. J Pediatr Gastroenterol Nutr 2012 Jul;55(1):4-13. [doi: 10.1097/MPG.0b013e318251550e] [Medline: 22437477]

Edited by N Zary; submitted 06.09.21; peer-reviewed by I Mack; comments to author 29.09.21; revised version received 14.12.21; accepted 07.04.22; published 06.05.22.

Please cite as:
Espinosa-Curiel IE, Pozas-Bogarin EE, Hernández-Arvizu M, Navarro-Jiménez ME, Delgado-Pérez EE, Martínez-Miranda J, Pérez-Espinosa H
HelperFriend, a Serious Game for Promoting Healthy Lifestyle Behaviors in Children: Design and Pilot Study
JMIR Serious Games 2022;10(2):e33412
URL: https://games.jmir.org/2022/2/e33412
doi:10.2196/33412
PMID:35522474

Original Paper

Virtual Reality Training Using Nintendo Wii Games for Patients With Stroke: Randomized Controlled Trial

Naveed Anwar[1,2*], MS-OMPT; Hossein Karimi[1*], PhD; Ashfaq Ahmad[1*], PhD; Syed Amir Gilani[1], PhD; Kehkshan Khalid[3], MS-SPT; Ahmed Sohaib Aslam[4], MSc; Asif Hanif[1], PhD

[1]University Institute of Physical Therapy, University of Lahore, Lahore, Pakistan

[2]Department of Physical Therapy, Nur International University, Lahore, Pakistan

[3]Department of Physical Therapy, Avicenna Medical College, Lahore, Pakistan

[4]Department of Physical Therapy, Kanaan Healthcare Center, Lahore, Pakistan

[*]these authors contributed equally

Corresponding Author:
Hossein Karimi, PhD
University Institute of Physical Therapy
University of Lahore
1-Km, Defence Road
Bhupatian Chowk, Off Raiwind Road
Lahore
Pakistan
Phone: 92 3336140916
Email: dr.hossein.karimi@gmail.com

Abstract

Background: Stroke is a leading cause of disability. It is difficult to devise an optimal rehabilitation plan once stroke survivors are back home. Conventional rehabilitative therapies are extensively used in patients with stroke to recover motor functioning and disability, but these are arduous and expensive. Virtual reality (VR) video games inspire patients to get involved in their therapeutic exercise routine in a fun way. VR in the form of games provides a fruitful, secure, and challenging learning environment for motor control and neural plasticity development in rehabilitation. The effects of upper limb sensorimotor functioning and balance are the main focus of this trial.

Objective: The aim of this study is to compare the effects of VR training and routine physical therapy on balance and upper extremity sensorimotor function in patients with stroke.

Methods: It was a single assessor-blinded randomized clinical trial. A total of 74 participants with their first chronic stroke were included and rehabilitated in a clinical setting. The lottery method was used to randomly assign patients to either the VR group (n=37) or the routine physical therapy group (n=37). The VR group received a 1-hour session of VR training for 3 weekdays over 6 weeks, and the routine physical therapy group received different stretching and strengthening exercises. The outcome measuring tools were the Berg Balance Scale for balance and the Fugl-Meyer Assessment (upper extremity) scale for sensorimotor, joint pain, and range assessment. The assessment was done at the start of treatment and after the 6 weeks of intervention. Data analysis was done using SPSS 22.

Results: The trial was completed by 68 patients. A significant difference between the two groups was found in the Berg Balance Scale score ($P<.001$), Fugl-Meyer Assessment for motor function ($P=.03$), and Fugl-Meyer Assessment for joint pain and joint range ($P<.001$); however, no significant difference ($P=.19$) in the Fugl-Meyer Assessment for upper extremity sensation was noted.

Conclusions: VR training is helpful for improving balance and function of the upper extremities in the routine life of patients with stroke; although, it was not found to be better than conventional training in improving upper limb sensation. VR training can be a better option in a rehabilitation plan designed to increase functional capability.

Trial Registration: Iranian Registry of Clinical Trials RCT20190715044216N1; https://www.irct.ir/user/trial/40898/view

(JMIR Serious Games 2022;10(2):e29830) doi:10.2196/29830

KEYWORDS

stroke; virtual reality; Fugl-Meyer score; rehabilitation; training; physical therapy; therapy; balance; function; randomized controlled trial

Introduction

Stroke is a leading cause of disability, and it is difficult to devise an optimal rehabilitation plan for patients with stroke once they are discharged from the hospital [1]. Almost 85% of patients with stroke have hemiparesis after stroke while 55% to 75% of stroke survivors have motor dysfunction. South Asian people (people of India, Pakistan, Sri Lanka, Bangladesh, Nepal, and Bhutan) have a higher risk of stroke because of compromised cardiac and metabolic rate [2,3]. Treatment for stroke is initiated with drugs [4], and surgery might be another option to repair any constriction or narrowing of blood vessels [5,6]. Patient rehabilitation is an important part of treatment. The main purpose of rehabilitation is to improve the quality of life for patients with stroke and make them independent [7,8]. Conventional rehabilitative therapies are extensively used to help patients with stroke recover motor functioning and disability. However, the application of conventional techniques is arduous and expensive, and requires transportation of patients to tertiary care hospitals especially in countries like Pakistan where hospitals are less in number. Virtual reality (VR) training in the form of games [9] provides a fruitful, secure, and challenging learning environment for motor control and neural plasticity development after stroke. VR video games inspire patients to get involved in their therapeutic exercise routine in a fun way [10]. Depending on the remodeling and reorganization of brain function, previous studies found that VR can be a great alternative for quick functional recovery after stroke [11]. Mirror neurons in the cortex can be activated through observational learning by VR training. Participants who received sensory input in VR training were also more likely to learn the desired motor behavior [12]. The feedback can help to promote the development of use-dependent cortical plasticity, which could lead to improved motor control. Furthermore, the functional improvement induced by VR training could significantly boost participants' confidence and self-efficacy in a new environment. Moreover, another advantage of VR is that it can greatly save on labor and cost of patients [13].

Stroke is seen to be more prevalent in countries like Pakistan, as the people are more inclined toward using local drugs like naswar, pipe smoking, and beetle leaf chewing (paan). Thus, there is higher incidence of stroke in middle-aged populations (<45 years) [14-16]. The study aimed for a younger population with stroke and found cost-effective treatment protocols at the same time. The unique needs of young people with stroke and the promising opportunity provided by a low-cost serious game would be a beneficial addition in treatment strategy. There is inadequate evidence in the literature to generalize effects on upper limb sensorimotor function and gait through commercial gaming in young patients with stroke. Studies on effectiveness of VR programs in comparison to traditional methods on functional-motor improvement of an upper limb are needed in low-resource countries to reduce cost and time through target-oriented interventions. This study was conducted to compare the effects of VR training and routine physical therapy on balance and function of upper limbs in patients with stroke from the lower- and middle-class populations. This study is conducted to accept or reject the hypothesis that VR has a significantly better effect on balance and upper limb function in patients with stroke.

Methods

Study Design and Participants

This study was a single assessor-blinded randomized clinical trial. Participants were recruited by convenient sampling at Kanaan Physiotherapy & Spine Clinic, Lahore, Pakistan, from September 2018 to December 2020. Diagnosed patients with subacute and chronic stroke were included in the study. The inclusion criteria were patients aged between 40 to 60 years irrespective of gender; unilateral involvement of extremity and the first episode of stroke was either hemorrhagic or ischemic in origin evident by computed tomography scan or magnetic resonance imaging [17]; at least a score of 2 or more on the medical research council scale; and patient is stable, alert, and able to follow the instructions of physical therapists. Patients with ischemic heart disease with unstable angina, history of seizures, Parkinson disease, severe aphasia that can limit participation or feedback, cognitive mental condition that can interfere with comprehension of commands, any systematic disease, BMI, or poor speech were excluded from the study. Written informed consent was obtained from each participant prior to data collection. The participants were randomly assigned to two groups by lottery method: VR (n=37) and routine physical therapy (n=37) groups [16].

Outcome Measures

All values were measured before and after the 6 weeks of intervention. The Berg Balance Scale (BBS) was used to assess balance. It is a 14-item list with each item consisting of a five-point ordinal scale ranging from 0 to 4, with 0 indicating the lowest level of function and 4 the highest level of function. A score of 56 indicates functional balance. A score of <45 indicates individuals may be at greater risk of falling. The Fugl-Meyer Assessment (FMA) tool for upper extremities (UEs) was used to assess sensorimotor function, joint range, and pain. The motor section of the FMA-UE has 33 points that evaluates aspects of movement, reflex, coordination, and speed. Each domain contains multiple items, each scored on a 3-point ordinal scale (0=cannot perform, 1=performs partially, 2=performs fully). Sensation has 6 points, while joint range and pain have 12 points each. Scoring is based on direct observation of performance. FMA-UE is a valid and reliable tool that measures the function of upper limbs, wrists, and hands while the BBS is also a valid and reliable tool to assess balance [18,19].

Interventions

Virtual Reality Training (Group I)

Wii comes with a console, adapter, infrared sensor bar, 2 wireless nunchucks, remote with wrist straps, sensor bar, Wii balance board, and Wii Sports kit. The Wii Sports (tennis and boxing), Wii balance board, and Wii Cooking Mama games were the main games used. Depending upon the participant's ability, the training complexity and intensity were increased by higher levels in the game. The therapist stood behind the participant for protection and support, and the participant was able to grab the handrail if they needed to avoid falling [20]. Patients in this group had a 1-hour session 3 days a week for a period of 6 weeks.

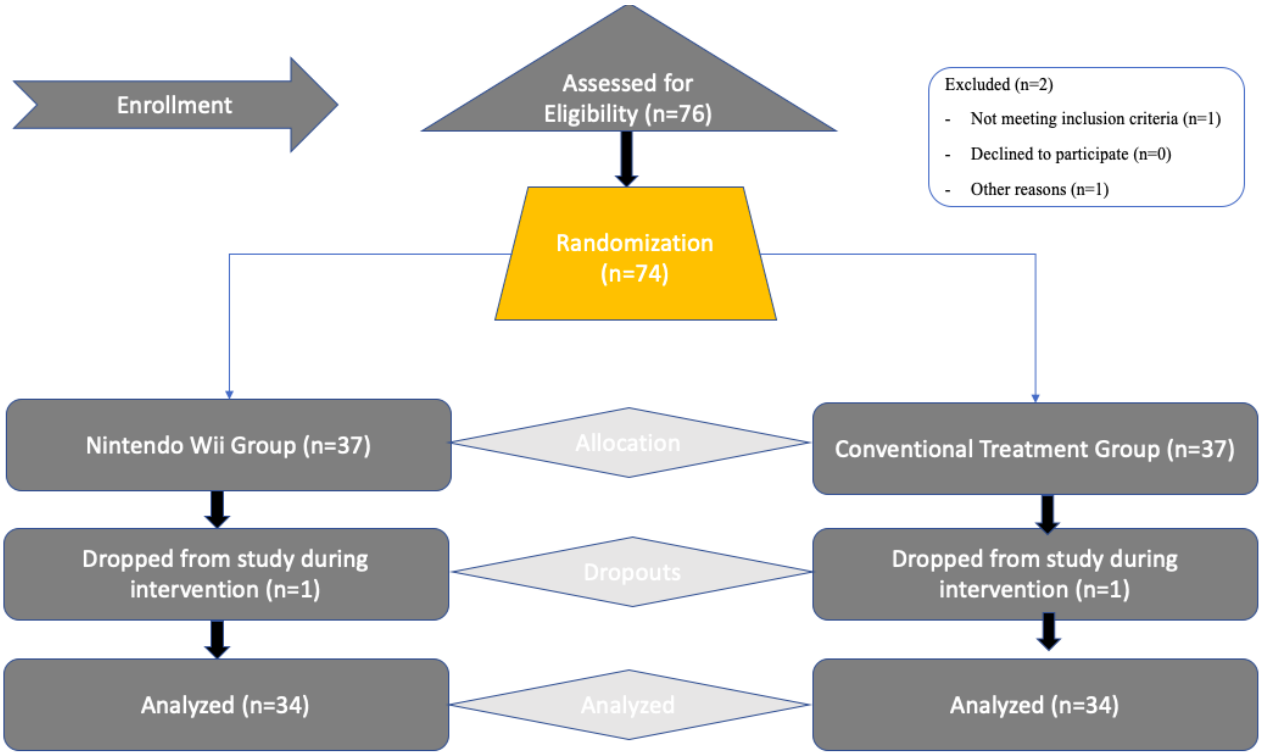
Routine Physical Therapy (Group II)

These include stretching exercises for tight muscles (eg, shoulder, elbow, and wrist flexors). The strengthening program includes exercises for weak extensor muscles and balance training, and coordination exercises to improve motor control and deficit. Each muscle group was targeted for strengthening exercises in upper limbs. Manual resistance was applied and increased according to the patient's condition. The patients in this group got 1-hour sessions, 3 days a week for a period of 6 weeks.

Sample Size

The sample size was calculated by sample size software [21,22]. The sampling technique used was convenient sampling with a statistical power of 80% and ∂=0.5. Recruitment of 76 participants was done with an expected dropout of (10%) during the intervention. One patient was excluded for not meeting the criteria, and 1 left due to personal reasons. During the intervention, 3 patients from each group were lost to follow-up as shown in Figure 1.

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



Ethical Approval

The study protocol was approved by the Ethical Committee of the University of Lahore (approval IRB-UOL-FAHS/373-III/2018).

Randomization

Eligible participants were initially screened by a research assistant. All participants were asked to sign the informed consent. Subsequently, the participants were assessed by a blinded independent assessor at baseline and were randomized. The randomization was done by lottery method. Patients were given the allotted intervention by a trained physical therapist. After 6 weeks of the intervention, another assessment was done by a blinded independent assessor.

Statistical Analysis

SPSS 22.0 (IBM Corp) was used for the statistical analysis. Descriptive statistics were applied for all outcome measures. To check the normality of data, the Shapiro-Wilk test was performed. Data were found to be normally distributed, so parametric tests were applied. An independent *t* test was used to compare data between the groups, while a paired *t* test was used for analysis within the group. Statistical significance was set at $P<.05$.

Results

The data were normally distributed. The mean age of the VR training group was 51.56 (SD 7.199) years, while the mean age for the routine physical therapy group was 51.35 (SD 5.783) years. The mean height of the VR training group was 1.68 (SD 0.11) meters, while the mean height for the routine physical therapy group was 1.69 (SD 0.09) meters. The mean weight of the VR training group was 86.50 (SD 11.41) Kg, while the mean weight for routine physical therapy was 86.35 (SD 11.7) Kg as shown in Table 1.

There was statistically significant differences between posttreatment BBS values of the 2 groups with *P*<.001. The balance increased to a greater extent post treatment for the VR training group with a mean value of 36.62 (SD 7.76) as compared to the routine physical therapy group (mean 26.94, SD 6.46) as shown in Table 2. The mean of the FMA score for motor function in the VR training group was 49.71 (SD 10.03) as compared to 33.47 (SD 11.07) in the routine physical therapy group, which was clinically more significant in the VR group. There was no significant difference in the FMA score for sensation across the 2 groups with *P*=.19. There was significant difference in the FMA for joint pain and joint range across the 2 groups with *P*<.001 as shown in Table 2. A paired sample *t* test was used to compare the values within each treatment group.

Table 1. Demographic characteristics (N=68).

Study groups	Patients, n	Value, mean (SD)
Age (years; virtual reality training)	34	51.56 (7.199)
Age (years; routine physical therapy)	34	51.35 (5.783)
Weight (kg; virtual reality training)	34	86.50 (11.41)
Weight (kg; routine physical therapy)	34	86.35 (11.70)
Height (m; virtual reality training)	34	1.68 (0.11)
Height (m; routine physical therapy)	34	1.69 (0.09)
BMI (kg/m^2; virtual reality training)	34	30.66 (4.24)
BMI (kg/m^2; routine physical therapy)	34	30.60 (4.67)

Table 2. Pre- and postintervention results for virtual reality and routine physical therapy treatment: t test between group statistics (N=68).

Berg Balance Scale and FMA[a] for upper extremity	Treatment group		*P* value
	Virtual reality training, mean (SD)	Routine physical therapy, mean (SD)	
Berg Balance Scale score			<.001
Pre	18.38 (5.19)	19.68 (5.23)	
Post	36.62 (7.76)	26.94 (6.46)	
FMA for motor function			<.001
Pre	19.97 (5.45)	21.03 (7.24)	
Post	49.71 (10.03)	33.47 (11.07)	
FMA for sensation			.19
Pre	7.62 (2.15)	7.24 (2.41)	
Post	8.44 (1.99)	8.44 (2.5)	
FMA for joint range			<.001
Pre	12.82 (4.26)	11.88 (3.88)	
Post	18.18 (3.71)	14.56 (4.13)	
FMA for joint pain			<.001
Pre	12.47 (3.79)	11.65 (3.16)	
Post	19.62 (2.86)	14.62 (4.00)	

[a]FMA: Fugl-Meyer Assessment.

Discussion

Principal Findings

The aim of the study is to compare the effects of two techniques, VR training and routine physical therapy, on balance and function of upper extremities in patients with stroke. Virtual training is a modern technique that has more interest now because of its affordability and that patients can easily use it at home as well. VR training improved balance, pain, range of motion, and motor function of the upper extremities. However, there was no significant effect on sensation.

This study had a significant result in the improvement of the balance during group analysis. Posttreatment balance in the VR group had a mean value of 36.62 (SD 7.76), as compared to the routine physical therapy group (mean 26.94, SD 6.46). Similar results were seen in a study conducted by Jeon et al [23] that suggested that VR in combination with balance training had significant improvement in patients with stroke as compared to the VR alone. Another study conducted by Aramaki et al [24] showed that VR training improved balance and upper limb function, and improved the quality of life in patients with stroke, who also reported that VR is more effective in improving dynamic balance as compared to conventional treatment. Specifically, balance in walking was the most underperformed activity of patients with stroke, so improvement in it would be a substantial help for them. Most of the patients considered playing games as a fun activity in their leisure time. Routine physical therapy was found to be somewhat boring for patients and less attractive [11]. The results of this study will help therapists to use VR for the rehabilitation of patients with stroke. A patient tries to use his muscles, and thus, coordination with the brain improves. Most of the postural muscles are involved in an upright position. On the contrary, routine physical therapy is somewhat passive, and thus improvement is slow. Postural muscles should be targeted in treatment protocols. Sometimes, stroke directly affects the patient's ability to manage the surrounding environment. So, it is difficult to balance yourself when you are unsure about your position in the surroundings [25-27]. This study proved that VR improves functional independence of patients after stroke.

In this study, both groups experienced positive improvements in the FMA on the motor function level. The study's posttreatment results of motor function in patients with stroke had a mean value of 49.71 (SD 10.03) in the VR group, while in the routine physical therapy group, where the mean value was 33.47 (SD 11.07) after treatment, the score was substantially higher than the baseline, suggesting a clinically meaningful increase in the VR training community relative to the routine physical therapy group. A study conducted by Maier et al [28] investigated whether VR training was beneficial for upper limb motor recovery as compared to conventional therapy. A review conducted by Levin and Demers [12] also concluded that VR training has significant results in improving balance and motor function of upper limbs in stroke rehabilitation.

Motor function in upper limbs improves during virtual training because this system provides extra spatial transformation and uncoupled eye-hand movements; it also enhances movement control and creates an entertaining environment, further motivating patients. Repeated motions improve motor learning and patient's functional and anatomical reorganization [29].

The study showed nonsignificant results of sensation by using VR in patients with stroke, the VR group posttreatment mean value was 8.44 (SD 1.99) as compared to the routine physical therapy group (mean 8.44, SD 2.5). A study conducted by Yeh et al [17] showed that the VR system helps promote functional movements and motor control like pinching and grasping activity, with significant results, but our sensory control patients strongly intended to continue it for a long time. Another study conducted by Serrada et al [30] highlights the sensory training, but limited evidence was available on the sensory impairment of patients with stroke. Further evaluation is required to analyze the effectiveness for active sensory training. Vigorous methods and high-class research is needed to measure outcomes of sensory rehabilitation. This largely overlooked topic is an important component for stroke rehabilitation.

This study showed significant results in joint range of motion in the VR group. The VR group had a posttreatment mean value of 18.18 (SD 3.71) as compared to the routine physical therapy group (mean 14.56, SD 4.13). A study conducted by Aşkın et al [31] stated that kinetic VR training improved UE motor function and range of motion in patients with stroke. Another study conducted by Huang et al [32] suggested that VR training improves fine hand movements and active ranges of motion, and promotes coordination, which support the results from our study.

A study conducted by Choi and Paik [33] showed that a VR mobile game–based program for upper limb rehabilitation had significant results with conventional therapy, and it was a good tool for rehabilitation of the UE, indicating that the joint range is achieved in patients with stroke by VR training. Different visual and auditory sounds engage patients in activity, and patients spend more time in performing activities. This creates a desire of interaction, and at the end, a desired joint movement is achieved [33], similar to kinetic VR where shoulder flexion, abduction external rotation, and elbow extension significantly improved by repetitive training more than once a day.

The study results showed significant results in improving pain in the poststroke treatment VR group, which had a mean value of 19.62 (SD 2.86), as compared to routine physical therapy group, which had a mean of 14.62 (SD 4.00). A study conducted by Shahrbanian et al [34] showed that VR training helped in pain management as compared to conventional therapy. A study conducted by Powell and Simmonds [35] suggested that musculoskeletal pain caused activity limitation, but VR training could improve pain and enhance movement speed, which also supports our study results.

Another study conducted in 2014 by Triberti et al [36] stated that VR training improved pain. It is a distraction and analgesic tool, and creates an environment in which the patients immerse themselves in a 3D computer-generated environment. This pleasurable and colorful environment diverts attention from noxious stimuli. It provides a psychological effectiveness, reduces anxiety, and promotes positive emotions so patients feel relaxed and experience less pain [36].

VR has a great neurobiological impact on neuroplasticity that results in an improved volume of gray matter, improved cognitive efficiency, and higher electroencephalographic beta wave concentrations. Innovative brain-computer interfaces assist clinical applications of VR by allowing a direct effect on the electric activity caused by different cortical areas of the brain to ensure efficient control of connected gaming devices. Healthy people may use VR as a storytelling tool to rewrite their own stories as part of an integrative process of self-improvement and personal growth [37,38].

This study has some limitations, as VR training is a comparatively new technique in Pakistan, which is why patients faced difficulties with familiarity despite the fact that it is being used widely for rehabilitation in higher-income countries. Few outcome measures were used in the trial, so we suggest that more outcome measures be used in future trials. This trial did not have long-term follow-up assessments to check continuous effects. These limitations need to be addressed in future studies.

Conclusions

In this study, VR training was an effective way to improve balance, sensorimotor function, joint range, and pain of the upper limbs. Routine physical therapy is beneficial, but VR training can be more target-oriented. This study has a lot of potential in the field of stroke rehabilitation, as it demonstrated that low-cost technologies can offer additional benefits to usual care. Moreover, further randomized controlled trials are required to find the effects of VR in different occupational performances in patients with stroke.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1143 KB - games_v10i2e29830_app1.pdf]

References

1. Charles D, Holmes D, Charles T, McDonough S. Virtual reality design for stroke rehabilitation. In: Rea PM, editor. Biomedical Visualisation: Volume 6. Cham: Springer; 2020:53-87.
2. Singh V, Prabhakaran S, Chaturvedi S, Singhal A, Pandian J. An examination of stroke risk and burden in South Asians. J Stroke Cerebrovasc Dis 2017 Oct;26(10):2145-2153. [doi: 10.1016/j.jstrokecerebrovasdis.2017.04.036] [Medline: 28579510]
3. Saposnik G, Teasell R, Mamdani M, Hall J, McIlroy W, Cheung D, Stroke Outcome Research Canada (SORCan) Working Group. Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle. Stroke 2010 Jul;41(7):1477-1484 [FREE Full text] [doi: 10.1161/STROKEAHA.110.584979] [Medline: 20508185]
4. Williams EI, Betterton RD, Davis TP, Ronaldson PT. Transporter-mediated delivery of small molecule drugs to the brain: a critical mechanism that can advance therapeutic development for ischemic stroke. Pharmaceutics 2020 Feb 14;12(2):154 [FREE Full text] [doi: 10.3390/pharmaceutics12020154] [Medline: 32075088]
5. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251.e1 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
6. Patel BM, Reinert NJ, Al-Robaidi K, Gao X, Fabio A, Esper SA, et al. Independent predictors of perioperative stroke-related mortality after cardiac surgery. J Stroke Cerebrovasc Dis 2020 May;29(5):104711. [doi: 10.1016/j.jstrokecerebrovasdis.2020.104711] [Medline: 32184023]
7. Long Y, Ouyang R, Zhang J. Effects of virtual reality training on occupational performance and self-efficacy of patients with stroke: a randomized controlled trial. J Neuroeng Rehabil 2020 Nov 13;17(1):150 [FREE Full text] [doi: 10.1186/s12984-020-00783-2] [Medline: 33187532]
8. Stinear CM, Lang CE, Zeiler S, Byblow WD. Advances and challenges in stroke rehabilitation. Lancet Neurol 2020 Apr;19(4):348-360. [doi: 10.1016/S1474-4422(19)30415-6] [Medline: 32004440]
9. Cortes M, Elder J, Rykman A, Murray L, Avedissian M, Stampas A, et al. Improved motor performance in chronic spinal cord injury following upper-limb robotic training. NeuroRehabilitation 2013;33(1):57-65. [doi: 10.3233/NRE-130928] [Medline: 23949034]
10. Van der Loos M. Virtual reality video games make rehab fun for younger patients. Vancouver Coastal Health Research Institute. 2019 Sep 23. URL: https://www.vchri.ca/stories/2019/09/23/virtual-reality-video-games-make-rehab-fun-younger-patients [accessed 2022-06-08]
11. Aramaki AL, Sampaio RF, Cavalcanti A, Dutra FCMSE. Use of client-centered virtual reality in rehabilitation after stroke: a feasibility study. Arq Neuropsiquiatr 2019;77(9):622-631 [FREE Full text] [doi: 10.1590/0004-282X20190103] [Medline: 31553392]
12. Levin MF, Demers M. Motor learning in neurological rehabilitation. Disabil Rehabil 2021 Dec;43(24):3445-3453. [doi: 10.1080/09638288.2020.1752317] [Medline: 32320305]
13. Demers M, Fung K, Subramanian SK, Lemay M, Robert MT. Integration of motor learning principles into virtual reality interventions for individuals with cerebral palsy: systematic review. JMIR Serious Games 2021 Apr 07;9(2):e23822 [FREE Full text] [doi: 10.2196/23822] [Medline: 33825690]

14. Nomani AZ, Iqbal M, Jamil U, Nabi S, Mughal S, Badshah M, et al. Etiology of stroke in young pakistani adults; results of a single center study. Pakistan J Neurol Sci 2015;10(4):18-22.
15. Hashmi M, Khan M, Wasay M. Growing burden of stroke in Pakistan: a review of progress and limitations. Int J Stroke 2013 Oct;8(7):575-581. [doi: 10.1111/j.1747-4949.2012.00827.x] [Medline: 22759392]
16. Anwar N, Karimi H, Ahmad A, Mumtaz N, Saqulain G, Gilani SA. A novel virtual reality training strategy for poststroke patients: a randomized clinical trial. J Healthc Eng 2021;2021:6598726. [doi: 10.1155/2021/6598726] [Medline: 34840701]
17. Yeh S, Lee S, Chan R, Wu Y, Zheng L, Flynn S. The efficacy of a haptic-enhanced virtual reality system for precision grasp acquisition in stroke rehabilitation. J Healthc Eng 2017;2017:9840273. [doi: 10.1155/2017/9840273] [Medline: 29230275]
18. Gladstone DJ, Danells CJ, Black SE. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair 2002 Sep;16(3):232-240. [doi: 10.1177/154596802401105171] [Medline: 12234086]
19. Downs S, Marquez J, Chiarelli P. The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review. J Physiother 2013 Jun;59(2):93-99 [FREE Full text] [doi: 10.1016/S1836-9553(13)70161-9] [Medline: 23663794]
20. Holden MK, Dyar T. Virtual environment training: a new tool for neurorehabilitation. Neurol Rep 2002;26(2):62-71. [doi: 10.1097/01253086-200226020-00003]
21. Sample size for before-after study (Paired T-test). Sample Size Calculators. URL: http://www.sample-size.net/sample-size-study-paired-t-test/ [accessed 2017-07-18]
22. Yoon J, Chun MH, Lee SJ, Kim BR. Effect of virtual reality-based rehabilitation on upper-extremity function in patients with brain tumor: controlled trial. Am J Phys Med Rehabil 2015 Jun;94(6):449-459. [doi: 10.1097/PHM.0000000000000192] [Medline: 25251249]
23. Jeon M, Moon J, Cho H. Effects of virtual reality combined with balance training on upper limb function, balance, and activities of daily living in persons with acute stroke: a preliminary study. Phys Ther Rehabil Sci 2019 Dec 30;8(4):187-193. [doi: 10.14474/ptrs.2019.8.4.187]
24. Aramaki AL, Sampaio RF, Reis ACS, Cavalcanti A, Dutra FCMSE. Virtual reality in the rehabilitation of patients with stroke: an integrative review. Arq Neuropsiquiatr 2019 May 13;77(4):268-278 [FREE Full text] [doi: 10.1590/0004-282X20190025] [Medline: 31090808]
25. Lee HS, Park YJ, Park SW. The effects of virtual reality training on function in chronic stroke patients: a systematic review and meta-analysis. Biomed Res Int 2019;2019:7595639. [doi: 10.1155/2019/7595639] [Medline: 31317037]
26. Noveletto F, Soares AV, Mello BA, Sevegnani CN, Eichinger FLF, Hounsell MDS, et al. Biomedical serious game system for balance rehabilitation of hemiparetic stroke patients. IEEE Trans Neural Syst Rehabil Eng 2018 Nov;26(11):2179-2188. [doi: 10.1109/TNSRE.2018.2876670] [Medline: 30334802]
27. Kim N, Park Y, Lee B. Effects of community-based virtual reality treadmill training on balance ability in patients with chronic stroke. J Phys Ther Sci 2015 Mar;27(3):655-658 [FREE Full text] [doi: 10.1589/jpts.27.655] [Medline: 25931701]
28. Maier M, Rubio Ballester B, Duff A, Duarte Oller E, Verschure PFMJ. Effect of specific over nonspecific VR-based rehabilitation on poststroke motor recovery: a systematic meta-analysis. Neurorehabil Neural Repair 2019 Feb;33(2):112-129 [FREE Full text] [doi: 10.1177/1545968318820169] [Medline: 30700224]
29. Huang Q, Wu W, Chen X, Wu B, Wu L, Huang X, et al. Evaluating the effect and mechanism of upper limb motor function recovery induced by immersive virtual-reality-based rehabilitation for subacute stroke subjects: study protocol for a randomized controlled trial. Trials 2019 Feb 06;20(1):104 [FREE Full text] [doi: 10.1186/s13063-019-3177-y] [Medline: 30728055]
30. Serrada I, Hordacre B, Hillier SL. Does sensory retraining improve sensation and sensorimotor function following stroke: a systematic review and meta-analysis. Front Neurosci 2019;13:402. [doi: 10.3389/fnins.2019.00402] [Medline: 31114472]
31. Aşkın A, Atar E, Koçyiğit H, Tosun A. Effects of Kinect-based virtual reality game training on upper extremity motor recovery in chronic stroke. Somatosens Mot Res 2018 Mar;35(1):25-32. [doi: 10.1080/08990220.2018.1444599] [Medline: 29529919]
32. Huang X, Naghdy F, Naghdy G, Du H, Todd C. The combined effects of adaptive control and virtual reality on robot-assisted fine hand motion rehabilitation in chronic stroke patients: a case study. J Stroke Cerebrovasc Dis 2018 Jan;27(1):221-228. [doi: 10.1016/j.jstrokecerebrovasdis.2017.08.027] [Medline: 28919312]
33. Choi Y, Paik N. Mobile game-based virtual reality program for upper extremity stroke rehabilitation. J Vis Exp 2018 Mar 08;133:56241 [FREE Full text] [doi: 10.3791/56241] [Medline: 29578520]
34. Shahrbanian S, Ma X, Aghaei N, Korner-Bitensky N, Moshiri K, Simmonds MJ. Use of virtual reality (immersive vs. non immersive) for pain management in children and adults: a systematic review of evidence from randomized controlled trials. Eur J Exp Biol 2012:1408-1422.
35. Powell W, Simmonds MJ. Virtual reality and musculoskeletal pain: manipulating sensory cues to improve motor performance during walking. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):390-396. [doi: 10.1089/cyber.2014.0061] [Medline: 24892203]

36. 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]
37. Georgiev D, Georgieva I, Gong Z, Nanjappan V, Georgiev G. Virtual reality for neurorehabilitation and cognitive enhancement. Brain Sci 2021 Feb 11;11(2):221 [FREE Full text] [doi: 10.3390/brainsci11020221] [Medline: 33670277]
38. Pintado-Izquierdo S, Cano-de-la-Cuerda R, Ortiz-Gutiérrez RM. Video game-based therapy on balance and gait of patients with stroke: a systematic review. Appl Sci 2020 Sep 15;10(18):6426. [doi: 10.3390/app10186426]

Abbreviations

BBS: Berg Balance Scale
FMA: Fugl-Meyer Assessment
UE: upper extremity
VR: virtual reality

Edited by N Zary; submitted 26.04.21; peer-reviewed by M Bashir, J Zhang, M Demers; comments to author 17.05.21; revised version received 30.06.21; accepted 13.05.22; published 13.06.22.

Please cite as:
Anwar N, Karimi H, Ahmad A, Gilani SA, Khalid K, Aslam AS, Hanif A
Virtual Reality Training Using Nintendo Wii Games for Patients With Stroke: Randomized Controlled Trial
JMIR Serious Games 2022;10(2):e29830
URL: https://games.jmir.org/2022/2/e29830
doi:10.2196/29830
PMID:35699989

Original Paper

The Role of Agency and Threat Immediacy in Interactive Digital Narrative Fear Appeals for the Prevention of Excessive Alcohol Use: Randomized Controlled Trial

Hendrik Engelbrecht[1], MSc; Laura Nynke van der Laan[1], Ir, PhD; Renske van Enschot[1], PhD; Emiel Krahmer[1], PhD

Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, Netherlands

Corresponding Author:
Hendrik Engelbrecht, MSc
Tilburg School of Humanities and Digital Sciences
Tilburg University
Warandelaan 2
Tilburg, 5037AB
Netherlands
Phone: 31 134663383
Email: h.engelbrecht@tilburguniversity.edu

Abstract

Background: Serious games for the training of prevention behaviors have been widely recognized as potentially valuable tools for adolescents and young adults across a variety of risk behaviors. However, the role of agency as a distinguishing factor from traditional health interventions has seldom been isolated and grounded in the persuasive health communication theory. Fear appeals have different effects on intentions to perform prevention behaviors depending on the immediacy of the consequences. Looking into how to increase self-efficacy beliefs for health behavior with distant consequences is the first step toward improving game-based interventions for adverse health outcomes.

Objective: This study aimed to investigate the effect of agency on self-efficacy and the intention to drink less alcohol in an interactive digital narrative fear appeal. Furthermore, the communicated immediacy of threat outcomes was evaluated as a potential moderator of the effect of agency on self-efficacy.

Methods: A web-based experimental study was conducted with university students (N=178). The participants were presented with a fear appeal outlining the consequences of excessive alcohol use in a fully automated web-based interactive narrative. Participants either had perceived control over the outcome of the narrative scenario (high agency) or no control over the outcome (low agency). The threat was either framed as a short-term (high immediacy) or long-term (low immediacy) negative health outcome resulting from the execution of the risk behavior (drinking too much alcohol).

Results: A total of 123 valid cases were analyzed. Self-efficacy and intention to limit alcohol intake were not influenced by the agency manipulation. Self-efficacy was shown to be a significant predictor of behavioral intention. The immediacy of the threat did not moderate the relationship between agency and self-efficacy.

Conclusions: Although agency manipulation was successful, we could not find evidence of an effect of agency or threat immediacy on self-efficacy. The implications for different operationalizations of different agency concepts, as well as the malleability of self-efficacy beliefs for long-term threats, are discussed. The use of repeated versus single interventions and different threat types (eg, health and social threats) should be tested empirically to establish a way forward for diversifying intervention approaches.

Trial Registration: ClinicalTrials.gov NCT05321238; https://www.clinicaltrials.gov/ct2/show/NCT05321238

(JMIR Serious Games 2022;10(2):e32218) doi:10.2196/32218

KEYWORDS

young adults; college students; alcohol abuse; drinking; EPPM; fear appeals; agency; serious games

Introduction

Background

Overview

Excessive alcohol abuse is an alarmingly large cause of death in European Union member states, as well as in Norway, Switzerland, and the United Kingdom [1]. This is especially the case among young adults, with 1 in 4 deaths between the ages of 20 and 24 years occurring because of alcohol-related illness or injury. College students, who make up a large part of this age group, are often confronted with the social facilitation of alcohol use, leading to problematic drinking behaviors [2]. Given the prevalence and risk of excessive alcohol use, prevention-focused approaches should receive more attention to educate and train young adults to moderate their drinking. Fostering the execution of preventive behaviors demands the development of interventions that are effective in eliciting the desired change in behavior to avert short-term (eg, alcohol poisoning) and long-term adverse health outcomes (eg, liver disease) of excessive alcohol use.

We argue that the use of interactive digital narratives (IDNs) could be a promising tool to enable the training of preventive behavior through direct interaction with the narrative as a protagonist. In contrast to traditional passive narratives, an IDN involves computer-based interactive storytelling that allows the user to intentionally influence a nonlinear narrative. This means that the interactor is not only experiencing fictional reality but can also be part of it by being able to make meaningful decisions that affect narrative outcomes [3]. What sets serious games apart from more traditional passive forms of media is the opportunity to interact with and affect what happens in the game, making it important to establish the effect of interaction on outcomes that are relevant to serious game interventions. This study aimed to connect the theoretical construct of agency of interactors as the perceived meaningfulness of interaction with self-efficacy beliefs about health behaviors for immediate and long-term health threats.

Fear Appeals

Fear appeals have been widely assessed and used by many health communication researchers for interventions that focus on behavioral changes. Fear appeal communication aims to stimulate behavior change through the provision of a threat that causes a fear response. The message receiver is assumed to be motivated to resolve the arousal caused by the fear response and is presented with a solution to avert the threat. If the appeal is successful, the message perceiver adopts the prevention behavior. The term *prevention behavior* is used to describe the behavior needed to avert a health threat. For example, in the case of messages promoting vaccinations to prevent an infection with a disease, receiving the vaccination constitutes prevention behavior.

Fear appeals are commonly used in many public health campaigns, making use of mostly text-based or pictorial stimuli to elicit fear responses. One of the most widely applied, as well as well-researched, theories on the underlying process of fear appeal communication is the extended parallel process model (EPPM) [4]. The core assumption of the EPPM is the processing of threat information as either eliciting a fear control process (ie, the mitigation of threat-related processing through reactance) or a danger control process, which results in the adoption of threat prevention behavior. A meta-analysis on the effectiveness of EPPM-based fear appeals in the general health domain showed positive effects on behavior (as well as attitudes and intentions), with this effect being increased by the use of messages perceived as high in severity and susceptibility [5]. In other words, the success of the intervention is determined by first appraising the threat as serious and personally relevant (severity and susceptibility) and, second, seeing the prevention behavior as effective in averting the threat (response efficacy) and believing that one is capable of executing the prevention behavior (self-efficacy). The perception of self-efficacy has been shown to be a core determinant of fear appeal effectiveness in influencing the execution of prevention behaviors when severity and susceptibility are high [6].

Witte [4] posited that high *self-efficacy* means that a person needs to believe that he or she has the ability to execute a response that is recommended to him or her in the context of a fear appeal (ie, preventive health behavior). According to the EPPM, perceived *efficacy* (*self-* and *response efficacy*) communicated in a fear appeal determines whether a fear message is accepted. It is important to note that the evidence for the abovementioned effects does not rely solely on studies dealing with the health threat of alcohol exclusively but rather encompasses studies across a variety of risk behaviors such as smoking, drug use, and risky driving. Previous work supports the notion that fear appeals, specifically those addressing alcohol use among young adults, can raise efficacy and affect intention to drink alcohol (see the studies of Stainback and Rogers [7] and Moscato et al [8]); however, work in this domain remains scarce. Moreover, studies on the assessment of interactive narrative interventions in this area are nonexistent.

Narrative Appeals

Fear appeals used in public campaigns often make use of simple narratives through pictures and textual information that are aimed at persuading the message receiver by providing a way of averting the threat (ie, increasing efficacy beliefs). These simple narratives (eg, a picture of a man losing his leg because of smoking) heighten the relevant affective responses for persuasion (ie, fear and compassion), leading to more persuasive fear appeals [9]. Although these simple narratives are seen as more personally relevant than pure factual information, they are limited by the brevity of the narrative content. Although concise messages and pictorial stimuli engage narrative processing of the content [10], for more elaborate narrative appeals, the difference lies in the elaboration of the threat and solution. The additional context of narrative content strengthens transportation and identification through a heightened loss of self-awareness [11]. Evidence for this can be found in many studies showing narrative persuasion through heightened identification with the protagonist and transportation into the story world [12-14]. The term *story world* describes the structure surrounding the events of the story, which includes the locations and persons relevant to the events taking place [15]. Transportation into the story world plays an important part in narrative persuasion as higher transportation can result in affective and cognitive responses that are consistent with what the protagonist in the story world experiences [16].

In their review of narrative interventions, De Graaf et al [17] concluded that further research on how content is embedded into the narrative is needed. More specifically, they pointed to the integration of narratives into causal structures as an important way forward for health narratives. De Graaf et al [17] based this conclusion on the work by Dahlstrom [18], who showed that the integration of persuasive appeals as having an effect on character actions leads to lower resistance to persuasion. Although the aforementioned study was not conducted in the health domain, we can assume a possible transfer of this *causality effect* to interactive game-like formats. Although a passive narrative appeal can affect character intentions to enhance persuasion, an interactive appeal could potentially increase this effect by making the message receiver an active participant in the behavior addressed by the appeal. Earlier research clearly points to evidence supporting the effectiveness of narrative appeals, but there is little work on the potential influence of providing the message receiver with agency in interactive narrative appeals that strengthen the causal connection between prevention behavior and threat aversion.

The Role of Agency for Self-efficacy Beliefs

The term *agency* has been widely used for different forms of interaction with games and IDNs. In IDN research, agency is most commonly conceptualized as the perceived meaningfulness of action with regard to narrative outcomes. As argued by Murray [19] and other researchers in the field, control over narrative progression enables higher affective engagement, which is important for effective persuasion in narrative appeals [20]. According to Murray [19], agency is the “power to take meaningful actions and see the results of our decisions and actions.” This definition was extended by Tanenbaum and Tanenbaum [21]. They described the need for players to develop *competence* in how their actions affect the story world, and this need can be fulfilled through the provision of agency. This competence in learning that player actions have consequences (also called *procedural literacy*) on the story world is a necessity for meaningful interaction (ie, the operationalization of *agency*). For example, the ingestion of certain foods within the game world results in an increase in the (visible) weight of the main character.

Many investigations have provided interactive elements in their message design (see the studies by Winskell et al [22] and Carvalho et al [23]) but do not contrast interactive to passive appeals directly, making generalizations about the impact of perceived agency difficult. Only a few studies have compared interactive with noninteractive fear appeals, and these studies found contradictory results. The study by Panic et al [24] on children suggested that the interactive game used for the appeal distracted participants, thereby preventing the processing of the appeal content. In contrast, Kim et al [25] found that a game-based fear appeal resulted in stronger intentions to quit adverse behavior than a brochure. However, it has to be noted that these findings are hard to ground in assumptions of interaction alone, as the difference in the format of the appeal introduces more differences between conditions than interaction alone.

Although seldom investigated experimentally, operationalizing agency through meaningful interaction seems promising in how this could relate to self-efficacy for threat prevention behaviors. Through the experience of agency, prevention behaviors can potentially be trained, and self-efficacy beliefs can be strengthened. Agency itself is only concerned with the experience of meaningful interactions within the narrative, whereas self-efficacy is concerned with the transfer of beliefs about behaviors in the real world. Self-efficacy is defined as the belief of an individual that they can execute a behavior recommended by a fear appeal message [4]. In existing work, self-efficacy is sometimes used interchangeably with efficacy beliefs in the game itself, where the experience of meaningful interactions within the game is assumed to be related to heightened self-efficacy beliefs for this in-game behavior for the remainder of the game [26]. However, for this study, our focus is not on in-game efficacy but on the transfer of experienced agency to real-life self-efficacy beliefs.

Real-life self-efficacy beliefs are amenable through experimental manipulation in persuasive communication [6]; however, no previous study has investigated whether an interactive fear appeal that is high in agency increases real-life self-efficacy for the behavior executed in the game, as opposed to a passive appeal. The relationship between in-game agency and self-efficacy needs to be further grounded in the assumptions of social learning theory [27] to explain the potential transfer from in-game behavior to real-life efficacy beliefs. The social learning theory posits self-efficacy as an outcome of social learning processes. By using narrative experiences, behavior can be modeled through the observation of role models, which increases the salience of attitude formation in real-life contexts and increases self-efficacy on the part of the person interacting with the narrative [28]. Although observation alone contributes to *self-efficacy* beliefs, the effect is deemed greater for the execution of the target behavior by the person himself or herself. The interaction in an IDN can be hypothesized to increase *self-efficacy* through the development of competence in prevention behavior (as rooted in perceived *agency*).

In line with this, in this study, we expect that self-efficacy beliefs about real-life prevention behaviors can be affected by the provision of agency in an interactive narrative fear appeal. Although it is only a stepping stone in examining the process of persuasive IDNs for behavior change, the role of interaction as a key differentiator in the experience between passive and active media has rarely been investigated in a highly controlled narrative environment.

The Issue of Distant Threats

Although fear appeals for immediate threats have been shown to be effective, there is less evidence showing EPPM’s effectiveness for temporally distant threats [5]. For example, although smoking in adolescents has an impact on future health (*low immediacy*), interventions informing about immediate negative outcomes (*high immediacy)*, such as bad breath, have been shown to be more effective in evoking behavior change [29]. This poses the question of how to design interventions that try to elicit prevention behaviors for threats that do not have immediate consequences or where immediate consequences are perceived as unlikely to occur.

As discussed by Klimmt and Hartmann [26], players are causal agents in IDNs, who derive their engagement with the game by experiencing *efficacy* from successfully conducting actions. The *immediacy* of a response causes temporal congruency between the cause and effect between player actions and game events. In the case of a cause-effect chain with high *immediacy*, there is little question on the part of the player with regard to their agency and, therefore, their self-efficacy to affect the story world. In terms of integrating a distant threat into an IDN, the narrative format allows the author to advance time in the story world, which opens up the opportunity to present players with an immediate causal connection between executed behavior and long-term consequences. Using this temporal flexibility of a narrative format, we hypothesized that the effect of agency on self-efficacy is moderated by the framing of the threat as either short-term (high immediacy) or long-term (low immediacy) adverse health outcomes.

Goals of This Study

This study aimed to assess the effect of agency and threat immediacy on self-efficacy and, consequently, the intention to perform prevention behaviors. The following hypotheses were tested using an IDN in this study:

1. Hypothesis 1: Higher agency in the narrative progression of the IDN fear appeal results in higher perceived self-efficacy for prevention behavior.
2. Hypothesis 2: The effect of agency on self-efficacy is moderated by framing the threat as an adverse health outcome with either high or low immediacy.
3. Hypothesis 3: Higher perceived self-efficacy will lead to a higher intention to perform the target behavior.

Methods

Design

Data were collected from November 26 to December 18, 2020. The hypotheses, study design, and planned analyses were preregistered using the preregistration platform provided by the Wharton Credibility Lab. The independent variables were agency (low or high) and threat immediacy (low or high). Self-efficacy and behavioral intention served as the dependent variables in this study.

Ethics Approval

This study was approved by the Research Ethics and Data Management Committee (REDC) of the Tilburg School of Humanities and Digital Sciences (TSHD) and was conducted as a 2×2 between-subjects experimental study (reference number: REDC 2020.141).

Participants

A total of 178 participants were recruited from the Human Subject Pool of Tilburg University. All participants were enrolled in a master’s or bachelor’s degree program and received 0.5 course credits for participation in the study. Participants were allocated equally across all 4 conditions, and the study was conducted entirely on the web. No personally identifiable data (including IP addresses) were collected to ensure the anonymity of all participants.

Data were collected in 2 phases. The first half of the participants (95/178, 53.4%) who enrolled in the study were allocated to the high-agency condition. After the completion of data collection for high-agency participants, the second half (83/178, 46.6%) of the participants were assigned to the low-agency condition. The low-agency narratives were matched in terms of narrative content to the narratives of the high-agency group who took part in the study before them. All participants were randomly assigned to either the high- or low-immediacy conditions. Given the short time frame of data collection and the absence of events that could have an influence on drinking-related behavior (eg, public events), we assumed the time gap between the participation of the first and second halves of the participants to not affect our measures. Sufficient computer literacy and English language capabilities were assumed among the student population used in this study.

Participants were included in the analysis if they (1) were aged >18 years, (2) spent >480 seconds completing the survey, (3) passed both attention checks relating to narrative content, (4) correctly identified their experimental condition in the manipulation check, (5) did not choose to *drink an alcoholic drink* for all decisions in the interactive narrative, and (6) did not show signs of alcohol abuse. Participants were excluded if they were aged <18 years (2/178, 1.1%) or completed the questionnaire in an unreasonably short amount of time (14/178, 7.9%). In addition, participants were excluded if they failed 1 of the 2 attention checks, indicating that the narrative was not read attentively (21/178, 11.8%), or they failed the manipulation check asking them to recall whether they were presented with short- or long-term (high or low immediacy) consequences of excessive drinking (9/178, 5.1%). All participants who chose an alcoholic drink for all decision points in the narrative were excluded (3/178, 1.7%). Choosing to drink an alcoholic drink for all presented decisions would lead to the participants not executing the prevention behavior (choosing not to drink alcohol); therefore, they were also not presented with a self-efficacy statement. This means that participants would not be exposed to a full fear appeal, as they only received a severity, susceptibility, and response efficacy statement. As elaborated earlier, according to the EPPM, without self-efficacy, a fear appeal is unlikely to lead to the adoption of the proposed prevention behavior. Hence, 1.7% (3/178) of participants had to be excluded as they chose to drink alcoholic drinks for all the presented decisions. Finally, 3.9% (7/178) of participants were excluded because of signs of alcohol dependence as they scored >15 on the Brief Young Adult Alcohol Consequences Questionnaire (B-YAACQ) [36]. As the inclusion of participants with dependence would have skewed the results for the self-efficacy measure, it was decided to exclude them so that the results were more representative of a prevention-focused intervention applied to a healthy population.

The final sample used for analysis comprised 123 participants, with 59 (48%) in the high-immediacy condition and 64 (52%) in the low-immediacy condition. The split between high and low agency was almost equal to 49.6% (61/123) of participants in the high-agency group and 50.4% (62/123) of participants in the low-agency group.

The sample was representative of a student sample with a mean age of 21.4 (SD 3.086) years. Most held either a high school (53/123, 43.1%) or bachelor's diploma (55/123, 55.3%), with few (2/123, 1.6%) participants having completed their master's degree. There was a large gender imbalance in the sample, with approximately 72% (88/123) of the sample being female.

Procedure

The entirety of this study took place on the web. Participants signed up through the Tilburg University Human Subject Pool. They were first presented with an information letter and then informed consent before agreeing to participate in the study. Next, demographic data (age, gender, and education) were collected, and participants were instructed to *imagine being in the place of the protagonist* for the entirety of the narrative. On the basis of their condition, they were presented with 1 of the 4 IDNs. After completion of the narrative, participants were asked to fill out several questionnaires. First, the 2 dependent measures—self-efficacy and behavioral intention—and the perceived fear measure were presented. Next, the participants had to fill in the attention check by answering 3 questions concerning the content of the narrative. This was followed by manipulation checks. Following this, severity, susceptibility, response efficacy, and perceived agency were measured. Finally, questions regarding disinhibition, perceived control over drinking, and frequency of drinking too much alcohol, as well as the alcohol dependence measure, were presented. After completing the questionnaires, participants were debriefed and offered to watch a video to potentially restore any adverse effects resulting from the study.

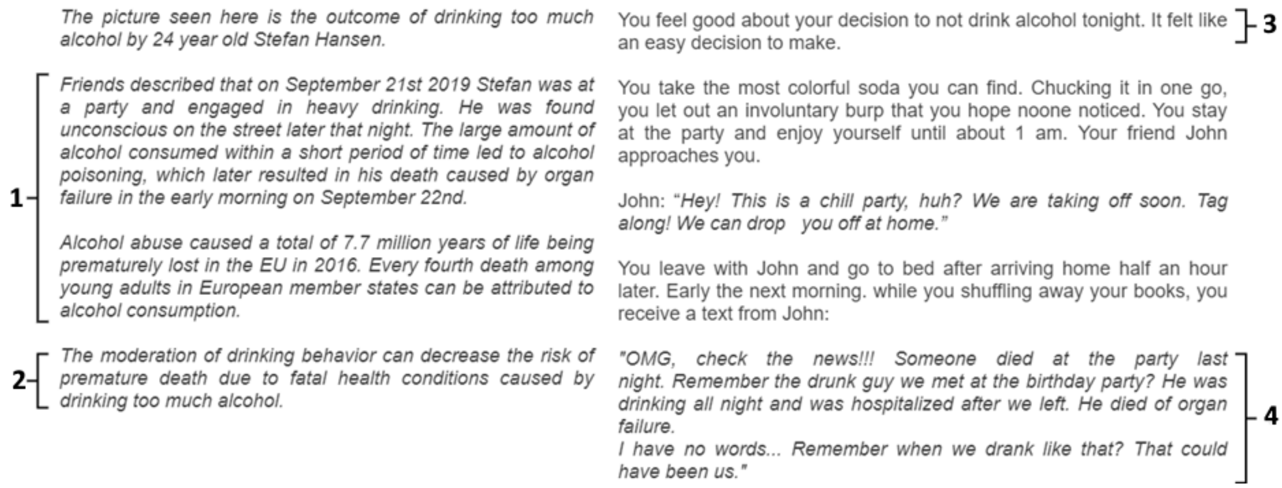
Stimulus Material

Overview

The narratives showed a fictional scenario in which the participant was presented with a fear appeal message, together with a pictorial fear stimulus (including severity, susceptibility, and response efficacy statements). Pictures were added throughout the story to (1) improve transportation into the story world, which has been shown to improve narrative persuasion [12], and (2) strengthen the fear response when participants are presented with the threat during the fear appeal message presentation [31].

The first part of the fear appeal message (Figure 1, left) was embedded into the narrative as a message that is displayed while the protagonist watches a video on the web during breakfast. The narrative subsequently presents participants with the risk situation—a house party—where the risk behavior—drinking alcohol—and, hence, the opportunity to perform the prevention behavior (declining a drink) is likely to be encountered. After completing the narrative scenario of the party, the protagonist of the story arrives at home. The last page of the narrative details the death of the protagonist or another person from organ failure, depending on whether the prevention behavior was executed by the participant (Figure 1, right). This event was described as having occurred either immediately after the party (high immediacy) or several months later (low immediacy). The entire narrative, as well as the decisions, were implemented in the survey using the survey platform Qualtrics.

Figure 1. Excerpts from the narrative fear appeal comprising (1) severity and susceptibility, (2) response efficacy, (3) self-efficacy, and (4) threat avoidance messages in the high-immediacy condition.


The picture seen here is the outcome of drinking too much alcohol by 24 year old Stefan Hansen.

1 *Friends described that on September 21st 2019 Stefan was at a party and engaged in heavy drinking. He was found unconscious on the street later that night. The large amount of alcohol consumed within a short period of time led to alcohol poisoning, which later resulted in his death caused by organ failure in the early morning on September 22nd.*

Alcohol abuse caused a total of 7.7 million years of life being prematurely lost in the EU in 2016. Every fourth death among young adults in European member states can be attributed to alcohol consumption.

2 *The moderation of drinking behavior can decrease the risk of premature death due to fatal health conditions caused by drinking too much alcohol.*

3 You feel good about your decision to not drink alcohol tonight. It felt like an easy decision to make.

You take the most colorful soda you can find. Chucking it in one go, you let out an involuntary burp that you hope noone noticed. You stay at the party and enjoy yourself until about 1 am. Your friend John approaches you.

John: "*Hey! This is a chill party, huh? We are taking off soon. Tag along! We can drop you off at home.*"

You leave with John and go to bed after arriving home half an hour later. Early the next morning. while you shuffling away your books, you receive a text from John:

4 *"OMG, check the news!!! Someone died at the party last night. Remember the drunk guy we met at the birthday party? He was drinking all night and was hospitalized after we left. He died of organ failure.*
I have no words... Remember when we drank like that? That could have been us."

The scenario of a house party was chosen as it is a familiar example to our target population and should, therefore, be easy to imagine and connect to real-life risk situations (to enable behavior transfer). Choosing organ failure as a threat made it possible to keep the threat congruent across different immediacy conditions while only changing the time frame. The threat needed to be severe to cause sufficiently high levels of perceived fear, whereas the pictorial stimuli needed to be congruent with the threat without invoking disgust. Disgust has been shown to potentially undermine appeal content when combined with fear stimuli [32]. The pictorial stimulus was taken from the Set of Fear Inducing Pictures developed by Michalowski et al [33] (picture index: *blood_60*) as part of the Nencki Affective Picture System [34]. This picture was chosen as it hides obvious features of the person and to avoid the influence of differences in identification with the patient and the moderate scores on valence and arousal around the midpoint in nonphobic individuals (as observed by Riegel et al [35]). According to

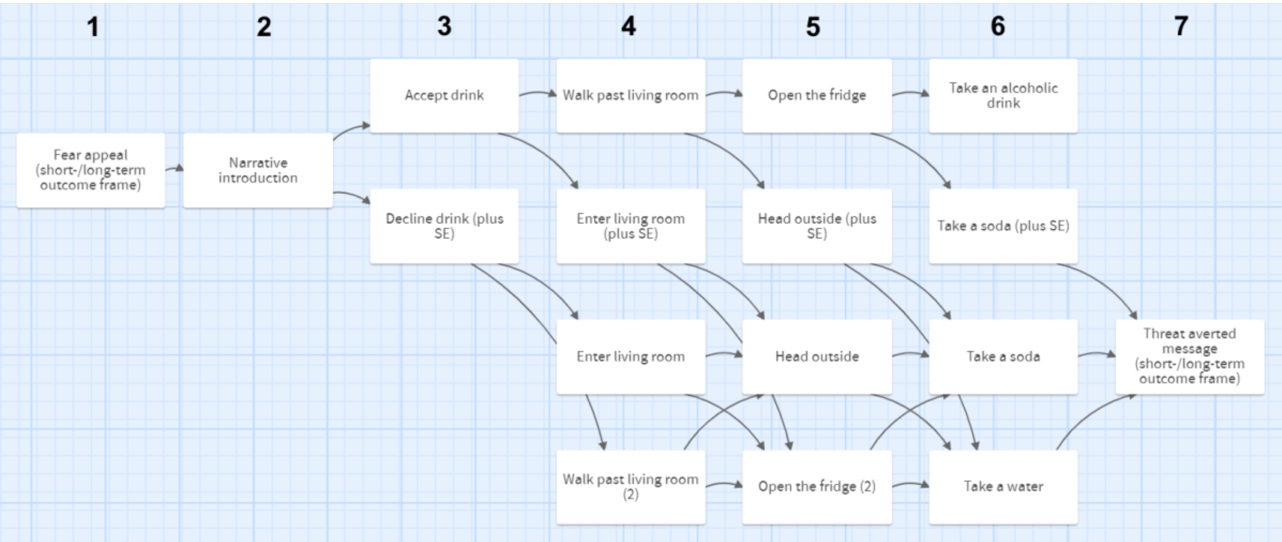
Tannenbaum et al [5], the effectiveness of a fear appeal underlies a u-shaped curve of the relationship between the amount of induced fear and successful persuasion of the individual. To avoid potential reactance because of the evocation of an excessively high fear response, a moderately fear-arousing pictorial stimulus was selected.

Agency

A total of 4 different versions of the narrative fear appeal were created (1 per condition).

As shown in Figure 2, agency was manipulated by giving participants choices at the end of every node (high agency) or by simply having them advance to the next page of the narrative by clicking the arrow button without making a decision (low agency). To be presented with the self-efficacy message in the high-agency condition, the participant had to execute the prevention behaviors (turning down an alcoholic drink) at the fictional party. To nudge participants to execute the prevention behaviors while retaining perceived agency, the participants were presented with 4 different decisions where they were able to decline or accept an alcoholic drink (Figure 2, nodes 2-5) using a foldback structure. This means that although the content presented to participants might differ slightly depending on the decision selected, the narrative would *fold back* to central events to enable the presentation of all participants with the same decision points. Once a participant declined the drink at any of the decision points, he or she was presented with a self-efficacy statement (Figure 2, row 2, denoted as *plus SE*) and was no longer able to choose options related to alcohol consumption. For example, a participant who accepted the alcoholic drink offer at the first decision point (Figure 2, node 2) will subsequently be presented with another 2 opportunities to turn down the alcoholic drink while the protagonist is making his way to the kitchen to obtain it. In decision 2, he can decide to walk past the living room and continue the pursuit of obtaining alcoholic drinks from the kitchen or decide not to drink that night (Figure 2, node 3) and enter the living room to chat with some friends who offer soft drinks. A participant who declines the initial drink offer for the first decision will still be asked to either enter the living room or continue toward the kitchen; however, these decisions (and all subsequent ones) will no longer be framed in terms of obtaining an alcoholic drink.

Figure 2. Overview of narrative nodes and progression for the high-agency condition. SE: self-efficacy.



Every participant received only 1 self-efficacy message, as subsequent choices were no longer related to the prevention behavior. The ending of the story highlights the meaningfulness of turning down the drink by showing the adverse consequences that drinking could have had that night (Figure 1, section 4) while the protagonist arrived home safely because of the decision made by the participant. Participants who chose to drink an alcoholic drink at all 4 decision points (Figure 2, node 2-5) were presented with a different ending (Figure 2, top row, node 6) that detailed the death of the protagonist from alcohol poisoning (high-immediacy condition) or long-term alcohol use (low-immediacy condition). Participants in the low-agency condition were presented with a preset combination of nodes (without decision options) matched with the participants in the high-agency condition. Low-agency participants simply read the story and could not influence the decisions of the protagonist.

Immediacy

To manipulate outcome immediacy, the fear appeal at the beginning of the narrative was presented as organ failure because of excessive alcohol consumption over a short period (high immediacy) or continuous consumption of large amounts of alcohol over a period of months (low immediacy). Similarly, the threat aversion message (Figure 1, right side) was also adjusted to reflect either the aversion to a long-term or short-term health threat.

Measures

Demographics

Demographic data were collected to account for appropriate randomization between the conditions. Participants were asked to indicate their age, gender, and educational background.

Dependent Measures

Self-efficacy beliefs were measured by adapting the item stems from Shi and Smith [36] to fit this study. Participants rated their perceived self-efficacy on 7-point Likert scales ranging from *strongly disagree* to *strongly agree* for 3 items. As the standards for what constitutes *too much alcohol* differ between individuals and are therefore difficult to specify in terms of the frequency of alcoholic drinks consumed [30], the items were formulated in terms of being able to *limit alcohol intake to low amounts*. For example, "I am able to limit my alcohol intake to low amounts." Reliability analysis of the 3-item self-efficacy measure (mean 6.049, SD 0.988) indicated good reliability of the scale (Cronbach α=.882).

Behavioral intention was measured using items adapted from Fisher et al [37]. Although other studies have used a single-item probability question (eg, see the study by Carrera et al [38]), the approach by Fisher et al [37] uses contextual cues to assess behavioral intentions. As we believe the risk behavior to be especially likely to be exhibited in these situations, the contextual questions are more likely to reflect behavioral intentions rooted in real-life experiences. For example, "I intend to limit my alcohol intake to low amounts when being with friends." The 3 items were scored on 7-point Likert scales from *strongly disagree* to *strongly agree*. The reliability analysis of the 3-item behavioral intention measure (mean 6.187, SD 1.193) indicated good reliability of the scale (Cronbach α=.882).

EPPM Measures

Items for the fear measure were taken from the Witte [4] recommendations for a different health threat. A total of 6 items were scored on 7-point Likert scales ranging from *not at all* to *very much*. The items asked participants about the intensity of affective response toward the message; for example, "How much did this message make you feel frightened?" Reliability analysis of the 6-item fear measure (mean 3.092, SD 1.357) indicated good reliability of the scale (Cronbach α=.898).

Items for perceived susceptibility and severity were adapted from Shi and Smith [36]. Although the items from Fisher et al [37] more closely resemble the items proposed by Witte [4], they do not explicitly connect the behavior and risk outcome for the response efficacy rating scales. Furthermore, Shi and Smith [36] used 3-item scales rather than the 2 items used by Fischer et al [37] and explicitly based their assumptions for their study on the Witte [4] EPPM rather than the focus on protection motivation theory used by Fisher et al [37].

Severity, susceptibility, and response efficacy items were measured with 3 items each on 7-point Likert scales from *strongly disagree* to *strongly agree*. As with the self-efficacy and behavioral intention measures, the measures are phrased in terms of either *drinking too much* or *drinking less* to account for individual differences in the perception of the appropriateness of alcohol intake.

The reliability analysis of severity (3 items; mean 6.333, SD 0.744; Cronbach α=.571), susceptibility (3 items; mean 2.832, SD 1.216; Cronbach α=.744), and response efficacy (3 items; mean 4.640, SD 1.308; Cronbach α=.792) showed moderate scale reliability.

Manipulation Check Perceived Agency

To validate agency manipulation, a scale was included to measure perceived agency after the participants experienced the narrative. The items for perceived agency were adapted from Fendt et al [39]. A question was dropped as it related to game enjoyment rather than perceived agency. The conceptualization of agency by Fendt et al [39] is congruent with the concept of agency applied in this study; for example, "I felt that the actions I took were meaningful within the context of the story." Items were measured on 7-point Likert scales from *strongly disagree* to *strongly agree*. Reliability analysis of the 5-item perceived agency measure (mean 4.524, SD 1.522) indicated good reliability of the scale (Cronbach α=.881).

Manipulation Check Immediacy

To check whether participants perceived a difference in the immediacy of the presented threat, they were asked to recall the content of the threat message presented at the beginning of the narrative. Participants had to indicate whether the protagonist died because of *drinking too much frequently over a long period* (low immediacy) or *drinking too much at a party over a short period* (high immediacy).

Sample Characteristics

To account for potential individual differences between participants, disinhibition, drinking frequency, alcohol dependence, and perceived control over drinking behavior were measured. Behavioral inhibition, measured by a subscale of the Sensation Seeking Scale-V by Zuckerman [40], has been extensively linked to differences in arousal to media messages. Specifically, subscale 3 (*disinhibition*) has been shown to indicate that low sensation seekers exhibit stronger affective responses to fear appeals addressing alcohol use [41] and are highly correlated with the risk of heavy drinking behavior [42]. The items to measure disinhibition were taken unchanged from the scale by Lee and Shinn [41] but altered from 9-point to 7-point Likert scales ranging from *strongly disagree* to *strongly agree*.

The subscales for measuring frequency and perceived control regarding excessive alcohol consumption were taken from Carrera et al [38]. The items on perceived control were changed from a binary question format to Likert scale statements.

Frequency was measured by asking participants how often they *drink alcohol in excess* and was measured on a 7-point Likert scale ranging from *never* to *very frequently*. Perceived control was measured by 2 items asking participants whether they could *control* and *stop* drinking alcohol whenever they wanted. These 2 items were measured on 7-point Likert scales ranging from *strongly disagree* to *strongly agree*.

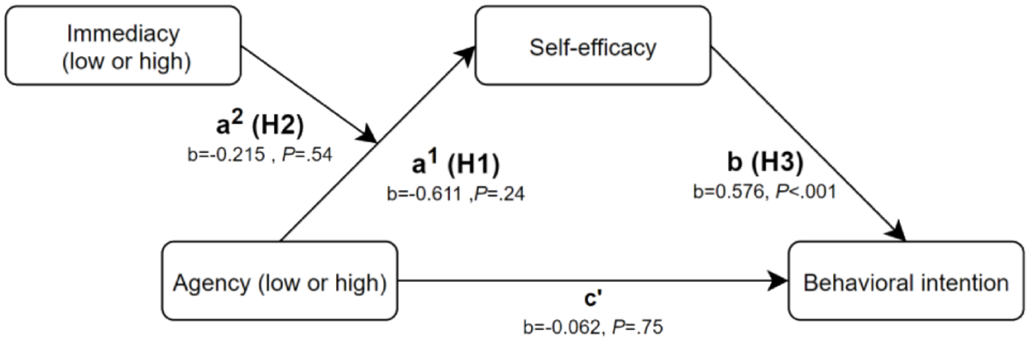
As alcohol dependence is difficult to measure accurately in student populations, the B-YAACQ was developed by Kahler et al [30] to measure the consequences of alcohol-related behaviors indicating dependence. An example item is "My drinking has gotten me into sexual situations I later regretted." The B-YAACQ comprises 24 statements scored with binary yes or no answers. A score of >15 positive answers indicates alcohol dependence.

Analysis

The conceptual model shown in Figure 3 was analyzed using the PROCESS [43] plug-in in SPSS in conjunction with template mode 7 (moderated mediation model). This enabled the analysis of the direct effect of agency on self-reported perceived self-efficacy (hypothesis 1; path a^1), moderation of the relationship between agency and self-efficacy by the outcome frame (hypothesis 2; path a^2), effect of self-efficacy on behavioral intention (hypothesis 3; path b), and potential direct effect of agency on behavioral intention (path c′). A total of 5000 bootstrap samples were used for the analysis.

The a priori power calculation assumed a medium effect size (f=0.25) with the desired power level of 0.8 for our 2×2 research design. To reach adequate power, 32 valid samples were needed per condition.

Figure 3. Moderated mediation model for the relationship between agency and behavioral intention. H:hypothesis.



Results

Deviation From Preregistration

For the main data analysis, a 2-way ANOVA was preregistered to test the effects of agency and immediacy on self-efficacy. However, we later realized that a more advanced model that also takes into account the possible moderating effect of immediacy provides a better fit for the data; hence, we report the analysis in the following sections. It should be noted that the preregistered ANOVA yielded essentially the same outcome.

Randomization Check

A total of 123 cases were analyzed. We first examined our control variables to ensure that their scores were equally distributed across all 4 conditions.

Three 1-way ANOVAs were conducted to test whether group assignment had an effect on the perceived control over alcohol intake, inhibition, or age. We obtained no significant results for group membership on perceived control ($F_{3,119}$=0.813; P=.49), inhibition ($F_{3,119}$=0.531; P=.66), and age ($F_{3,119}$=2.287; P=.08), indicating that the conditions did not have significantly different distributions of scores for alcohol intake, inhibition, and age. Chi-square tests were administered to ascertain the potential differences in the distribution of drinking frequency and gender across conditions. For drinking frequency, participants were grouped as nondrinkers, rare drinkers, or occasional and frequent drinkers. There were no significant differences in distributions across conditions for gender (N=123; χ^2_3=3.8; P=.28) or drinking frequency (N=123; χ^2_6=5.8; P=.44). In summary, no significant differences in perceived control over drinking behavior, inhibition, drinking frequency, age, and gender were found between the conditions. Thus, the randomization was successful.

Manipulation Check

An independent-sample t test (2-tailed) was performed to check for the effectiveness of the agency manipulation by comparing the condition (low or high agency) with the perceived agency measures. There was a significant difference between the low-agency (mean 4.132, SD 1.596) and high-agency (mean 4.921, SD 1.342) conditions (t_{121}=−2.965; P=.004), indicating an effective manipulation of agency.

Hypothesis Tests

The analysis of path a_1 indicates that the different levels of agency (low or high) are not significant predictors of the self-efficacy measure (path a_1: t_{119}=−1.182; P=.24; 95% bias-corrected and accelerated [BCa] CI −1.634 to 0.412), indicating that we had found no evidence to support hypothesis 1. There was also no interaction effect between agency and outcome frame as a predictor of self-efficacy (path a_2: t_{119}=.610; P=.54; 95% BCa CI −0.483 to 0.913). This indicates that the relationship between agency (low or high) and self-efficacy was not affected by the difference in outcome frames (high immediacy and low immediacy) presented to the participants (hypothesis 2). Looking at path b, we see that self-efficacy is a significant predictor of behavioral intention (path b: t_{120}=5.135; P<.001; 95% BCa CI 0.354-0.798; hypothesis 3). Given that self-efficacy is the mediator in this model, agency did not have a direct effect on behavioral intention (path c′: t_{120}=−0.325; P=.75; 95% BCa CI −0.437 to 0.314). Therefore, we observed no evidence to assume that the manipulation of agency and threat immediacy affected behavioral intention measures. An overview of these effects can be found in Table 1.

Table 1. Effects overview of the hypothesis tests for the moderated mediation model.

	Estimate (SE)	*t* test (*df*)	*P* value	95% CI
Direct effects				
Agency→self-efficacy	−0.611 (0.517)	−1.182 (119)	.24	−1.634 to 0.412
Self-efficacy→behavioral intention	0.576 (0.112)	5.135 (120)	<.001	0.354 to 0.798
Interaction effect				
Agency×immediacy→self-efficacy	0.215 (0.352)	0.610 (119)	.54	−0.483 to 0.913
Total effects				
Agency→behavioral intention	−0.062 (0.189)	−0.325 (120)	.75	−0.437 to 0.314

Exploratory Findings

Additional exploratory analyses were conducted to contextualize the null effects. A median split was conducted by analyzing the effect of agency on self-efficacy for participants scoring high on susceptibility to ascertain whether the appeal or the pre-existing characteristics of the sample caused a potential floor effect for the susceptibility measure. The adjusted sample (75/123, 61%) showed no significant relationship between the agency and self-efficacy measures (t_{73}=1.928; *P*=.06). Similarly, we evaluated whether a possible ceiling effect of self-efficacy measures might have influenced the overall effect of agency on behavioral intention. After conducting the median split, it was found that only a small number of participants (40/75, 53%) scored below the median, with a relatively large proportion scoring exactly on the median (24/75, 32%). Given the low variability in the sample and, thus, only a limited sample, it was decided not to conduct any further analysis on the remaining 32.5% (40/123) of participants.

Discussion

Principal Findings

Although the manipulation of agency was found to be successful, this study found no effect of agency on perceived self-efficacy, and this relationship was also not influenced by the framing of the threat as a short-term (high immediacy) or long-term (low immediacy) adverse health outcome. A significant effect was found for the relationship between self-efficacy and behavioral intention. To point forward for future studies, the null findings of this study need to be discussed with regard to their contribution by contrasting the different approaches taken for this study.

Operationalization of Agency

In this study, agency had no significant effect on the self-efficacy perceptions of the participants. Agency was operationalized as the one-time execution of an action with a meaningful impact on the outcome of the narrative. To achieve this, a foldback structure was used to ensure that participants would encounter the same narrative content and be presented with a complete fear appeal where the prevention behavior was executed, and a self-efficacy message could be displayed. Although the manipulation check showed a difference in perceived agency between high- and low-agency conditions, the perceived agency might still not have been impactful enough to cause a change in the self-efficacy measure. Owing to the nonexistence of related work on the effects of agency on self-efficacy, these results are difficult to compare with other operationalizations of agency in the field. However, looking at the different theoretical conceptualizations of agency, we can derive possible explanations.

The manipulation in this study could conceivably be seen as a low-agency condition if agency is defined as the balance of affordances and constraints provided by the system [44]. Using the foldback structure increases constraints for affordances (ie, the number of choices) in our narrative, leading to less agency than using a system that provides more choices (ie, branches) to the user. It can be argued that agency is restricted to only one possible outcome in our narrative. According to Crawford [45], the degree of interactivity is determined by the functional significance of a choice and its perceived completeness. Perceived completeness means that the number of choices corresponds to the readers' imagined number of possible choices. Although functional significance was high in this study, relating decisions back to story outcomes, perceived completeness was very limited. This highlights the antagonistic relationship of targeted interventions using a game-like format. With less authorial control over the story, the intervention can exert little control over the presentation of the message in terms of the content and chronology of persuasive aspects. This is especially the case when applying a health intervention framework (such as the EPPM), which segments the intervention into distinct parts.

Heightening perceived completeness could be achieved, given the successful use of illusory agency, presenting options that are perceived to be more complete even if they fold back to the same outcomes, or by loosening authorial control in more open experiences such as a sandbox game that enables emergent storytelling. The question is whether the elicitation of self-efficacy through agency is reliant on control over the chronology of message components or whether heightened agency through emergent storytelling can make up for the loss in control by providing a closer link between the protagonist and the message receiver in terms of perceived completeness of the available actions. Taking this a step further, there is evidence to form the assumption that agency does have an effect on persuasion, even if the agency manipulation is completely decoupled from the persuasion attempt and simply induced as a state before the intervention (eg, see the study by Damen et al [46]).

The wide range of used agency definitions and their heterogeneous operationalization make generalizations about agency effects in health interventions difficult. For future studies, the plural modality of agency should be taken into account [47] and tested experimentally to establish discriminant validity for different agency constructs. This means that not only the kind of agency needs to be taken into account but also how this can be operationalized most effectively, within or outside the intervention, to heighten the extent of perceived agency in participants.

Immediacy and the Malleability of Self-efficacy Beliefs

In this study, we found no significant effect of threat immediacy on the relationship between agency and self-efficacy. Organ damage was chosen as the threat as it can be a result of either short-term or long-term risk behavior and was henceforth communicated as a consequence of a single night or multiple months of drinking. There is ample work evaluating the effect of fear appeals on one-time versus repeated risk prevention behaviors, showing that appeals advising the one-time execution of behaviors are more likely to succeed [5]. However, comparisons of short- and long-term outcomes in fear appeal threats are scarce, making it difficult to base the manipulation on existing work in the field. Although the manipulation checks ensured that participants noticed the manipulation, this might not have made the framing of threat immediacy salient enough to result in a significant effect on the relationship between agency and self-efficacy.

In most IDN research, as well as game studies, behavior transfer between digital and real-life behavior is often assumed to be high when the digital execution matches that of real-life analog behavior. The prevention behavior for the threat in this study does not constitute a one-time behavior, as it needs to be executed multiple times for the threat to be averted. With regard to immediacy in this study, there is, therefore, a mismatch between both our short- and long-term threats and our digital narrative in which the prevention behavior (rejecting a drink) is only executed once. Although a single appeal message can influence self-efficacy for repeated prevention behaviors (eg, see the study by Smith and Stutts [29]), in this study, a single execution of the behavior might not have been enough to affect self-efficacy beliefs. A possible explanation could be the depth of processing of the message itself, which hinders message acceptance and, therefore, undermines possible effects on short- and long-term threat prevention. Presentation with a single persuasive message activates the systematic processing of this message. Heuristic processing occurs only after repeated exposure, which breaks down resistance to persuasion and heightens fear appeal effectiveness [36].

A possible direction for future studies would be to make the long-term, as well as the short-term, impact of risk behavior more salient through the provision of repeated interventions. As discussed by Shi and Smith [36], EPPM-based fear appeals with long-term threats are more effective after 3 exposures than after a single exposure. It is unclear whether this also holds for short-term threats. Future research should investigate whether intervention repetition affects self-efficacy differently for threats with high and low immediacy.

Lack of Susceptibility

The null effects for agency and threat immediacy could also result from the low perceived susceptibility of the target population (young adults) to the threat used in this study (organ failure because of excessive drinking). The so-called *optimism* concerning the likelihood of adverse health effects has been well established in other studies and potentially creates reactance during persuasive attempts for risk behaviors involving drinking, drunk driving, and smoking in young adults [48-50]. Organ failure is an extreme health threat that might be too abstract (ie, perceived as unrealistic) for young adults, leading to low perceived susceptibility. It can be argued that although organ failure from one night of binge drinking is immediate in a temporal sense, it is quite distant when it comes to the perceived likelihood of affecting young adults. This means that although general severity and response efficacy evaluations can be perceived as high, the link to personal relevance does not come into play because of low perceived susceptibility and the high baseline of perceived self-efficacy for this population (ie, optimism).

The issue of perceived threat susceptibility can be further explained by our exclusion criteria. The criteria specifically excluded participants who showed signs of alcohol dependence when they might actually have been participants for whom the threat would be most relevant and the preappeal efficacy would be low. However, as the aim is to increase the effectiveness of early intervention approaches, more effective ways of eliciting prevention behaviors for normal populations should still be investigated more thoroughly.

An alternative approach could be the use of appeals that do not focus on health threats to increase susceptibility as a prerequisite for successful narrative persuasion. Health threats themselves have been the main focus of health-related interventions and are effective for older cohorts; however, this might not always be the case for younger populations. As suggested by Pechmann et al [51], the use of social threats seems to be more effective for younger populations as they are more immediate than health threats are perceived, leading to higher perceived susceptibility. Therefore, it would be valuable to extend this line of research by trying to evaluate the effectiveness of appeals addressing short-term versus long-term social threats. As argued previously, the potentially weak cause-effect chain perceived by young participants could also potentially be strengthened through the use of social threats, which in turn might strengthen the effect of agency. If the threat is seen as relevant and the prevention behavior as effective, having participants execute the prevention behavior to avert a social threat might heighten self-efficacy, given high perceived agency over the narrative outcome.

Methodological Limitations

This study did not use a pre- or postmeasures design to account for the potential influence of baseline beliefs already present in participants before the appeal message was presented. However, it has to be noted that disinhibition, drinking frequency, alcohol dependence, and perceived control over drinking behavior were measured to account for individual differences between participants. These were measured after the appeal message so as to minimize the influence on the responses to the IDN appeal.

Although we believe that these constructs relate to either stable traits or general behavior and therefore should not have been influenced by the manipulation, we cannot fully discount the potential influence of the IDNs on these constructs.

Furthermore, it should be noted that the current sample was predominantly women (88/123, 72%). Previous studies have shown gender differences in the effectiveness of fear appeals, which might have influenced the results. In general, this sample might not be representative of the general population.

Finally, this study fell short of its goal of collecting a total of 140 valid samples, leading to a potential shortcoming in statistical power, given the complex study design. Furthermore, it could even be that the authors assumed a medium effect size for a single intervention transferring virtual behaviors to real-world beliefs and behavioral intentions.

Contextual Factors

Finally, the context of this study must be mentioned. The study was conducted during the COVID-19 pandemic, which limited the participants' exposure to risk situations and could have skewed the general perception of the appeal as personally relevant. Although there was no active lockdown in place at the time of data collection, there was a limit to the number of people allowed to gather more generally. This made the otherwise common risk environments that facilitate drinking, such as parties, almost nonexistent. As social facilitation of drinking is one of the core determinants of risk behavior related to excessive alcohol use for college students [2], the context during which the data collection took place could have affected self-efficacy beliefs because of nonexposure to the threat in real life. Considering the abovementioned limitations of this study, going forward, the findings should be replicated in a nonpandemic context to be generalizable for the target population.

Conclusions

In this preregistered study, no significant effects were found for agency on self-efficacy and behavioral intention. In addition, we observed no effect for the influence of threat immediacy on the relationship between agency and self-efficacy.

The multitude of conceptual distinctions of different forms of agency and the different ways of integrating them into IDNs provides a challenge that needs further evaluation to affect self-efficacy and, more generally, behavioral change. There is a need to empirically test the difference between conceptually different agency concepts to establish their impact on persuasiveness in narrative health interventions. Although there is evidence to assume the effects of interaction on antecedents of narrative persuasion (eg, identification and transportation), more work is needed to understand how different agency conceptualizations directly affect the processing of health messages. In addition, more work is needed that empirically contrasts passive and interactive interventions to isolate the effects. Much of the previous work done in this field use fully-fledged serious game experiences without passive controls, making it difficult to ascribe effects to single factors in such complex systems. With regard to the theoretical assumptions underlying narrative fear appeals, the personal relevance of the content should be ensured through pilot testing before the deployment of the intervention. Without creating the personal relevance of the narrative fear appeal, the intended effects could be undermined because of a lack of perceived susceptibility to the threat. Understanding how different kinds of threats are perceived by different target populations should be investigated more thoroughly to adapt interventions more effectively. Furthermore, the effect of intervention repetition for different kinds of prevention behaviors is poorly understood, making generalizations across different fear appeal threats difficult. A systematic approach contrasting one-time and repeated prevention behaviors for different kinds of threats could establish a clearer picture of the potential necessity of adaptations for threats that are the result of repeated risk behaviors.

In conclusion, despite the null findings for the effects of agency on self-efficacy, this study highlights the potential for further exploration of agency concepts and threat types in terms of their embedding into a narrative intervention.

Authors' Contributions

HE designed the study and wrote the manuscript. LNvdL, RvE, and EK supervised the entire process and provided feedback on the study and the written manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Editorial notice: This randomized study was retrospectively registered, given that the authors were not aware that a study with a student sample needs to be registered as a clinical trial for publication with JMIR. The editor granted an exception from ICMJE 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.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 9223 KB - games_v10i2e32218_app1.pdf]

References

1. Fact sheet on alcohol consumption, alcohol-attributable harm and alcohol policy responses in European Union Member States, Norway and Switzerland. World Health Organization Regional Office for Europe. 2018. URL: https://www.euro.who.int/__data/assets/pdf_file/0009/386577/fs-alcohol-eng.pdf [accessed 2022-01-15]
2. Beck KH, Caldeira KM, Vincent KB, Arria AM. Social contexts of drinking and subsequent alcohol use disorder among college students. Am J Drug Alcohol Abuse 2013 Jan;39(1):38-43 [FREE Full text] [doi: 10.3109/00952990.2012.694519] [Medline: 22746152]
3. Roth C, Koenitz H. Evaluating the user experience of interactive digital narrative. In: Proceedings of the 1st International Workshop on Multimedia Alternate Realities. 2016 Oct 16 Presented at: AltMM '16; October 16, 2016; Amsterdam, The Netherlands p. 31-36. [doi: 10.1145/2983298.2983302]
4. Witte K. Fear control and danger control: a test of the extended parallel process model (EPPM). Commun Monogr 1994;61(2):113-134. [doi: 10.1080/03637759409376328]
5. Tannenbaum MB, Hepler J, Zimmerman RS, Saul L, Jacobs S, Wilson K, et al. Appealing to fear: a meta-analysis of fear appeal effectiveness and theories. Psychol Bull 2015 Nov;141(6):1178-1204 [FREE Full text] [doi: 10.1037/a0039729] [Medline: 26501228]
6. Ruiter RA, Kessels LT, Peters GY, Kok G. Sixty years of fear appeal research: current state of the evidence. Int J Psychol 2014 Apr;49(2):63-70. [doi: 10.1002/ijop.12042] [Medline: 24811876]
7. Stainback RD, Rogers RW. Identifying effective components of alcohol abuse prevention programs: effects of fear appeals, message style, and source expertise. Int J Addict 1983 Apr;18(3):393-405. [doi: 10.3109/10826088309039356] [Medline: 6874160]
8. Moscato S, Black DR, Blue CL, Mattson M, Galer-Unti RA, Coster DC. Evaluating a fear appeal message to reduce alcohol use among "Greeks". Am J Health Behav 2001;25(5):481-491. [Medline: 11518342]
9. Ooms JA, Jansen CJ, Hoeks JC. The story against smoking: an exploratory study into the processing and perceived effectiveness of narrative visual smoking warnings. Health Educ J 2020 Mar 1;79(2):166-179. [doi: 10.1177/0017896919867436]
10. Bilandzic H. Narrative persuasion. In: Dillard JP, Shen L, editors. The SAGE Handbook of Persuasion: Developments in Theory and Practice. Thousand Oaks, CA, USA: Sage Publications; 2013:200-219.
11. Cohen J. Defining identification: a theoretical look at the identification of audiences with media characters. Mass Commun Soc 2001 Aug;4(3):245-264. [doi: 10.1207/S15327825MCS0403_01]
12. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. J Pers Soc Psychol 2000 Nov;79(5):701-721. [doi: 10.1037//0022-3514.79.5.701] [Medline: 11079236]
13. de Graaf A, Hoeken H, Sanders J, Beentjes JW. Identification as a Mechanism of Narrative Persuasion. Commun Res 2011 May 16;39(6):802-823. [doi: 10.1177/0093650211408594]
14. Ooms J, Hoeks J, Jansen C. "Hey, that could be me": the role of similarity in narrative persuasion. PLoS One 2019 Apr 18;14(4):e0215359 [FREE Full text] [doi: 10.1371/journal.pone.0215359] [Medline: 30998694]
15. Duchan JF, Bruder GA, Hewitt LE. Deixis in Narrative: A Cognitive Science Perspective. East Sussex, UK: Psychology Press; Dec 06, 2012.
16. van Laer T, de Ruyter K, Visconti LM, Wetzels M. The extended transportation-imagery model: a meta-analysis of the antecedents and consequences of consumers' narrative transportation. J Consum Res 2014 Feb 01;40(5):797-817 [FREE Full text] [doi: 10.1086/673383]
17. de Graaf A, Sanders J, Hoeken H. Characteristics of narrative interventions and health effects: a review of the content, form, and context of narratives in health-related narrative persuasion research. Rev Commun Res 2016;4:88-131. [doi: 10.12840/issn.2255-4165.2016.04.01.011]
18. Dahlstrom MF. The persuasive influence of narrative causality: psychological mechanism, strength in overcoming resistance, and persistence over time. Media Psychol 2012 Jul;15(3):303-326. [doi: 10.1080/15213269.2012.702604]
19. Murray J. Hamlet on the Holodeck: The Future of Narrative in Cyberspace. Cambridge, MA, USA: MIT Press; 2017.
20. Appel M, Richter T. Transportation and need for affect in narrative persuasion: a mediated moderation model. Media Psychol 2010 May 28;13(2):101-135. [doi: 10.1080/15213261003799847]
21. Tanenbaum K, Tanenbaum TJ. Agency as commitment to meaning: communicative competence in games. Digit Creat 2010 May 26;21(1):11-17. [doi: 10.1080/14626261003654509]
22. Winskell K, Sabben G, Obong'o C. Interactive narrative in a mobile health behavioral intervention (Tumaini): theoretical grounding and structure of a smartphone game to prevent HIV among young Africans. JMIR Serious Games 2019 May 08;7(2):e13037 [FREE Full text] [doi: 10.2196/13037] [Medline: 31066692]
23. Carvalho DB, Domingueti DB, Almeida Santos SM, Dias DR. A game to deal with alcohol abuse (Jib): development and game experience evaluation. JMIR Serious Games 2019 Oct 15;7(4):e11151 [FREE Full text] [doi: 10.2196/11151] [Medline: 31617849]
24. Panic K, Cauberghe V, De Pelsmacker P. Promoting dental hygiene to children: comparing traditional and interactive media following threat appeals. J Health Commun 2014;19(5):561-576. [doi: 10.1080/10810730.2013.821551] [Medline: 24393019]

25. Kim J, Song H, Merrill Jr K, Jung Y, Kwon RJ. Using serious games for antismoking health campaigns: experimental study. JMIR Serious Games 2020 Dec 02;8(4):e18528 [FREE Full text] [doi: 10.2196/18528] [Medline: 33263549]
26. Klimmt C, Hartmann T. Effectance, self-efficacy, and the motivation to play video games. In: Vorderer P, Bryant J, editors. Playing Video Games: Motives, responses, and consequences. New York, NY, USA: Routledge; Mar 28, 2006:28-45.
27. Bandura A. Social Learning Theory. Englewood Cliffs, NJ, USA: Prentice Hall; Oct 1977.
28. Slater MD. Integrating application of media effects, persuasion, and behavior change theories to communication campaigns: a stages-of-change framework. Health Commun 1999 Oct;11(4):335-354. [doi: 10.1207/s15327027hc1104_2]
29. Smith KH, Stutts MA. Effects of short-term cosmetic versus long-term health fear appeals in anti-smoking advertisements on the smoking behaviour of adolescents. J Consum Behav 2003 Dec;3(2):157-177. [doi: 10.1002/cb.130]
30. Kahler CW, Strong DR, Read JP. Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: the brief young adult alcohol consequences questionnaire. Alcohol Clin Exp Res 2005 Jul;29(7):1180-1189. [doi: 10.1097/01.alc.0000171940.95813.a5] [Medline: 16046873]
31. Noar SM, Hall MG, Francis DB, Ribisl KM, Pepper JK, Brewer NT. Pictorial cigarette pack warnings: a meta-analysis of experimental studies. Tob Control 2016 May;25(3):341-354 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051978] [Medline: 25948713]
32. Leshner G, Bolls P, Thomas E. Scare' em or disgust 'em: the effects of graphic health promotion messages. Health Commun 2009 Jul;24(5):447-458. [doi: 10.1080/10410230903023493] [Medline: 19657827]
33. 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]
34. Marchewka A, Zurawski Ł, Jednoróg K, Grabowska A. The Nencki Affective Picture System (NAPS): introduction to a novel, standardized, wide-range, high-quality, realistic picture database. Behav Res Methods 2014 Jun;46(2):596-610 [FREE Full text] [doi: 10.3758/s13428-013-0379-1] [Medline: 23996831]
35. Riegel M, Żurawski Ł, Wierzba M, Moslehi A, Klocek Ł, Horvat M, et al. Characterization of the Nencki Affective Picture System by discrete emotional categories (NAPS BE). Behav Res Methods 2016 Jun;48(2):600-612 [FREE Full text] [doi: 10.3758/s13428-015-0620-1] [Medline: 26205422]
36. Shi J, Smith SW. The effects of fear appeal message repetition on perceived threat, perceived efficacy, and behavioral intention in the extended parallel process model. Health Commun 2016;31(3):275-286. [doi: 10.1080/10410236.2014.948145] [Medline: 26305152]
37. Fisher JJ, Almanza BA, Behnke C, Nelson DC, Neal J. Norovirus on cruise ships: motivation for handwashing? Int J Hosp Manag 2018 Sep;75:10-17. [doi: 10.1016/j.ijhm.2018.02.001]
38. Carrera P, Muñoz D, Caballero A. Mixed emotional appeals in emotional and danger control processes. Health Commun 2010 Dec;25(8):726-736. [doi: 10.1080/10410236.2010.521914] [Medline: 21153989]
39. Fendt MW, Harrison B, Ware SG, Cardona-Rivera R, Roberts DL. Achieving the illusion of agency. In: Proceedings of the 5th International Conference on Interactive Storytelling. 2012 Nov 12 Presented at: ICIDS '12; November 12-15, 2012; San Sebastian, Spain p. 114-125. [doi: 10.1007/978-3-642-34851-8_11]
40. Zuckerman M. The sensation seeking scale V (SSS-V): still reliable and valid. Pers Individ Differ 2007 Oct;43(5):1303-1305. [doi: 10.1016/j.paid.2007.03.021]
41. Lee MJ, Shin M. Fear versus humor: the impact of sensation seeking on physiological, cognitive, and emotional responses to antialcohol abuse messages. J Psychol 2011;145(2):73-92. [doi: 10.1080/00223980.2010.532519] [Medline: 21449245]
42. Hittner JB, Swickert R. Sensation seeking and alcohol use: a meta-analytic review. Addict Behav 2006 Aug;31(8):1383-1401. [doi: 10.1016/j.addbeh.2005.11.004] [Medline: 16343793]
43. Hayes AF. PROCESS: a versatile computational tool for observed variable mediation, moderation, and conditional process modeling. Ohio State University. 2012. URL: http://www.afhayes.com/public/process2012.pdf [accessed 2022-01-15]
44. Mateas M, Stern A. Interaction and narrative. In: Tekinbaş KS, Zimmerman E, editors. The Game Design Reader: A Rules of Play Anthology. Cambridge, MA, USA: MIT Press; 2006:642-669.
45. Crawford C. Chris Crawford on Interactive Storytelling. 2nd edition. San Francisco, CA, USA: New Riders; 2013.
46. Damen TG, Müller BC, van Baaren RB, Dijksterhuis A. Re-examining the agentic shift: the sense of agency influences the effectiveness of (self)persuasion. PLoS One 2015 Jun 8;10(6):e0128635 [FREE Full text] [doi: 10.1371/journal.pone.0128635] [Medline: 26053303]
47. Jennings SC. A meta-synthesis of agency in game studies. trends, troubles, trajectories. GAME Italian J Game Stud 2019 Dec 30;1(8):87.
48. Kang J, Lin CA. Effects of message framing and visual-fear appeals on smoker responses to antismoking ads. J Health Commun 2015;20(6):647-655. [doi: 10.1080/10810730.2015.1012242] [Medline: 25868549]
49. Masiero M, Riva S, Oliveri S, Fioretti C, Pravettoni G. Optimistic bias in young adults for cancer, cardiovascular and respiratory diseases: a pilot study on smokers and drinkers. J Health Psychol 2018 Apr;23(5):645-656. [doi: 10.1177/1359105316667796] [Medline: 27624614]
50. Davies EL, Joshi MS. "Here's to a night of drunken mistakes": exploring experiences, regrets, and optimism in young adult drinkers. Subst Use Misuse 2018 Nov 10;53(13):2174-2183. [doi: 10.1080/10826084.2018.1461227] [Medline: 29676652]

51. Pechmann C, Zhao G, Goldberg ME, Reibling ET. What to convey in antismoking advertisements for adolescents: the use of protection motivation theory to identify effective message themes. J Mark 2003 Apr 1;67(2):1-18. [doi: 10.1509/jmkg.67.2.1.18607]

Abbreviations

BCa: bias-corrected and accelerated
B-YAACQ: Brief Young Adult Alcohol Consequences Questionnaire
EPPM: extended parallel process model
IDN: interactive digital narrative

Edited by N Zary; submitted 19.07.21; peer-reviewed by K Winskell, D Smith; comments to author 24.09.21; revised version received 19.12.21; accepted 16.03.22; published 13.06.22.

Please cite as:
Engelbrecht H, van der Laan LN, van Enschot R, Krahmer E
The Role of Agency and Threat Immediacy in Interactive Digital Narrative Fear Appeals for the Prevention of Excessive Alcohol Use: Randomized Controlled Trial
JMIR Serious Games 2022;10(2):e32218
URL: https://games.jmir.org/2022/2/e32218
doi:10.2196/32218
PMID:35699976

Original Paper

Identifying Player Types to Tailor Game-Based Learning Design to Learners: Cross-sectional Survey using Q Methodology

A E J Van Gaalen[1], MD; J Schönrock-Adema[2], PhD; R J Renken[3], PhD; A D C Jaarsma[2,4], Prof Dr; J R Georgiadis[1], PhD

[1]Anatomy & Medical Physiology Section, Department of Biomedical Sciences of Cells and Systems, University Medical Center Groningen, University Groningen, Groningen, Netherlands

[2]Wenckebach Institute for Education and Training, University Medical Center Groningen, Groningen, Netherlands

[3]Cognitive Neuroscience Center, Department of Biomedical Sciences of Cells and Systems, University Medical Center Groningen, University Groningen, Groningen, Netherlands

[4]Faculty of Veterinary Medicine, Utrecht University, Utrecht, Netherlands

Corresponding Author:
J R Georgiadis, PhD
Anatomy & Medical Physiology Section
Department of Biomedical Sciences of Cells and Systems
University Medical Center Groningen, University Groningen
Antonius Deusinglaan 1
Groningen, 9713AV
Netherlands
Phone: 31 648527717
Email: j.r.georgiadis@umcg.nl

Abstract

Background: Game-based learning appears to be a promising instructional method because of its engaging properties and positive effects on motivation and learning. There are numerous options to design game-based learning; however, there is little data-informed knowledge to guide the choice of the most effective game-based learning design for a given educational context. The effectiveness of game-based learning appears to be dependent on the degree to which players like the game. Hence, individual differences in game preferences should be taken into account when selecting a specific game-based learning design.

Objective: We aimed to identify patterns in students' perceptions of play and games—player types and their most important characteristics.

Methods: We used Q methodology to identify patterns in opinions on game preferences. We recruited undergraduate medical and dental students to participate in our study and asked participants to sort and rank 49 statements on game preferences. These statements were derived from a prior focus group study and literature on game preferences. We used by-person factor analysis and varimax rotation to identify common viewpoints. Both factors and participants' comments were used to interpret and describe patterns in game preferences.

Results: From participants' (n=102) responses, we identified 5 distinct patterns in game preferences: the social achiever, the explorer, the socializer, the competitor, and the troll. These patterns revolved around 2 salient themes: sociability and achievement. The 5 patterns differed regarding cheating, playing alone, story-telling, and the complexity of winning.

Conclusions: The patterns were clearly interpretable, distinct, and showed that medical and dental students ranged widely in how they perceive play. Such patterns may suggest that it is important to take students' game preferences into account when designing game-based learning and demonstrate that not every game-based learning-strategy fits all students. To the best of our knowledge, this study is the first to use a scientifically sound approach to identify player types. This can help future researchers and educators select effective game-based learning game elements purposefully and in a student-centered way.

(JMIR Serious Games 2022;10(2):e30464) doi:10.2196/30464

KEYWORDS

gamification; serious games; game-based learning; medical education; computers; new technology; focus group; play; qualitative; player types; taxonomy; theory; framework

Introduction

In health professional education, there has been a growing interest in game-based learning because of its engaging properties and positive effects on students' motivation and learning [1]. Yet, the understanding of how and when to implement game-based learning in educating health professionals remains in its infancy [1] as well as in other educational domains [2]. Although there are myriad ways to design game-based learning strategies [2-4], there is little good-quality evidence to guide the choice of the most effective game-based learning design in a given educational context [1]. This, in turn, may increase the likelihood of choosing suboptimal or even counterproductive game-based learning strategies [5]. Hence, there is a need for empirical research to inform future game-based learning design [1].

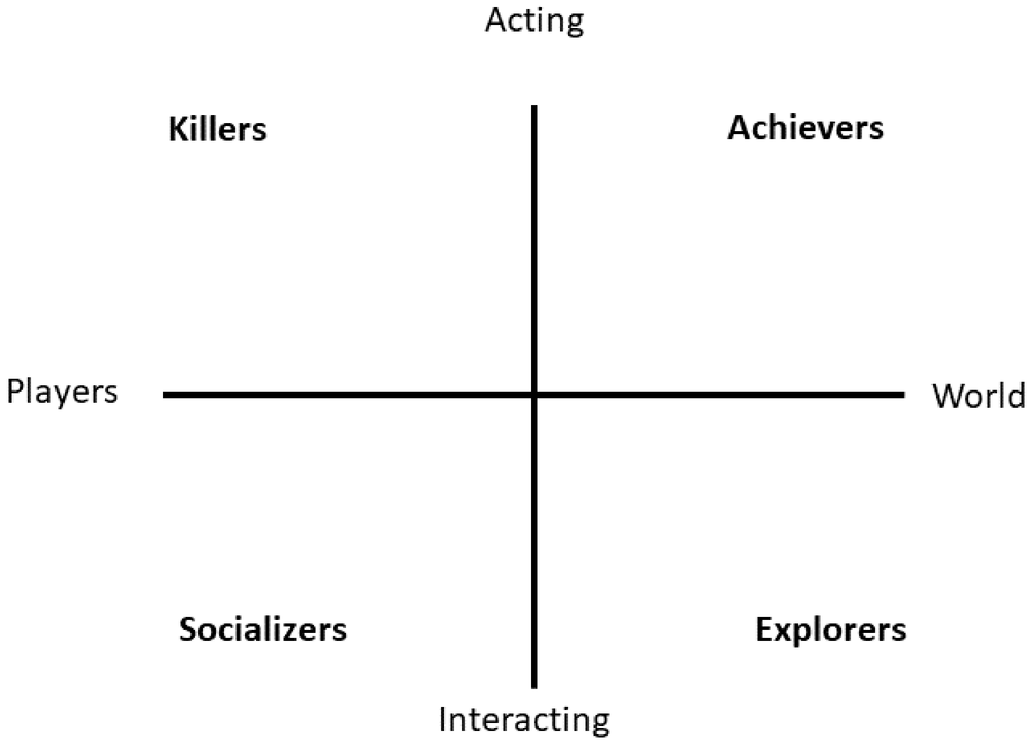
Some scholars have stated that educational games are designed by academics who do not understand the culture, art, and science of games [6-8]. This may result in educational learning tools that can either be a success or a failure with respect to playability and engagement. On the other hand, games developed by game designers with little or no understanding of the theory and practice of game-based learning can be fun to play but are also hit-or-miss with respect to educational goals and outcomes. Indeed, designing an educationally sound game-based learning tool is a challenging task and depends highly on the synergy between pedagogy and engagement [5-8].

Practical applications of game-based learning have not been substantiated by a significant body of scientific research [1,3,9], which could be interpreted as corroboration for the abovementioned assertions. Researchers in health professional education generally take an educational approach to game-based learning without considering the body of knowledge available in the field of game research. For example, most game-based learning research in health professional education focused on one specific game attribute (ie, the effects of scoring and rewards) [1,2,10], although many other game attributes have also been investigated [3,11]. Moreover, game elements that motivate some learners may actually demotivate others [12-14] indicating that personal preference is a crucial element for motivation to play [15-18]. Game [19] and game-based learning [5,20] research consistently demonstrated that people vary greatly in what they like in play and games. Outside the domain of education for health professionals, individual differences in age, gender, culture, and personality play a role in a person's preferences for specific types of play, games, and responses to different game-based learning designs [20]. Linking personality traits with game-based learning design solutions that best fit each particular trait has been shown to improve learner experience (eg, perceived playfulness) [12,21-27], motivation [28-31], and performance [28,30]. Hence, preferences should be considered in designing game-based learning strategies to engage and motivate an entire cohort of students (not only a subgroup).

In the field of game research, the concept of player types is used to characterize users who share preference for a specific type of play, which enables game designers to create an optimal user experience [32]. In an earlier and probably best-known player typology [18], users of a multiplayer role-playing game were classified on the basis of two in-game behaviors—(1) acting versus interacting and (2) world versus player—which resulted in 4 different player types: socializers (users who like to interact with other players, eg, the game is a tool to meet other people), explorers (users who like to interact with the world, eg, discover new areas, and immerse themselves in the game world), achievers (users who like to act on the world, eg, prefer gaining rewards, points, and equipment from the game world), and killers (users who like to act upon other players, eg, thrive on beating other people) (Figure 1). Since then, various player types have been proposed [17,19,33-35]. However, there are major concerns with these player typologies. Many are not supported by empirical evidence [35]. Instead, they are based on researchers' prior experience in developing games [18,33], on nonscientific literature [36,37], or on combinations of some of the aforementioned player types [38]. Player typologies based on empirical data [17,19,39] tend to be based on research into a specific game genre, which means the typologies may be biased and incomplete. In addition, surveys that were used (eg, Likert-scale surveys [17,19,39]) may have unnecessarily limited respondents' answers and, thereby, researchers' interpretations. Hence, important information may have been overlooked [40].

In this study, we aimed to identify player types among a representative group of education for health professional students, independent of game context. As the first study of its kind, we aimed to explore the widest possible range of preferences for game and play in this group. We formulated the following research question: What patterns in students' perceptions of play and games (ie, player types) can be identified and what are their most important characteristics?

Figure 1. Typology [18] with 4 player types based on 2 themes: player vs world and acting vs interacting.



Methods

Design

To investigate students' perceptions of play, we applied Q-methodology [41], which is a mixed methods research technique that aims to account for all key subjective viewpoints on a certain matter [42]. The qualitative component of Q-methodology allows the expression of subjective opinions to be considered, and the quantitative component uses statistical analysis in order to group participants with shared opinions. Q-methodology is used to cluster individuals based on shared opinions rather than based on latent variables, which is the case in regular factor analysis (or R-methodology). The Q-methodology technique has been used before in medical education, although for different purposes (eg, for identifying different patterns of self-regulating learning behavior [43-45]). Q-methodology is characterized by two main stages: (1) designing a set of statements and let participants sort that set of statements, and (2) by-person correlation and factor analysis of a sample of the included participants.

Ethics

We obtained ethical approval for this study from the Netherlands Association for Medical Education (NVMO 2019.1.11).

Setting and Participants

This study was conducted in May 2019 at the Faculty of Medical Sciences of the University of Groningen, the Netherlands. The 6-year undergraduate medical and the 6-year undergraduate dental curriculum both consist of a 3-year Bachelor and a 3-year Master's phase. In both curricula, teachers used face-to-face and web-based teaching methods and sometimes apply game-based learning, but not on a regular or structural basis.

We invited all medical and dental students (3000 eligible undergraduate students) to participate in our study by posting an announcement on the web-based learning environment (called Nestor) of the university. Participants were informed about the purpose and procedures of the study. Their participation was anonymous, voluntary, and confidential. Participants had the right to withdraw from the study at any time. All participants provided informed consent. In appreciation for their time and effort, each participant received a 5€ (approximately US $5.66) gift certificate.

Although there is no decisive minimum or maximum number of participants for performing Q-methodology research [43], generally, the number of statements should exceed the number of participants [46], and 40 to 60 participants is considered adequate [47,48]. To achieve a highly diverse sample—which is recommended for Q-sort analysis [46,48]—we purposively selected participants. Therefore, we asked participants to complete a sociodemographic questionnaire (eg, age, gender, and whether or not they considered themselves a gamer) prior to the sorting process that also included a question about the participant's favorite game. For our purposive sample, we included only participants who had specified a favorite game. We identified the game genre to ensure each game genre was represented evenly in the final sample. In addition, only participants who had performed the sorting task in 12 minutes or more were included. We conducted pilot testing and found that the average sorting time was 25 minutes (range 2-3 minutes) and that reading the statements and swiftly sorting the statements took at least 12 minutes; thus, for less than 12 minutes, the sort was regarded as ill-considered. If there were participants with identical favorite games, only one participant was included. Participants' preferences for modality (digital or analog) had to be distributed as evenly as possible across game genres; the male-to-female ratio had to be evenly distributed across game genres; and medical and dental student had to be evenly distributed across game genres. If a decision about inclusion or exclusion of a participant could not be made based the preceding criteria, the decision was made by rolling dice.

Statement Set

There is no single correct way to compile a set of statements in Q-methodology [41]. In general, the Q sample size is 40 to 80 statements [41,48], and the number of statements should exceed the number of participants [46]. A set containing too many statements can make the sorting process an exhausting and burdensome task, whereas a set containing too few statements may result in inadequate coverage of the topic of game preferences [41]. By sorting and prioritizing each statement from the statement set, individual participants provided us with a model of their view on their own game preferences. Statements should be carefully selected since their nature limits what can be expressed by a participant [49].

We aimed to develop a set of statements in which each statement was unique and made its own original contribution, and all statements together covered the full range of game preferences. Statements were based on the findings of an earlier focus group study [5] among medical and dental students (n=58) with no experience in game-based learning but widely varying experiences in play and games that had been conducted to obtain perspectives on leisure time and academic education. To make sure that the statement set covered as many aspects of game preferences as possible, we also examined player type studies [17-19,50] that possibly addressed different game preferences. This resulted in an initial set of 136 statements. We grouped the statements into 28 themes, duplicates were removed, and statements were translated into English and reworded to start with the phrase "I like games that…." in order to improve clarity and make sorting more intuitive for participants [41]. The final set (Table 1) consisted of 49 statements and was piloted by 3 medical students. Based on their feedback, we considered the final statement set to meet our abovementioned aims.

Table 1. Statement set and factor array.

Statement	Factors				
	1	2	3	4	5
1. I like games in which people help each other.	1	0	−2	3	−1
2. I like to see how others learn a new game.	−2	0	−1	−3	−3
3. I like games with easy wins.	−3	−3	−1	2	−2
4. I like games which create an atmosphere of sociability.	3	0	−3	1	1
5. I like games that let me build relationships.	2	0	−3	1	0
6. I like games that let me play in teams.	2	−1	0	2	0
7. I like to play games to maintain relationships.	1	−3	−2	2	0
8. I like games that let me play on my own.	−3	3	3	0	−1
9. I like games in which I can create something.	1	1	0	2	2
10. I like games that allow different ways of winning.	2	3	1	2	−2
11. I like games that use luck to enhance my odds of winning.	−1	−3	−1	−1	−4
12. I like games with a good storyline.	1	4	1	3	1
13. I like games in which I can influence the storyline.	0	4	0	0	−1
14. I like games in which I need to actively participate.	3	1	0	4	1
15. I like games in which I know the other players.	2	0	−2	0	3
16. I like games in which I can solve a difficult part / puzzle.	1	2	3	4	3
17. I like to improve my gameplay by searching for new techniques.	0	2	0	−2	−3
18. I like games in which I learn new things (eg, knowledge/skills)	3	2	3	1	0
19. I like games in which I can act differently than I usually do in real life.	−2	1	−2	0	0
20. I like games which make you feel immersed in your own world.	−2	3	−1	−1	4
21. I like games that let me apply a strategy.	4	2	2	0	1
22. I like games in which I can bluff.	−1	−1	−2	0	3
23. I like games that have trading elements.	0	−2	−1	1	2
24. I like games in which I can negotiate.	1	1	−1	0	0
25. I like games that can be played differently than they are intended.	−1	1	−3	−2	2
26. I like games in which I can cheat.	−4	−2	−3	−4	3
27. I like games in which other players cheat.	−4	−4	−4	−4	−3
28. I like games in which I can be fanatic.	3	−2	1	−1	1
29. I like games in which I can play strictly by the rules.	0	−3	1	2	−2
30. I like games to which I can bring modifications.	−2	1	0	−2	0
31. I like games in which I can obtain as many points as possible.	−1	0	2	0	1
32. I prefer losing with lots of rewards over winning with very few.	−3	−1	−1	−1	−2
33. I like games which show me my progression.	2	2	2	1	0
34. I like games which let me have items that others don’t manage to collect.	−2	−2	0	−1	0
35. I like games which have a reward at stake.	0	1	0	−1	−1
36. I like games in which I can get my revenge after losing.	0	−2	1	−1	1
37. I like games that show everyone that I’ve won.	−1	−2	1	−3	−3
38. I like games in which I can prove to the other players that I am the best.	−1	−1	2	−2	−2
39. I like games that use competition as a way to improve myself.	2	0	3	−1	−2
40. I like games in which I can annoy other players.	−1	−4	−2	−3	2
41. I like games that use competition to defeat other players.	1	0	1	−3	1

Statement	Factors				
	1	2	3	4	5
42. I like to be the best in a game.	1	−1	4	−2	2
43. I'm a good loser.	0	−1	−4	0	−4
44. Winning is important to me.	−1	−1	4	−2	2
45. I like games in which I can play alone against a game or computer	−3	3	2	1	−1
46. I like games which let me stay anonymous.	−2	1	1	1	−1
47. I like to get better in a game.	4	2	2	3	3
48. I like games that use a lot of different materials (eg, dices, cards, fake money)	0	0	0	1	−1
49. I like games in which losing is okay.	0	−1	−1	3	−1

Sorting Procedure

Participants performed the sorting procedure using a web app (Q-sorTouch), in which the 49 statements were randomly presented. Participants were asked to drag and drop each statement into 3 piles: agree, neutral, disagree. After sorting all the statements, they had to refine their 3 piles by ranking the statements into a Q-sort grid ranging from −4, extremely disagree, to +4, extremely agree. In Q-methodology, the number of statements that can be assigned to each scale point are fixed and represent a quasi-normal distribution (Table 2) [41]; thus, participants placed the 2 statements with which they disagreed most under −4 and the 2 with which they agreed most under +4.

Table 2. Quasi-normal distribution.

Position	−4	−3	−2	−1	0	+1	+2	+3	+4
Number of items	2	4	6	8	9	8	6	4	2

The sorting procedure ended when all statements were placed in the fixed distribution and the participants felt that the final sort represented their viewpoint. In the final stage of the data collection, participants provided answers to open-ended questions to elaborate on the rationale behind their sort (eg, why statements were assigned to the extreme ends).

Statistical Analysis

To identify groups of participants with shared, but distinct, viewpoints (ie, who subjectively ranked the 49 statements in a similar way), we conducted by-person factor analysis using dedicated software (PQMethod, version 2.35; developer: J Atkinson), which we later verified with formulas [48] in MATLAB (version R2020a; The MathWorks).

Because each sort was correlated with every other sort, the correlation matrix of the participants' sorted statement sets (ie, *sorts*) was used to identify factors (ie, groups of respondents whose Q-sorts were statistically similar) by subjecting the correlation matrix to varimax rotation [41]. Varimax rotation generates a factor solution according to the best mathematical solution (while maintaining an orthogonal basis) [48]. Only factors with eigenvalues >1 and on which at least 2 participants are loaded significantly (P<.01) were accepted [41,48,49], which corresponded to a factor loading >0.37, calculated using 2.58 × (1 / √ (number of items in the Q set) [41,48]. Since our aim was to extract patterns that were unique, participants loading on more than one factor were not used for the construction of a factor. This is in line with the procedures applied in other Q-methodology studies [41,43,44,48].

A range of factor solutions were generated. To describe patterns of the participants' game preferences, each factor solution was interpreted in conjunction with qualitative data from participants' responses in the final stage of the sort. To facilitate factor interpretation, ideal Q-sorts were computed for each factor. These so-called factor arrays are weighted averages of sorts loading on that factor [41,49]. A group of 9 independent researchers individually interpreted all factor solutions and were asked to identify the solution with the highest number of viewpoints while providing distinct and clearly interpretable factors.

Results

Overview

A total of 102 students volunteered to participate in our study and completed the sorting procedure. On the basis of their statements about their favorite games, we identified 7 game genres: action games (n=7), adventure games (n=6), party games (n=13), simulations or sports games (n=15), strategy games (n=35), puzzle games (n=14), and role-playing games (n=10). Consequently, we excluded 60 participants: 2 participants did not provide their favorite game; 10 participants performed the sort in less than 12 minutes; 36 participants had duplicate favorite games (eg, 11 participants stated the game *Settlers of Catan*); 9 participants (4 favorite digital games, 2 females and 3 male students) to ensure a more even distribution of these variables; and 3 participants, by the roll of the dice. The sample consisted of 42 participants (dental students: n=13; medical students: n=29) having 41 different favorite games, of whom 31 were female and 11 were male, with a mean age of 23.3 years (SD 4.0; range 18-42). Of the 42 participants, 15 participants considered themselves to be gamers. Nine sorts were confounded, and 3 sorts did not load significantly on any of the factors (factor loading <0.37; Table 3).

Table 3. Factor matrix.

Q sort	Factor				
	1	2	3	4	5
1	0.383	0.675[a]	0.211	0.182	0.170
2	0.559[a]	0.096	0.089	0.282	0.166
3	−0.041	0.670[a]	0.109	0.284	−0.019
4[b]	0.521	0.522	0.294	0.132	0.312
5	0.578[a]	0.183	0.147	0.031	0.025
6	0.226	0.753[a]	−0.057	0.188	0.048
7	−0.176	0.124	0.584[a]	0.171	0.109
8[c]	0.168	0.364	0.087	0.336	−0.094
9[c]	0.303	−0.296	0.316	0.361	−0.255
10	0.423[a]	0.288	−0.089	0.354	−0.065
11[b]	0.208	0.100	0.521	0.414	−0.334
12	0.501[a]	0.104	0.143	0.160	0.176
13[b]	0.531	0.253	0.535	0.241	−0.045
14	0.344	0.054	0.014	0.526[a]	0.060
15	0.368	0.562[a]	−0.065	0.387	0.110
16[b]	0.402	0.645	0.178	−0.003	0.106
17	−0.113	0.358	0.135	0.639[a]	−0.131
18	0.357	0.166	−0.137	−0.035	0.569[a]
19	0.191	0.703[a]	0.134	0.134	−0.392
20[b]	0.607	0.214	0.073	0.621	0.094
21	0.047	0.686[a]	0.188	−0.011	0.292
22	0.550[a]	0.097	0.038	0.328	−0.293
23	−0.011	0.695[a]	0.308	−0.158	0.055
24	0.708[a]	−0.067	0.237	0.159	0.009
25	0.341	0.175	−0.323	0.556[a]	−0.092
26	0.068	0.380	0.721[a]	0.222	0.131
27[b]	0.658	0.162	−0.202	0.448	−0.175
28	−0.041	0.089	0.246	−0.025	0.598[a]
29	0.320	0.019	0.052	0.479[a]	0.257
30	0.391	0.046	0.276	0.572[a]	−0.045
31[b]	−0.324	0.504	0.439	0.133	0.057
32	0.709[a]	−0.288	0.065	0.265	0.099
33	0.357	0.118	0.658[a]	−0.274	−0.137
34	0.551[a]	0.058	−0.100	0.042	−0.137
35	0.512[a]	0.226	0.223	0.195	−0.064

Q sort	Factor				
	1	2	3	4	5
36	0.508[a]	0.128	−0.086	0.129	0.210
37	0.138	0.064	0.625[a]	−0.030	0.071
38[c]	0.173	0.096	0.190	0.294	−0.299
39[b]	0.199	0.435	0.043	0.073	0.452
40	0.474[a]	0.108	0.055	0.026	0.041
41[b]	0.467	0.327	0.580	−0.106	−0.030
42	0.596[a]	0.272	0.272	0.233	−0.070

[a]A defining sort for a specific factor.

[b]A confounded Qsort (multiple loadings).

[c]A Q-sort with a factor loading <0.37.

Factor Interpretation

Overview

Solutions with up to 5 factors were obtained. The 5-factor solution was retained after analysis by 9 independent researchers because it represented 5 clearly distinguishable patterns in students' perceptions of play and games and had the highest percentage agreement between researchers (88.9%).

Of the 42 included sorts, 30 loaded significantly (factor loading >0.37; Table 3) on 1 of the 5 factors. These patterns are presented below, with sociodemographic information about the participants and a relevant statement to illustrate each pattern. For example, in pattern 1, statement 21 is in the *extremely agree* position (21: +4) in that factor array (Table 1). To give a concise (but oversimplified) overview of the patterns, we chose a descriptor for each that reflected its interpretation in a broad sense.

Social Achiever

Pattern 1 comprised 12 participants with significant factor loadings (female: 9, male: 3; age: mean 23.7 years, range 18-42 years), of whom 7 were medical students, and 5 were dental students. Of the 12 participants, 5 self-identified as gamers. Favorite game genres were strategy (n=5), action (n=3), party (n=2), and simulation or sport games (n=2). Preferred modality was distributed evenly; 6 participants favored analog games, and 6 participants favored digital games.

Participants in Pattern 1 shared the opinion that playing is a social act (4: +3; 5: +3), playing alone or in an individual competition with the self was, therefore, generally disliked (8: −3; 45: −3).

> *What I really like in games is to collaboratively achieve something meaningful.* [Student 3]

The act of social togetherness was not enough for these participants, as they also expressed the need to obtain something meaningful through play (18: +3). Participants loading on this pattern tended to work hard and fanatically toward that goal (28: +3; 3: −3).

Notably, strategy was liked to a great extent (21: +4), which seemed attributable to the fun of being able to play socially and achieving something together (18: +3). Student 67 mentioned,

> *In my opinion, games are way more fun when you play them with friends .... besides, they will give us way longer fun when it is possible to apply a strategy.... This keeps the game interesting and fun for a longer time.*

Explorer

Pattern 2 comprised 7 participants with significant factor loadings (female: 5, male: 2; age: mean 23.1 years, range 20-31 years), of whom 6 were medical students, and 1 was a dental student. Of the 7 participants, 5 participants self-identified as gamers. Favorite games genres were adventure (n=3), role-playing (n=2), action (n=1), and puzzle games (n=1). The majority (n=5) favored a digital modality over an analog modality.

Pattern 2 was characterized by a need for immersion (20: +3), which was especially satisfied through story-driven games (12: +4, 13: +4). Student 21 stated,

> *A good game must drag me into the story and not let go until I am finished.*

These participants generally liked games that granted them substantial autonomy (10: +3; 29: −3) to explore and alter the game (25: +1; 30: +1). They seemed to be drawn to exploring the potential of the game rather than searching for sociability in play (8: +3; 45: +3; 7: −3). Participant 12 stated,

> *For me, gaming is something that I can do primarily on my own.*

These participants played for their own sake or individual pleasure. (6: −1; 7: −3).

Competitor

Pattern 3 comprised 4 participants with significant factor loadings (female: 2, male: 1; age: mean 22.2 years, range 21-23 years). Two participants were medical, and 2 participants were dental students. Two participants self-identified as gamers. Favorite game genres in this group were puzzle (n=2) and simulation or sports games (n=2). Of the 4 participants, 3 students favored a digital modality over an analog modality.

Hunger for competition was the defining aspect of pattern 3 (39: +3)—not only winning or being the best (42: +4; 44: +4), but also parading their supremacy was considered important compared to the other patterns (37: +1; 38: +2). As stated by Student 93,

> *I am very competitive, I want to win every game and I want to show that to everyone.*

Losing was therefore greatly disliked (43: −4). These participants shared the opinion with those described by pattern 2 that other game players were not important to them and they would rather play alone (8: +3); however, whereas participants described by pattern 2 had neither a strong preference nor a dislike for social togetherness as a characteristic of play (4: 0; 5: 0), participants described by pattern 3 found sociability in play unnecessary (4: −3; 5: −3; 7: −2; 15: −2). As Student 23 stated,

> *I play for myself, not for others.*

Thus, *competitors* like competition that does not involve collaboration with others but is directed against other players (since they want to prove they are the best (37: +1; 38: +2)) or a nonplayable character. Student 23 stated,

> *I like to play independently of other players but with an opponent; so, against a computerized opponent.*

Socializer

Pattern 4 comprised 5 participants with significant factor loadings (female: 4, male: 1; mean age 26.2 years; range 2-9 years), of whom 4 were medical students, and 1 was a dental student. Of the 5 participants, 2 participants identified themselves as gamers. Favorite game genres in this group were party (n=2), role-playing (n=2), and action games (n=1). The majority favored an analog modality (n=4) over a digital modality (n=1).

Participants described by Pattern 4 and Pattern 1 had similar characteristics. They valued collaborative play (5: +1; 6: +2); however, whereas being fanatic was important in Pattern 1, in pattern 4, participants did not have the urge to focus on winning (44: −2) or being fanatic (28: −1). They generally disliked competition (41: −3; 42: −2; 44: −2).

> *Winning is not important to me, I just enjoy working together with others and having a good time together.* [Student 68]

This concept of "having a good time" seemed to be a recurrent characteristic for Pattern 4. Games were seen as a means for social togetherness (7: +2; 1: +3) that should depend on nothing but sociability. Losing should be okay (49: +3) and winning should be easy (3: +2); however, participants felt that active participation would be needed to have a good time (14: +4).

Troll

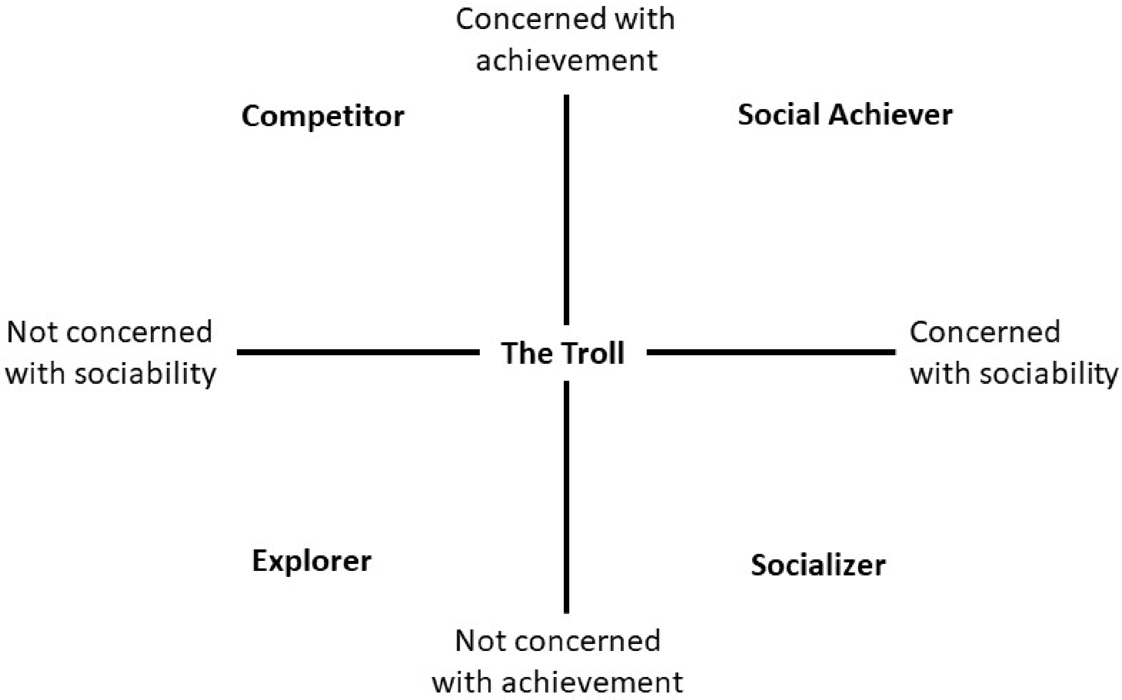
Pattern 5 comprised 2 participants with significant factor loadings (female: 2; age: mean 23.5 years, range 2-5 years), and both were both medical students. One student self-identified as a gamer. One student favored action games, the other student favored simulation or sports games. Both students favored a digital modality (n=2).

Having the ability to exploit game mechanics to cheat (26: +3), annoy other players (40: +3), and bluff (22: +3) was important for these 2 participants compared with participants described by the other patterns. Such behavior seemed to be the result of boredom or laziness and not really being interested in the game itself. Notably, these participants were not inclined to invest time to learn new techniques (17: −3) but, paradoxically, wanted to get better in a game (47: +3), did not like to see others learn the game (2: −3), and were inclined to play games differently than intended when the game would take too much time (25: +2; 29: −2).

> *I like it when a game requires little prior knowledge. It is much simpler and easier to play.* [Student 51]

Figure 2 presents a theoretical framework illustrating different player types in relation to sociability and achievement themes.

Figure 2. Theoretical framework illustrating player types in relation to sociability and achievement themes.



Discussion

Principal Findings

We aimed to improve the understanding of game-based learning design, in general, and of game-based learning in health professional education, in particular. We contended that player typology, a concept that has been used to inform game design and game play, may be relevant to game-based learning design. To the best of our knowledge, this study is the first to investigate player types in a student cohort, outside of a game design context, using a methodology deliberately aimed to accommodate the largest variety in game preferences. We found that, in a cohort of medical and dental students with likely similar academic interests and intellectual ability, there was considerable variability in play preferences. We identified 5 distinct and clearly interpretable patterns in game preferences, which can be considered player types: the social achiever, the explorer, the socializer, the competitor, and the troll. Given that only a few game elements are applied in education for health professionals research—predominantly points and rewards [1]—our findings indicate that there is room for improvement within game-based learning design; the current variety of game-based learning designs in education for health professionals seems too limited to be able to tailor game-based learning to students' game preferences to improve learning through motivation and engagement.

Each player type has distinct characteristics. *Social achievers* like to play collaboratively to achieve something meaningful. They like competition and difficult games, preferably in teams. In contrast, *explorers* are drawn to the game's story and immersive elements of play rather than winning and team play. *Competitors* on the other hand, thrive by winning and competition and would rather not depend on others. *Socializers* (much like *social achievers*) play for sociability, and interaction with other players is important to them. Yet, whereas competition is important to *social achievers*, *socializers* would rather play to find a sense of sociability and togetherness. They look for easy wins just to have a good time. *Trolls* like games in which they can annoy other players, bluff, and cheat.

Two themes (Figure 2) are salient in the player types that we identified, and likewise in scientific and grey literature on play and player typologies [17,19,35,39,51-56]: sociability and achievement. *Competitors* and *social achievers* like the achievement aspect of play, however, *explorers* and *socializers* instead preferred playing for the storyline or to enjoy playing together, respectively. *Social achievers* and *socializers* are driven by sociability, collaboration, and interaction; *explorers* and *competitors*, however, seem less prone to these traits or only need others to prove their supremacy. *The troll* is more ambivalent about sociability and achievement aspects of games than other player types and is, therefore, situated at the intersection of achievement and sociability.

Interestingly, the main themes identified in our theoretical framework bear similarity to the very first and often-cited player typology [18], which was based on a sample with homogeneity in terms of the preferred game, instead of homogeneity in terms of academic interest and which lacked any empirical basis. The fact that our study (which followed a more rigorous scientific approach) resulted in a similar typology may suggest that the existence of player types in a cohort is stable. Our scientific approach led to the identification of the *explorer* and the *socializer*, player types that have also previously been described [18]. We identified the *social achiever*, a player type that seems comparable with that of "achiever [18]." We additionally identified 2 other player type—*the competitor* and *the troll*—whereas in [18] only one other player type, namely the "killer [18]," which, upon close inspection, includes troll-like aspects (eg, annoying other players) as a social component (acting on other players). In our study, these characteristics appeared in other player types. We identified *the troll* and *the competitor* as separate player types. The achiever [18], with its social component was therefore interpreted as *social achiever*. The reason for these differences between both typologies may be that we also included games involving teamplay as a play genre in our framework, since we aimed to avoid selection bias from using only one or a few specific game genres to identify player types. The earlier typology [18] did not include teamplay, probably because it was based on a multiuser dungeon game that included role playing, player versus player, and chat functions but rarely team effort. We also found differences in relation to the dimensions on which the player types varied. Whereas the dimensions *world versus player* and *action versus interaction* have been previously described [18] our empirical evidence supported *achievement* and *sociability* as player type dimensions. As a result, *competitors* and *socializers* were opposites in our framework (instead of "killers" and "explorers" [18]).

The *troll* as a player type has not been identified in previous studies [20]. Remarkably, however, the troll phenomenon is well known in the field of problematic gaming and internet use [57,58]. *Trolling* is defined as deliberately trying to create distress or conflict via provocation, for instance, for the purpose of deception or disruption [58]. More than one-third of American millennials said they engaged in the act of trolling [59] and an immensely popular digital game, called Among Us, is based on the concept of trolling (ie, sabotaging and causing chaos [60]). This suggests that the game-related behavior of trolling is not rare or marginal. Although the relevance of this player type to game-based learning design is unclear, this player type might also be pertinent outside the field of education for health professionals.

Strength and Limitations

The player types in this study represent a broad spectrum of views on games and play. One of the strengths of this study is that the comprehensive set of statements was derived from prior research among medical and dental students [5] and supplemented with statements taken from existing player type studies. Furthermore, a solid scientific method was used to account for all key subjective viewpoints on game preferences, and we included of a variety of participants (independent from game context) to prevent selection bias on game genre. In addition, we discussed multiple factor solutions, sought advice from expert authors [41], and verified Q-methodology software results. In doing so, we added a new perspective to literature on player types and game-based learning by identifying 5 patterns that were distinct, characteristic, and could be considered player types.

This study had some limitations: (1) In the interpretation of our patterns, we cannot (and do not) claim to be exhaustive with respect to all viewpoints on game preferences in the entire population. While Q-methodology is a method that aims to capture variety and heterogeneity, our participant group was relatively homogenous (medical and dental students). Therefore, we cannot claim that replication of our study in a different educational context would yield the same outcomes. However, by adding statements from prior (nonmedical) studies on player types in the statement set, and by using stratification to provide profuse and varied participants' opinions, we feel that the quantitative aspect of the Q-methodology (ie, analyzing participants' rankings using multivariate data reduction techniques) helped us detect meaningful patterns and connections in game preferences. This, in turn, may provide future researchers with a starting point to investigate the generalizability of our results. (2) In a recent study [5], we showed that game elements are possibly context dependent (ie, aspects that motivate play may not necessarily play a motivating role in game-based learning). For instance, although competition was liked and named trivial in play in leisure time, students considered it stressful and unwanted in play focused on learning. Since we did not ask participants to keep a specific learning environment in mind when they answered the question about their game preferences, their answers may not reflect their game-based learning preferences. (3) We aimed to reduce selection bias by selecting participants independent of game context, however, we do not know whether they had a specific game or context in mind when they performed the sorting procedure. (4) We chose to adopt the 5-factor solution after rigorous discussions and with the help of 3 independent researchers. Although this allowed us to detect a new player type (the *troll*), few students had significant factor loadings on this player type. Nevertheless, this player type adhered to the widely accepted rules for including a factor in Q-methodology and helped explain the largest variety in play preferences [58].

By using Q-methodology, we aimed to explain as much variety in existing game preferences as possible; thus, our player types are extreme ends of a spectrum on game preferences. The factor arrays that construct these player types are the combined average of all sort loadings on that player type. Therefore, there is very little chance that a participant's sort will load 100% on a specific player type and fully match its definition [41]. Indeed, all sorts demonstrated characteristics of all player types, and no sort loaded 100% on one player type. Yet, most sorts loaded clearly on one player type.

Practical Implications and Future Research

Systematic reviews indicate that, often, game-based learning strategies are selected based on researchers' personal opinions rather than theory or a conceptual framework [1,2,61]. Additionally, there is a tendency in game-based learning strategies to use scoring and reward, especially in gamification [1,61]. Our taxonomy provides a novel theoretical framework that may help to tailor game-based learning strategies to student preferences. Future research is needed to investigate whether such tailoring would result in increased effectiveness of applying game-based learning in education.

Based on our findings, all player types except *explorers* might need the presence or participation of other players to be optimally motivated to continue playing. To develop game-based learning strategies that optimally engage and motivate the majority of students, multiplayer options appear to be critical.

However, this feature is currently overlooked in game-based learning strategies in current practice [1,11,61].

Our theoretical framework and corresponding factor arrays indicate that preferences for multiplayer modalities can be diverse and are not limited to sociability [62], social media [63], a chat function [34], and message boards [64]. *Competitors*, for instance, need other players or computerized opponents to triumph over and show their supremacy, *social achievers* need other players to work with, *trolls* need other players to annoy, and *socializers* need other players to have a good time together. By including each player type in a game-based learning-strategy, the complex and dynamic interaction between player types can turn game-based learning into a meaningful strategy for every student. For example, although *trolls* might only make a small contribution to the overall player population, their actions can have major impact on social play and interaction [65-68], much more than, for example, the actions of *social achievers*. The inclusion of *trolls* in game-based learning design can unite socially oriented players by giving them a common foe. Future research should explore how each player type can contribute to multiplayer game-based learning strategies to enhance collaborative learning.

Future research can focus on investigating whether the range of opinions on play vary significantly across students as a function of the academic level or discipline they are enrolled in, for instance, a medical or a nonmedical group, or medical specialization. Such findings would provide an understanding for future student-specific game-based learning designs. Game preferences might be dependent on context [5] or the players' current needs [31]. For instance, in the playground game called Tag, one player is *it* and chases the other players in an attempt to tag them by touching them. Then the tagged player becomes *it* and starts chasing the others to tag someone else. This means that, when being *it*, a player must adopt the *competitor* player type (ie, competing and winning from the others), while the others (who are getting chased) can adopt the *social achiever* or even *troll* player type to act as a group against the one that is *it*. Likewise, other digital games (eg, Among Us) perhaps also use changing player types, where one is sometimes a troll and, at other times, needs to take on the role of the social achiever [60]. This raises the question whether player types are in search of a specific game design or does the game design elicit different types of behavioral responses (ie, player types). This might also suggest that game designers should adhere to the entire diversity of player types to ensure inclusion of all participants of the game-based learning strategy.

As a first step in this direction, we aimed to investigate the prevalence of player types among medical and dental students. This may not only provide more evidence for the existence of the currently identified typology in education for health professional students, it may also shed light on the true diversity of player types within medical and dental education. Furthermore, it may improve our understanding of whether the current educational strategy focusing on the achievement-oriented player type is effective and can be justified or whether it might be better to tailor game-based learning strategies to individual player types.

Conclusion

We identified 5 clear and distinct patterns of game preferences. These patterns represent player types that differ in terms of the player type dimensions *achievement* and *sociability*. Our taxonomy and accompanying factor arrays can be used to tailor game-based learning design to students' game preferences to optimize game-based learning effectiveness.

Acknowledgments

The authors would like to thank the students for participating in this study. We also thank Tineke Bouwkamp-Timmer for editorial assistance. We thank Kiki Spanjers, Gerben Ruesink, and Anne-Marijke Kosta, for their valuable input in choosing the optimal pattern solution, and Irena Middeljans, for her valuable input during statement selection. Last but certainly not least, we are grateful for the helpful email correspondence with Simon Watts and Paul Stenner.

Conflicts of Interest

None declared.

References

1. van Gaalen AEJ, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma ADC, Georgiadis JR. Gamification of health professions education: a systematic review. Adv Heal Sci Educ 2021:683-711 [FREE Full text] [doi: 10.1007/s10459-020-10000-3] [Medline: 33128662]
2. Dichev C, Dicheva D. Gamifying education: what is known, what is believed and what remains uncertain: a critical review. Int J Educ Technol High Educ 2017 Feb 20;14(1). [doi: 10.1186/s41239-017-0042-5]
3. Bedwell WL, Pavlas D, Heyne K, Lazzara EH, Salas E. Toward a taxonomy linking game attributes to learning: an empirical study. Simul Gaming 2012 May 14;43(6):729-760. [doi: 10.1177/1046878112439444]
4. Marczewski A. 52 gamification mechanics and elements. Gamified UK. 2017. URL: https://www.gamified.uk/user-types/gamification-mechanics-elements/ [accessed 2022-02-25]
5. Van Gaalen AEJ, Jaarsma ADC, Georgiadis JR. Medical students' perceptions of play and learning: qualitative study with focus groups and thematic analysis. JMIR Serious Games 2021 Jul 28;9(3):e25637 [FREE Full text] [doi: 10.2196/25637] [Medline: 34319237]

6. Papert S. Does easy do it? children, games, and learning. Game Developer 1998;88:1 [FREE Full text]
7. DiSalvo B. Pink boxes and chocolate-dipped broccoli: bad game design providing justifications for reluctant learners. 2015 Presented at: 11th Games + Learning + Society Conference; July 7-10; Chicago, Illinois, USA p. 11.
8. Van ECK R. Digital Game-Based Learning: It's Not Just the Digital Natives Who Are Restless. EDUCAUSE Review 2006 Apr:1-16 [FREE Full text]
9. Landers RN, Bauer KN, Callan RC, Armstrong MB. Psychological theory and the gamification of learning. In: Gamification in Education and Business. New York: Springer; 2015:86.
10. Böckle M, Novak J, Bick M. Towards adaptive gamification: a synthesis of current developments. 2017 Presented at: 25th European Conference on Information Systems; June 5-10; Guimarães, Portugal.
11. Dicheva D, Dichev C, Agre G, Angelova G. Gamification in education: A systematic mapping study. Educ Technol Soc 2015:1-14 [FREE Full text]
12. Monterrat B, Lavoué É, George S. Adaptation of gaming features for motivating learners. Simul Gaming 2017 Jun 30;48(5):625-656. [doi: 10.1177/1046878117712632]
13. Harviainen J. Critical challenges to gamifying education: A review of central concept. 2014 Presented at: Game On! Conference; 03-17-2014; Moscow.
14. Burgers C, Eden A, van Engelenburg MD, Buningh S. How feedback boosts motivation and play in a brain-training game. Comput Human Behav 2015 Jul;48:94-103. [doi: 10.1016/j.chb.2015.01.038]
15. Chen C, Chang S. An exploration of the tendency to online game addiction due to users liking of design features. Asian J Heal Inf Sci 2008:38-51 [FREE Full text]
16. Robinson M, Fischer A, Ahuja A, Lesser E, Maniates H. Roles of "wanting" and "liking" in motivating behavior: gambling, food, and drug addictions. Curr Top Behav Neurosci 2016;27:105-136. [doi: 10.1007/7854_2015_387] [Medline: 26407959]
17. Yee N. Motivations for play in online games. Cyberpsychol Behav 2006 Dec;9(6):772-775. [doi: 10.1089/cpb.2006.9.772] [Medline: 17201605]
18. Bartle R. Hearts, clubs, diamonds, spades: players who suit MUDs. Research Gate. 1996. URL: https://www.researchgate.net/publication/247190693_Hearts_clubs_diamonds_spades_Players_who_suit_MUDs [accessed 2022-02-25]
19. Yee N, Ducheneaut N, Nelson L. Online gaming motivations scale: Development and validation. 2012 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 05/12/2012 - 05/10/2012; Texas. [doi: 10.1145/2207676.2208681]
20. Klock ACT, Gasparini I, Pimenta MS, Hamari J. Tailored gamification: a review of literature. Int J Hum Comput Stud 2020 Dec;144:102495. [doi: 10.1016/j.ijhcs.2020.102495]
21. Codish D, Ravid G. Detecting playfulness in educational gamification through behavior patterns. IBM J Res Dev 2015 Nov;59(6):1-14. [doi: 10.1147/jrd.2015.2459651]
22. Codish D, Rabin E, Ravid G. User behavior pattern detection in unstructured processes – a learning management system case study. Interact Learn Env 2019 May 19;27(5-6):699-725. [doi: 10.1080/10494820.2019.1610456]
23. Codish D, Ravid G. Personality Based Gamification. Educational Gamification for Extroverts and Introverts. 2014 Presented at: 9th Chais Conference for the Study of Innovation and Learning Technologies: Learning in the Technological Era; 02/11/2014 - 02/12/2014; Israel.
24. Codish D, Ravid G. Personality based gamification: How different personalities percive gamification. 2014 Presented at: 22nd European Conference on Information Systems; 06/09/2014 - 06/11/2014; Israel.
25. Codish D, Ravid G. Gender Moderation in Gamification: Does One Size Fit All? 2017 Presented at: 50th Hawaii International Conference on System Sciences; 01/04/2017 - 01/07/2014; Hawaii. [doi: 10.24251/hicss.2017.244]
26. Denden M, Tlili A, Essalmi F, Jemni M. An investigation of the factors affecting the perception of gamification and game elements. 2017 Presented at: International Conference on Information and Communication Technology and Accessbility, ICTA; Dec 19-21; Muscat. [doi: 10.1109/icta.2017.8336019]
27. Tlili A, Essalmi F, Ayed LB, Jemni M. A smart educational game to model personality using learning analytics. 2017 Presented at: IEEE 17th International Conference on Advanced Learning Technologies; July 3-7; Romania. [doi: 10.1109/icalt.2017.65]
28. Buckley P, Doyle E. Individualising gamification: an investigation of the impact of learning styles and personality traits on the efficacy of gamification using a prediction market. Comput Educ 2017 Mar;106:43-55. [doi: 10.1016/j.compedu.2016.11.009]
29. Davis K, Sridharan H, Koepke L, Singh S, Boiko R. Learning and engagement in a gamified course: investigating the effects of student characteristics. J Comput Assist Learn 2018 Apr 19;34(5):492-503. [doi: 10.1111/jcal.12254]
30. Pedro L, Lopes A, Prates B, Vassileva J, Isotani S. Does gamification work for boys and girls? An exploratory study with a virtual learning environment. 2015 Presented at: ACM Symposium on Applied Computing; April 13-17; Spain. [doi: 10.1145/2695664.2695752]
31. Tu C, Yen C, Sujo-Montes L, Roberts GA. Gaming personality and game dynamics in online discussion instructions. Educ Media Int 2015 Sep 11;52(3):155-172. [doi: 10.1080/09523987.2015.1075099]
32. Charles D, Kerr A, McNeill M. Player-centred game design: Player modelling and adaptive digital games. 2005 Presented at: Digital Games Research Conference; 16/06/2005 - 20/06/2005; Vancouver, British Columbia, Canada.

33. Barata G, Gama S, Jorge J, Goncalves D. Early prediction of student profiles based on performance and gaming preferences. IEEE Trans Learning Technol 2016 Jul 1;9(3):272-284. [doi: 10.1109/tlt.2016.2541664]
34. Klock A, da Cunha LF, de Carvalho MF, Rosa B, Anton A, Gasparini I. Gamification in e-learning systems: a conceptual model to engage students and its application in an adaptive e-learning system. In: Zaphiris P P, Ioannou A, editors. Learning and Collaboration Technologies. Cham: Springer; 2015:9192.
35. Nacke LE, Bateman C, Mandryk RL. BrainHex: a neurobiological gamer typology survey. Entertain Comput 2014 Jan;5(1):55-62. [doi: 10.1016/j.entcom.2013.06.002]
36. Borges S, Mizoguchi R, Durelli V, Bittencourt I, Isotani S. A link between worlds: towards a conceptual framework for bridging player and learner roles in gamified collaborative learning contexts. In: Koch F, Koster A, Primo T, Guttmann C, editors. Advances in Social Computing and Digital Education. Cham: Springer; 2016.
37. Marczewski A. Even Ninja Monkeys Like to Play: Gamification, Game ThinkingMotivational Desig. In: Gamified UK. San Fransisco: BLURB Incorporated; 2015.
38. Ferro L, Walz S, Greuter S. Towards personalised, gamified systems: An investigation into game design, personality and player typologies. 2013 Presented at: ACM International Conference Proceeding Series; 10/6/2013 - 10/9/2013; United Kingdom. [doi: 10.1145/2513002.2513024]
39. Yee N. The demographics, motivations, and derived experiences of users of massively multiuser online graphical environment. Presence (Camb) 2006 Jun 01;15(3):309-329. [doi: 10.1162/pres.15.3.309]
40. Danielson S. Q method and surveys: three ways to combine Q and R. Field Methods 2009 Apr 06;21(3):219-237. [doi: 10.1177/1525822x09332082]
41. Watts S, Stenner P. Doing Q methodology: theory, method and interpretation. Qual Res Psychol 2005. [doi: 10.1191/1478088705qp022oa]
42. Brown SR, Danielson S, van Exel J. Overly ambitious critics and the Medici effect: a reply to Kampen and Tamás. Qual Quant 2014 Feb 28;49(2):523-537. [doi: 10.1007/s11135-014-0007-x]
43. Berkhout JJ, Teunissen PW, Helmich E, van Exel J, van der Vleuten CPM, Jaarsma DADC. Patterns in clinical students' self-regulated learning behavior: a Q-methodology study. Adv Health Sci Educ Theory Pract 2017 Mar 27;22(1):105-121 [FREE Full text] [doi: 10.1007/s10459-016-9687-4] [Medline: 27235123]
44. Fokkema J, Teunissen P, Westerman M, van der Lee N, van der Vleuten CPM, Scherpbier AJJA, et al. Exploration of perceived effects of innovations in postgraduate medical education. Med Educ 2013 Mar;47(3):271-281. [doi: 10.1111/medu.12081] [Medline: 23398013]
45. Meade L, Caverzagie K, Swing S, Jones R, O'Malley CW, Yamazaki K, et al. Playing with curricular milestones in the educational sandbox: Q-sort results from an internal medicine educational collaborative. Acad Med 2013 Aug;88(8):1142-1148. [doi: 10.1097/ACM.0b013e31829a3967] [Medline: 23807106]
46. Watts S, Stenner P. Factor interpretation. In: Doing Q Methodological Research: Theory, Method and Interpretation. Los Angeles: SAGE Publications; 2012:147-169.
47. Solomon H, Stephenson W. The study of behavior: Q-technique and its methodology. J Am Stat Assoc 1955 Dec;50(272):1415. [doi: 10.2307/2281274]
48. Brown SR. Political subjectivity: applications of Q methodology in political science. J Politics 1981 Aug;43(3):912-914. [doi: 10.2307/2130649]
49. Stenner PH, Cooper D, Skevington SM. Putting the Q into quality of life; the identification of subjective constructions of health-related quality of life using Q methodology. Soc Sci Med 2003 Dec;57(11):2161-2172. [doi: 10.1016/s0277-9536(03)00070-4]
50. Sherry J, Greenberg B, Lucas K, Lachlan K. Video game uses and gratifications as predictors of use and game preference. In: Playing Video Games: Motives, Responses, and Consequences. New York: Routledge; 2006:213-224.
51. Hamari J, Tuunanen J. Player types: a meta-synthesis. Trans Digit Games Res Assoc 2014 Mar 24;1(2). [doi: 10.26503/todigra.v1i2.13]
52. Tseng F. Segmenting online gamers by motivation. Expert Syst Appl 2011 Jun;38(6):7693-7697. [doi: 10.1016/j.eswa.2010.12.142]
53. Whang L, Chang G. Lifestyles of virtual world residents, living in the on-line game "Lineage". Cyberpsychol Behav 2004:592-600. [doi: 10.1109/cyber.2003.1253430]
54. Fullerton T. A playcentric approach to creating innovative games. In: Game Design Workshop. New York: CRC Press; 2008.
55. Berger BM, Caillois R, Barash M. Man, Play, and Games. Am Sociol Rev 1963 Aug;28(4):651. [doi: 10.2307/2090095]
56. Huizing J. Homo Ludens: A Study of the Play-Element in Culture. New York: Angelico Press; Apr 1951.
57. Jones D, Paulhus D. Differentiating the dark triad within the interpersonal circumplex. In: Handbook of Interpersonal Psychology: Theory, Research, Assessment, and Therapeutic Interventions. New Jersey: John Wiley & Sons; 2012.
58. Buckels E, Trapnell P, Paulhus D. Trolls just want to have fun. Personality and Individual Differences 2014 Sep;67:97-102. [doi: 10.1016/j.paid.2014.01.016]
59. Over a quarter of Americans have made malicious online comments. YouGov. URL: https://today.yougov.com/topics/politics/articles-reports/2014/10/20/over-quarter-americans-admit-malicious-online-comm [accessed 2022-02-25]

60. Among us. Wikipedia. URL: https://en.wikipedia.org/wiki/Among_Us [accessed 2022-02-25]
61. 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]
62. Christou G, Lai-Chong Law E, Zaphiris P, Ang CS. Challenges of designing for sociability to enhance player experience in massively multi-player online role-playing Games. Behav Inf Technol 2013 Jul;32(7):724-734. [doi: 10.1080/0144929x.2012.754497]
63. Petrucci AM, Kaneva P, Lebedeva E, Feldman LS, Fried GM, Vassiliou MC. You have a message! social networking as a motivator for FLS training. J Surg Educ 2015 May;72(3):542-548. [doi: 10.1016/j.jsurg.2014.10.007] [Medline: 25496795]
64. Wongso O, Rosmansyah Y, Bandung Y. Gamification framework model, based on social engagement in e-learning. 2014 Presented at: 2nd International Conference on Technology, Informatics, Management, Engineering and Environment; December 19-21; Indonesia. [doi: 10.1109/time-e.2014.7011583]
65. Bishop J. The psychology of trolling and lurking: the role of defriending and gamification for increasing participation in online communities using seductive narratives. In: Gamification for Human Factors Integration: Social, Education, and Psychological Issues. Pennsylvania: IGI Global; 2015:160-176.
66. Ruiz C, Domingo D, Micó JL, Díaz-Noci J, Meso K, Masip P. Public sphere 2.0? the democratic qualities of citizen debates in online newspapers. Int J Press Polit 2011 Sep 22;16(4):463-487. [doi: 10.1177/1940161211415849]
67. Shachaf P, Hara N. Beyond vandalism: Wikipedia trolls. J Inf Sci 2010 Apr 13;36(3):357-370. [doi: 10.1177/0165551510365390]
68. Thacker S, Griffiths M. An exploratory study of trolling in online video gaming. Int J Cyber Behav Psychol Learn 2012:17-33. [doi: 10.4018/ijcbpl.2012100102]

Edited by N Zary; submitted 15.05.21; peer-reviewed by L Roper, S Azer; comments to author 03.08.21; revised version received 15.09.21; accepted 24.09.21; published 04.04.22.

Please cite as:
Van Gaalen AEJ, Schönrock-Adema J, Renken RJ, Jaarsma ADC, Georgiadis JR
Identifying Player Types to Tailor Game-Based Learning Design to Learners: Cross-sectional Survey using Q Methodology
JMIR Serious Games 2022;10(2):e30464
URL: https://games.jmir.org/2022/2/e30464
doi:10.2196/30464
PMID:35377333

Original Paper

Using Virtual Reality to Improve Classroom Behavior in People With Down Syndrome: Within-Subjects Experimental Design

Stefan Carlo Michalski[1], BPsych(Hons); Ancret Szpak[1], PhD; Caroline Ellison[1], PhD; Rowena Cornish[2], BSc; Tobias Loetscher[1], PhD

[1]UniSA Justice and Society, University of South Australia, Adelaide, Australia

[2]Orana Australia Ltd, Adelaide, Australia

Corresponding Author:
Stefan Carlo Michalski, BPsych(Hons)
UniSA Justice and Society
University of South Australia
Magill Campus
GPO Box 2471
Adelaide, 5001
Australia
Phone: 61 088302 ext 2611
Email: stefan.michalski@mymail.unisa.edu.au

Abstract

Background: People with Down syndrome face various learning challenges. Introducing new and enjoyable experiences in learning settings may improve learning outcomes. Immersive and interactive technologies such as virtual reality can be used to deliver rich visual experiences in classrooms.

Objective: The aim of this study was to investigate the feasibility and benefits of virtual reality exposure for people with Down syndrome in learning settings.

Methods: To address this aim, we used a within-subjects design to assess the effect of a brief virtual reality drawing experience and conventional drawing experience on subsequent behavior in 16 participants.

Results: Large positive effects were found for virtual reality drawing (t_{15}=5.020, P<.001) and conventional drawing (t_{15}=3.720, P=.002) in improving subsequent behavior in a learning setting. Irrespective of the intervention, the participant's mood, attention, and overall behavior significantly improved. No significant differences were found between the interventions (t_{15}=–0.648; P=.53).

Conclusions: This study's results are encouraging for researchers and educators interested in using virtual reality for people with Down syndrome, as virtual reality was found to be highly feasible. Recommendations are made for researchers and educators interested in providing virtual reality experiences for people with Down syndrome.

(JMIR Serious Games 2022;10(2):e34373) doi:10.2196/34373

KEYWORDS

virtual reality; Down syndrome; intellectual disability; drawing; art; behavior; mood; attention; classroom; self-report

Introduction

People with Down syndrome often encounter significant challenges in learning settings [1]. Inattentiveness, impulsive behavior, excessive fidgeting, and other nondirected motor activity are signs of distress and stereotypical behaviors for people with Down syndrome [2]. Immersive technologies such as virtual reality (VR) hold great potential in delivering enjoyable and therapeutic experiences [3]. VR is commonly being considered by researchers and educators to provide safe access to realistic experiences that may otherwise be logistically difficult, dangerous, or impractical to implement [4,5]. Exposure to VR settings that provide a sense of distance from routine can reduce stress and improve mood [6]. There is encouraging evidence supporting the use of nonimmersive virtual environments to provide useful learning [7], rehabilitation [8], and leisure experiences for people with intellectual disabilities. Therefore, immersive VR applications may also be an effective way to improve motivation and engagement for people with Down syndrome in learning settings.

Applications of VR are proliferating in psychology, health care, and education [9]. Weiss et al [10] and Yalon-Chamovitz and Weiss [11] were few of the first researchers to study the use of VR to improve the leisure experiences of people with intellectual disabilities. In these experiments, virtual environments were found to increase enjoyment and engage participants with cerebral palsy and an intellectual disability [10,11]. Flat-screen displays were used to deliver the experience, offering a low level of immersion. Given these positive early results in people with intellectual disabilities, it is surprising that there is limited research investigating the use of modern and more immersive virtual environments such as head-mounted displays (HMDs). Modern technology offers new opportunities, and there is reason to believe that more specific, realistic, and engaging VR applications may be useful for people with intellectual disabilities.

HMDs may be preferred over less-immersive displays as it can provide rich visual experiences that elicit greater feelings of presence in the user [12]. However, implementing immersive technology in vulnerable or VR-inexperienced groups needs to be carefully introduced and monitored for possible negative experiences specific to that population. People with Down syndrome, for example, have structural eye abnormalities, which may diminish their vision even when corrected [13]. There is a high prevalence of near vision impairments in people with Down syndrome, and 25%-60% have strabismus [13], which will affect their visual perception when using stereoscopic displays. A convergence insufficiency (inability to maintain binocular function) will likely lead to difficulty seeing depth in VR, which may also increase the likelihood of a negative VR experience [14,15].

People with Down syndrome may be predisposed to experience cybersickness. Cybersickness has been related to a visual-vestibular mismatch between VR and the real world, leading to symptoms such as nausea or disorientation [16]. Vergence-accommodation conflicts may also exacerbate oculomotor symptoms such as eye strain and fatigue [14,17]. It is unclear how enjoyable VR is for people with eye abnormalities, given many typical users still experience adverse effects. In addition, people with Down syndrome typically have difficulties with fine motor skills due to low muscle tone and joint hypermobility [18], which may present challenges when interacting in virtual environments. Thus, it is unclear how suitable using a headset and a handheld controller is. A thorough investigation of users' experiences with HMDs is critical.

Self-report measures are typically used to assess VR aftereffects, though such measures must be interpreted with caution in this population. Widely reported in the literature are concerns that people with intellectual disabilities tend to positively self-report or overestimate their responses [19,20]. For example, Yalon-Chamovitz and Weiss [11] found in their study on young adults with cerebral palsy and moderate intellectual disability that self-reported success and enjoyment in VR significantly differed from staff observations. Researchers and educators alike encounter significant challenges in obtaining valid self-reports from participants with intellectual disabilities owing to challenges in communication and comprehension, especially among nonverbal participants [21]. Utilizing methods that do not use complex language, such as observation, may prove to be more effective when assessing behavior in learning settings.

The aim of this study was to investigate the feasibility and benefits of VR exposure for people with Down syndrome. To address this aim, we explored the effect of a brief VR drawing experience and conventional drawing experience on subsequent behavior in a learning setting. Drawing was selected as it is a familiar activity, and participants could engage in free play.

Methods

Design

This study contrasted 2 drawing activities in a within-subjects design: drawing in a VR application (Tilt Brush, developed by Google) and conventional drawing. The researcher removed the participants from the learning setting to complete each activity. Once completed, the participants returned to the learning setting for observation. Participants were required to wait a minimum of 24 hours before completing their second activity (counterbalanced order).

Participants

Seventeen people (mean age 25.25 [SD 6.61] years) diagnosed with Down syndrome were recruited from a nonprofit disability services organization in South Australia. One participant was excluded as they were unable to complete the VR experience. Thus, 16 participants were included in the analyses (7 females and 9 males). All participants in this study were considered to have a severe-to-profound intellectual disability. The severity of the intellectual disability was classified on the ability to perform daily skills as per the Diagnostic and Statistical Manual of Mental Disorders, fifth edition criteria [22]. All participants attended a program that aimed to improve life skills in young adults with Down syndrome. Class sizes varied, though the number of clients never exceeded 12 per session. At a minimum, 1 support staff member was present per 4 clients. Informed consent was obtained from the participant, caregiver, and a staff member at the organization. An easy-to-read consent form with pictures was developed to ensure that participants clearly understood what was involved in the study. The appropriate sample size was calculated using the G*Power 3 software [23]. Yalon-Chamovitz and Weiss [11] found a large effect size for perceived level of enjoyment in a VR leisure activity in people with physical and intellectual disabilities. Using a large effect size (0.80) as an estimate for the power analysis, it was calculated that for 1-sample 1-tailed *t* tests, at least 15 participants would be needed to suffice power with α=.05.

Ethics Approval

This study was granted ethics approval from the University of South Australia Human Research Ethics Committee (202640).

Materials

VR Apparatus

The Oculus Quest (developed by Facebook Technologies, LLC) was used. Immersive HMDs such as the Oculus Quest enable users to view a 3D environment that moves in real time following their movements. Users were required to hold a controller to interact within the environment while wearing the HMD. Corrective glasses were worn in the device if needed.

VR Application

Tilt Brush (developed by Google) was used. Users were immersed in a 360° virtual environment where they could paint and observe their artwork in 3D space. A controller was used to simulate a paintbrush in the virtual environment.

Conventional Drawing

Participants were provided a blank A4 paper and their favorite color pencil.

Measures

Learner Behavior

Learner behavior data were collected from support staff at the disability services organization. Staff observed participants and provided ratings for changes on the following 6 factors (proceeding examples were also listed): (1) mood, (2) attention (eg, listening to instructions, not distracted, not looking around), (3) activity (eg, jumping out of the seat, walking around class inappropriately), (4) impulses (eg, blurting out answers before questions completed, interrupting others, butt into conversations, failing to wait turn), (5) anxiety (eg, fidgeting, bite hands/nails, twitch, pace, shake, hand/feet tapping, tense expression), and (6) withdrawal (eg, staring blankly, daydream, fiddling with objects, detached). Fifteen minutes after returning to the learning setting, 2 staff rated changes in behavior on a 7-point scale ranging from better (+3) to worse (–3), with 0 being no change. The 2 rater scores were averaged. Individual subscales and the total scores (sum of all subscales) were analyzed. The staff remained blinded to which intervention the participant completed (ie, VR drawing or conventional drawing). Further, staff were asked to report any noticeable observations in the participant's behavior (ie, reports of sickness or suspicion of the activity completed). The learner behavior form was adapted from Part 2 of Mather and Jaffe's [24] classroom behavior form, which was designed to observe problem behaviors in a classroom. This measure demonstrated good internal consistency in this study (Cronbach α=.72). The learner behavior form used in this study is available upon request.

Drawing Time

The number of minutes participants opted to remain in each drawing activity was recorded.

Cybersickness

The researcher asked participants if they were feeling dizzy or sick. Specifically, the researcher handed the participants a sheet that stated, "I felt dizzy or sick…" Below this statement, there were 3 response options: no, not sure, or yes. Each option had an emoticon underneath, a smiley face, confused face, and a nauseous face, respectively. Participants were required to select an option by either circling the response on paper or by saying the word aloud. The researcher guided the participants through the question to ensure comprehension.

Choice Paradigm

The researcher asked questions regarding preference of 3 different activities: drawing in VR, drawing on paper, and watching TV. Questions were phrased in 3 different ways:

1. Single-choice question: participants provided a yes or no response to each activity individually, indicating whether they enjoy the activity and would like to do it again in the future. The researcher asked, for example, "Do you like drawing in virtual reality?"
2. Paired-choice question: participants were asked to select their favorite activity out of 2 options. The researcher asked, for example, "What do you like better: drawing in virtual reality or drawing on paper?" The paired choice was completed when participants responded to each of the 3 pair combinations. Based on the responses, their favorite item was determined.
3. Multiple-choice question: participants were asked to select their overall favorite out of the 3 activities. Specifically, the researcher asked, "What is your favorite activity out of drawing in virtual reality, drawing on paper, or watching TV?"

Procedure

In a counterbalanced order, participants completed 2 interventions: VR drawing and conventional drawing. Participants were not given specific instructions on what to draw, as the researcher indicated they had free time. The range of options in the VR drawing was replicated in the conventional drawing as best as possible. For example, in-game sounds and effects were removed. Furthermore, in both conditions, the researcher asked participants for their favorite color and that was the only color used. Participants held only 1 controller and 1 pencil in each experience. Similarly, participants did not have erasers, and both conditions were completed on a neutral background.

Participants were instructed that 7 minutes had been allocated to each activity. Once 7 minutes elapsed, the researcher asked the participants if they would like to continue for an additional minute or stop entirely. If participants opted to continue, this process was repeated at the end of each minute until a maximum of 10 minutes was reached. Although participants were informed that 7 minutes had been allocated to the activity, they were reminded they could withdraw at any time. After both interventions, the researcher asked the participants if they were feeling dizzy or sick and recorded notes.

The choice paradigm was completed following the second intervention, where the researcher asked the single-choice, paired-choice, and multiple-choice questions. Physical props were used (ie, VR headset, paper and pencils, and an image of a TV) to ensure participants understood the questions.

Following completion of each activity, learner behavior data were collected. Two assigned staff members were asked to observe the participant for 15 minutes upon returning to the learning setting. Then, the staff members were asked to complete the learner behavior form by providing independent ratings based on their observations. Staff were asked to note if they had a suspicion what activity the participant completed. Staff carried on with the regular class routine, which means they may have worked with multiple clients simultaneously.

Results

Feasibility

Twenty-two people were invited to participate, but only 17 agreed. VR exposure was highly feasible in this sample, with 16 out of 17 participants able to complete the VR activity. One participant was excluded as they did not engage in the task and were nonresponsive; therefore, the researcher opted to discontinue the experience. Two out of 16 participants (13%) elected to end the VR experience before the allocated time expired. Five out of 16 participants (31%) reported cybersickness symptoms after VR exposure. The researcher followed up on the participants' symptoms: 3 reported eye strain, 1 reported dizziness, and 1 was unable to provide further information. All symptoms reported were mild and short-lived, as there was no evidence of discomfort during observations (15 minutes after the exposure). A paired-sample *t* test revealed no difference of drawing time between VR drawing (mean 8.0 [SD 2.28] minutes) and conventional drawing (mean 7.5 [SD 2.03] minutes; t_{15}=0.6; *P*=.54; Cohen *d*=0.16).

Learner Behavior

A series of paired-samples *t* tests were conducted to assess whether there were significant differences in learner behavior following the VR drawing and conventional drawing interventions. Paired-samples *t* tests revealed no significant differences between the 2 interventions (Table 1).

A series of 1-sample *t* tests were conducted to assess whether the interventions changed behavior from 0 (representing no change, Table 2). One-sample *t* tests revealed that the total score of learner behavior was significantly different from 0 for VR drawing (*P*<.001) and conventional drawing (*P*=.002) interventions. Mood and attention scores were also significantly different from 0 after both interventions. Notably, the effect sizes were large following the VR intervention and bigger in comparison to that following conventional drawing. Furthermore, activity scores increased after VR, while withdrawal scores increased after conventional drawing. Nonsignificant differences were found in the remaining variables.

Table 1. Learner behavior differences between virtual reality drawing and conventional drawing interventions.

Variable	Mean (SD)	*t* value *(df)*[a]	*P* value	Cohen *d*[b]
Mood	0.61 (0.5)	–0.131 (15)	.89	0.03
Attention	0.66 (0.4)	–0.324 (15)	.75	0.08
Activity	0.17 (0.2)	0.863 (15)	.40	0.21
Impulses	0.22 (0.3)	–0.307 (15)	.76	0.07
Anxiety	0.19 (0.3)	–1.103 (15)	.29	0.27
Withdrawal	0.23 (0.3)	–1.454 (15)	.17	0.36
Total[c]	2.08 (1.4)	–0.648 (15)	.53	0.16

[a]Instances of negative *t* values indicate higher scores in conventional drawing as compared to those in virtual reality drawing.

[b]Cohen *d* effect size interpretation: 0.2=small effect size, 0.5=medium effect size, and 0.8=large effect size.

[c]Total score indicates the sum of all subscales.

Table 2. Learner behavior differences for each variable in each intervention.

Variable, intervention	Mean (SD)[a]	*t* value *(df)*	*P* value	Cohen *d*[b]
Mood				
Virtual reality drawing	0.59 (0.5)	4.842 (15)	*<.001* [c]	1.21
Conventional drawing	0.63 (0.8)	2.953 (15)	*.01* [c]	0.74
Attention				
Virtual reality drawing	0.63 (0.5)	5.371 (15)	*<.001* [c]	1.34
Conventional drawing	0.69 (0.6)	4.198 (15)	*<.001* [c]	1.05
Activity				
Virtual reality drawing	0.25 (0.4)	2.449 (15)	*.03* [c]	0.61
Conventional drawing	0.09 (0.4)	0.899 (15)	.38	0.22
Impulses				
Virtual reality drawing	0.19 (0.4)	1.861 (15)	.08	0.46
Conventional drawing	0.25 (0.6)	1.732 (15)	.10	0.43
Anxiety				
Virtual reality drawing	0.09 (0.4)	1.000 (15)	.33	0.25
Conventional drawing	0.28 (0.5)	2.058 (15)	.06	0.51
Withdrawal				
Virtual reality drawing	0.09 (0.3)	1.379 (15)	.19	0.34
Conventional drawing	0.38 (0.6)	2.423 (15)	*.03* [c]	0.61
Total[d]				
Virtual reality drawing	1.84 (1.5)	5.020 (15)	*<.001* [c]	1.25
Conventional drawing	2.31 (2.5)	3.720 (15)	*.002* [c]	0.93

[a]A mean score of 0 represents no change. Higher scores reflect better behavior. If the blinded staff members guessed the activity a participant completed, they were correct only at chance level (17/32, 53%).

[b]Cohen *d* effect size interpretation: 0.2=small effect size, 0.5=medium effect size, and 0.8=large effect size.

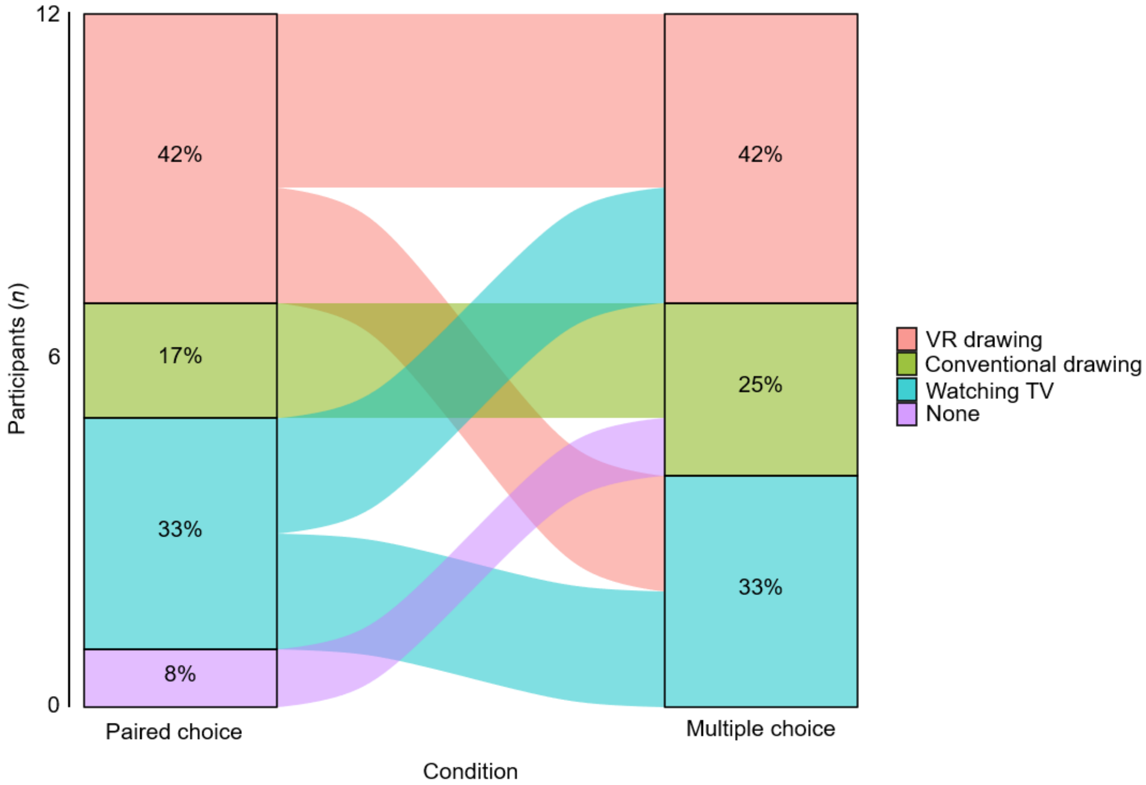
[c]Significant values (*P*<.05) are italicized.

[d]Total score indicates the sum of all subscales.

Self-reported Activity Preference

Twelve participants responded to the choice paradigm. The remaining 4 participants were unable to complete the choice paradigm as they were nonresponsive to the questions. Regarding single stimulus responses, all participants provided a yes response for both VR and conventional drawing, indicating their enjoyment of each activity. Figure 1 shows the percentage of participants' responses in paired and multiple-choice responses. Five out of 12 participants (42%) who completed the choice paradigm were inconsistent upon a comparison of their paired choice and multiple-choice responses. The alluvial plot highlights the inconsistency in the participant's self-report when asking questions in different formats.

Figure 1. Alluvial plot showing self-reported activity preference for virtual reality drawing, conventional drawing, and watching TV under different formats (paired choice versus multiple choice). The alluvial plot highlights the percentage of participants who had an inconsistent response when asked about the same topic (activity preference) in a different way. VR: virtual reality.



Discussion

Overall, learner behavior was found to improve after both VR drawing and conventional drawing in people with Down syndrome and a severe-to-profound intellectual disability. As assessed by blinded staff members, there was no evidence that one intervention was more effective than the other. Perhaps participants found the break from the learning setting most valuable, irrespective of the activity. Although the total score of learner behavior significantly improved, not all factors improved. After both activities, considerable improvements were found for mood and attention, while the differences for impulses and anxiety were nonremarkable. Impulsiveness is rigid and perhaps less likely to change from a brief intervention. Attention deficits in people with Down syndrome are well documented, though effective ways to improve them are less understood. The findings from this study are encouraging, as 10 minutes or less of conventional and VR drawing improved learner behavior, which may subsequently improve learning success.

It stands to reason that there is scope to amplify the observed positive effects. First, drawing is likely not the preferred activity for all participants. An advantage of VR is to easily tailor an experience or activity to the specific preferences of a user. Perhaps being more selective with the activities could have improved engagement and the subsequent effects on behavior. Second, some participants may have preferred clearly structured activities as opposed to free play as this is more common at the organization. Third, the researcher removed participants from the learning setting at any time to start the activity. It may have been more effective to conduct the intervention during instances when participants were exhibiting poor or undesirable learning behavior. Consideration of these factors is an important direction for further study.

The key aim of this study was to assess the feasibility aspects of using VR in people with Down syndrome. Of the 22 people invited to participate, 17 agreed, indicating the willingness of the participants and their guardians to participate in research using VR technology. Sixteen out of 17 participants were able to engage in the VR activity successfully. The findings from this study demonstrate the feasibility of VR use in young adults with Down syndrome and severe-to-profound intellectual disability. Overall, VR was tolerated well. Of the 16 participants, 2 (13%) elected to end the VR experience before the allocated time expired, and 5 (31%) reported visual discomfort symptoms after VR, including eye strain and dizziness.

Given that participants' verbal abilities were limited, they were not able to clearly quantify the severity of their symptoms. Behavioral observations were therefore essential to detect if participants appeared distressed after exposure to the intervention. During these observations, no staff member reported unusual behavior or other significant concerns. Furthermore, behavior in the learning setting improved. We take this as good evidence that there were no serious negative effects of VR exposure. If many participants were sick, we would expect to see a negative impact on behavior. Yet, we found the opposite. The improved ratings and lack of distress identified in behavioral observations indicate acceptable levels of cybersickness in this study, as all symptoms were mild and short-lived. It is important to note that participants may have presented with greater positive behavior despite negative aftereffects owing to a novelty effect, as this was perhaps their first time using VR technology. It remains an open question to what degree a novelty effect contributed to the results.

Although there were no obvious major concerns of cybersickness, it is highly recommended that the interpupillary distance (IPD) be measured and adjusted in headsets to reduce the likelihood of visual discomfort. IPD is the distance between the pupils of both eyes, which facilitates the correct positioning of VR headset lenses. IPD range is essential for optimal image quality, comfort, and has been related to cybersickness in HMDs [25]. In VR, specific points on the lenses must coincide with the center of each eye's pupil (visual axis) for the display image to be in focus [25]. In our study, participants were unable to adjust the IPD in VR. If a VR headset does not allow such eye lens alignment, eyestrain and headaches can be expected [25]. It is plausible that the user's inability to adjust IPD in this study contributed to the experiences of cybersickness.

There was a large variability found between participants' abilities in understanding instructions when using the device. In this study, participants engaged in a simple activity by using basic functions on a handheld controller. The authors note that basic functionality in the application was restricted as 2 controllers are typically used to select in-game options. It is unclear if the success found in this study would translate to VR tasks that require more complex interactions, as it was found that many participants needed assistance before engaging in the task independently.

The choice paradigm was designed to assess self-reported activity preference; yet, this measure's findings highlight a more significant issue. It was found that 5 out of 12 participants (42%) provided inconsistent responses when selecting their preference. For example, a participant may have selected activity A as preferred over activities B and C during the paired-choice options, but when asked to select their favorite overall, they selected B. This measure highlights the difficulty of obtaining valid self-report measures in people with Down syndrome [19]. Objective measures are critical for assessing the safety of VR for this population. Standardized measures such as the simulator sickness questionnaire [26] are typically used to capture cybersickness. Stereoacuity measures would also help measure stereovision (the ability to perceive depth), which is another essential component for an enjoyable VR experience. However, there are no suitable and standardized measures for this group. Despite the challenges with self-report in this population, it is important that measures of cybersickness and stereoacuity are not neglected. Valid assessments are needed that accommodate the language barriers in this sample.

This study's results are encouraging, as VR usage was found to be highly feasible for people with Down syndrome. Participants enjoyed the VR experience and engaged in the task well. There were some experiences of cybersickness; however, all were mild and short-lived. Large positive effects were found for a brief drawing experience to improve overall learner behavior following both VR and conventional interventions. This suggests that immersive VR exposure may provide similar benefits to traditional (paper and pencil) options. For researchers and educators interested in using VR in similar samples, it is recommended that measures are carefully operationalized and there is limited reliance on self-report. Assessing cybersickness is essential to ensuring that users engage in a positive VR experience. An observation checklist worked well in this study to determine the frequency of positive and negative behaviors after exposure to the VR intervention. In this experiment, VR was used as a tool for leisure. Based on the success identified in this study, researchers could investigate the potential for people with Down syndrome to complete tasks and develop real-world skills via training in VR. VR provides promise as a tool to practice real-world skills in a safe, repeatable, and controlled environment [27], though further research is required.

Acknowledgments

SCM was supported by the Australian Government Research Training Program Scholarship. TL was funded by National Health and Medical Research Council Dementia Research Leadership Fellowship (GNT1136269).

Authors' Contributions

The study was designed by SCM and TL with input from all others. Data recruitment and collection was done by SCM. Data analysis was done by SCM and TL. This paper was written by SCM with initial review by TL. Further critical revision and final approval was performed by AS, CE, and RC.

Conflicts of Interest

None declared.

References

1. Wishart J. Motivation and learning styles in young children with Down syndrome. Downs Syndr Res Pract 2001 Oct;7(2):47-51 [FREE Full text] [doi: 10.3104/reports.113] [Medline: 11721529]
2. Shalev N, Steele A, Nobre AC, Karmiloff-Smith A, Cornish K, Scerif G. Dynamic sustained attention markers differentiate atypical development: The case of Williams syndrome and Down's syndrome. Neuropsychologia 2019 Sep;132:107-148. [doi: 10.1016/j.neuropsychologia.2019.107148]

3. Alex M, Wünsche BC, Lottridge D. Virtual reality art-making for stroke rehabilitation: Field study and technology probe. International Journal of Human-Computer Studies 2021 Jan;145:102481. [doi: 10.1016/j.ijhcs.2020.102481]
4. de Oliveira Malaquias FF, Malaquias RF. The role of virtual reality in the learning process of individuals with intellectual disabilities. Technology and Disability 2017 Jan 03;28(4):133-138. [doi: 10.3233/tad-160454]
5. Cobb SVG. Virtual environments supporting learning and communication in special needs education. Topics in Language Disorders 2007;27(3):211-225. [doi: 10.1097/01.TLD.0000285356.95426.3b]
6. Anderson AP, Mayer MD, Fellows AM, Cowan DR, Hegel MT, Buckey JC. Relaxation with immersive natural scenes presented using virtual reality. Aerospace Medicine and Human Performance 2017 Jun 01;88(6):520-526. [doi: 10.3357/amhp.4747.2017]
7. Ramdoss S, Lang R, Fragale C, Britt C, O'Reilly M, Sigafoos J, et al. Use of computer-based interventions to promote daily living skills in individuals with intellectual disabilities: a systematic review. J Dev Phys Disabil 2011 Oct 29;24(2):197-215. [doi: 10.1007/s10882-011-9259-8]
8. Standen PJ, Brown DJ. Virtual reality in the rehabilitation of people with intellectual disabilities: review. Cyberpsychol Behav 2005 Jun;8(3):272-282. [doi: 10.1089/cpb.2005.8.272] [Medline: 15971976]
9. Pizzoli SFM, Mazzocco K, Triberti S, Monzani D, Alcañiz Raya ML, Pravettoni G. User-centered virtual reality for promoting relaxation: an innovative approach. Front Psychol 2019;10:479 [FREE Full text] [doi: 10.3389/fpsyg.2019.00479] [Medline: 30914996]
10. Weiss PLT, Bialik P, Kizony R. Virtual reality provides leisure time opportunities for young adults with physical and intellectual disabilities. Cyberpsychol Behav 2003 Jun;6(3):335-342. [doi: 10.1089/109493103322011650] [Medline: 12855092]
11. Yalon-Chamovitz S, Weiss P. Virtual reality as a leisure activity for young adults with physical and intellectual disabilities. Res Dev Disabil 2008;29(3):273-287. [doi: 10.1016/j.ridd.2007.05.004] [Medline: 17590313]
12. Wilson PN, Foreman N, Stanton D. Virtual reality, disability and rehabilitation. Disabil Rehabil 1997 Jun;19(6):213-220. [doi: 10.3109/09638289709166530] [Medline: 9195138]
13. Krinsky-McHale SJ, Silverman W, Gordon J, Devenny DA, Oley N, Abramov I. Vision deficits in adults with Down syndrome. J Appl Res Intellect Disabil 2014 May;27(3):247-263 [FREE Full text] [doi: 10.1111/jar.12062] [Medline: 23784802]
14. Turnbull PRK, Phillips JR. Ocular effects of virtual reality headset wear in young adults. Sci Rep 2017 Nov 23;7(1):16172 [FREE Full text] [doi: 10.1038/s41598-017-16320-6] [Medline: 29170432]
15. Boon MY, Asper LJ, Chik P, Alagiah P, Ryan M. Treatment and compliance with virtual reality and anaglyph-based training programs for convergence insufficiency. Clin Exp Optom 2020 Nov;103(6):870-876. [doi: 10.1111/cxo.13057] [Medline: 32201990]
16. Saredakis D, Szpak A, Birckhead B, Keage HAD, Rizzo A, Loetscher T. Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Front Hum Neurosci 2020;14:96 [FREE Full text] [doi: 10.3389/fnhum.2020.00096] [Medline: 32300295]
17. Szpak A, Michalski SC, Saredakis D, Chen CS, Loetscher T. Beyond feeling sick: the visual and cognitive aftereffects of virtual reality. IEEE Access 2019;7:130883-130892. [doi: 10.1109/access.2019.2940073]
18. Spanò M, Mercuri E, Randò T, Pantò T, Gagliano A, Henderson S, et al. Motor and perceptual-motor competence in children with Down syndrome: variation in performance with age. Eur J Paediatr Neurol 1999;3(1):7-13. [doi: 10.1053/ejpn.1999.0173] [Medline: 10727186]
19. Schwartz C, Rabinovitz S. Life satisfaction of people with intellectual disability living in community residences: perceptions of the residents, their parents and staff members. J Intellect Disabil Res 2003 Feb;47(Pt 2):75-84. [doi: 10.1046/j.1365-2788.2003.00436.x] [Medline: 12542572]
20. Emerson E, Felce D, Stancliffe RJ. Issues concerning self-report data and population-based data sets involving people with intellectual disabilities. Intellect Dev Disabil 2013 Oct;51(5):333-348. [doi: 10.1352/1934-9556-51.5.333] [Medline: 24303821]
21. Graves RJ, Graff JC, Esbensen AJ, Hathaway DK, Wan JY, Wicks MN. Measuring health-related quality of life of adults with Down syndrome. Am J Intellect Dev Disabil 2016 Jul;121(4):312-326. [doi: 10.1352/1944-7558-121.4.312] [Medline: 27351699]
22. Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders; Board on the Health of Select Populations; Board on Children, Youth, and Families; Institute of Medicine; Division of Behavioral and Social Sciences and Education; The National Academies of Sciences, Engineering, and Medicine. Clinical characteristics of intellectual disabilities. In: Boat TF, Wu JT, editors. Mental Disorders and Disabilities Among Low-Income Children. Washington (DC): National Academies Press; Oct 28, 2015.
23. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146]
24. Mather N, Jaffe LE. Woodcock-Johnson III: Reports, Recommendations, and Strategies. Hoboken, NJ: John Wiley & Sons; 2002.

25. Stanney K, Fidopiastis C, Foster L. Virtual reality is sexist: but it does not have to be. Front Robot AI 2020;7:4 [FREE Full text] [doi: 10.3389/frobt.2020.00004] [Medline: 33501173]
26. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
27. Michalski SC, Ellison C, Szpak A, Loetscher T. Vocational training in virtual environments for people with neurodevelopmental disorders: a systematic review. Front Psychol 2021;12:627301 [FREE Full text] [doi: 10.3389/fpsyg.2021.627301] [Medline: 34305704]

Abbreviations

HMD: head-mounted display
IPD: interpupillary distance
VR: virtual reality

Edited by N Zary; submitted 20.10.21; peer-reviewed by M Mosher, F Redgrove; comments to author 10.11.21; revised version received 28.11.21; accepted 03.02.22; published 07.04.22.

Please cite as:
Michalski SC, Szpak A, Ellison C, Cornish R, Loetscher T
Using Virtual Reality to Improve Classroom Behavior in People With Down Syndrome: Within-Subjects Experimental Design
JMIR Serious Games 2022;10(2):e34373
URL: https://games.jmir.org/2022/2/e34373
doi:10.2196/34373
PMID:35124567

Original Paper

Adapting the Use of Digital Content to Improve the Learning of Numeracy Among Children With Autism Spectrum Disorder in Rwanda: Thematic Content Analysis Study

Theoneste Ntalindwa[1], MSc; Mathias Nduwingoma[1], PhD; Alphonse Uworwabayeho[1], PhD; Pascasie Nyirahabimana[1], MSc; Evariste Karangwa[2], PhD; Tanjir Rashid Soron[3], MD, MSc; Thomas Westin[4], PhD; Thashmee Karunaratne[4], PhD; Henrik Hansson[4], PhD

[1]School of Education, University of Rwanda, Kigali, Rwanda
[2]School of Inclusive and Special Needs Education, University of Rwanda, Kigali, Rwanda
[3]Telepsychiatry Research and Innovation Network Ltd, Dhaka, Bangladesh
[4]Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden

Corresponding Author:
Tanjir Rashid Soron, MD, MSc
Telepsychiatry Research and Innovation Network Ltd
114, Kazi Nazrul Islam Avenue
Dhaka, 1205
Bangladesh
Phone: 880 1718827138
Email: tanjirsoron@gmail.com

Abstract

Background: Many teachers consider it challenging to teach children with autism spectrum disorder (ASD) in an inclusive classroom due to their unique needs and challenges. The integration of information communication technology (ICT) in the education system allows children with ASD to improve their learning. However, these ICT tools should meet their needs to lead a productive life.

Objective: This study aimed to examine the possibilities of re-creating and adapting digital content to improve the learning of numeracy among children with ASD in inclusive school settings.

Methods: We conducted 7 focus group discussions (FGDs) with 56 teachers from 7 schools and 14 parents from April to November 2019. Each of the FGDs took around 1 hour. Two clustered sets of questions were used: (1) general knowledge about teaching children with ASD and (2) analysis of selected online educational video content of early math (specifically, counting numbers). The researchers used video to understand current methodologies used in teaching children with ASD, possibilities of adaptation of the content in the current teaching environment, future challenges when the content is adapted, and possible solutions to overcome those challenges. All data, including audio recordings, field notes, and participants' comments, were transcribed, recorded, and analyzed following the steps recommended in qualitative data analysis.

Results: The researchers identified ten themes from the analysis of the data: (1) awareness of the existence of ASD among children in schools and the community, (2) acceptance of children with ASD in an inclusive classroom and the community, (3) methods and models used when teaching children with ASD, (4)realia used to improve the learning of children with ASD, (5) the design of educational digital content, (6) the accessibility of online educational content, (7) quality of the content of the educational multimedia, (8) the opportunity of using the translated and re-created content inside and outside the classroom, (9) the relevance of the digital content in the Rwandan educational system, and (10) enhancement of the accessibility and quality of the digital content. We found that participants assumed that the content translation, gamification, and re-creation would help teach children with ASD. Moreover, they recommended contextualizing the content, increasing access to digital devices, and further research in the education of different subjects.

Conclusions: Although many studies have identified the possibilities of using ICT to support children with ASD, few studies have documented the possibilities of integrating the existing technologies tested in the international community. This study is charting new territory to investigate online content to suit the context of schools. This study recommends further exploration of possible methodologies, such as applied behavior analysis or verbal behavior therapy, and the development of contextualized technologies that respond to the educational needs of children with ASD.

(JMIR Serious Games 2022;10(2):e28276) doi:10.2196/28276

KEYWORDS

autism; learning; ICT; e-learning; education; children; ASD; teaching; teachers; communication; communication technology; online content; Rwanda; gamification; school; school-age children; behavior

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that appears during the developmental period when the accommodation process starts [1]. The prevalence of autism has increased in recent years, and many people with this condition live in low- and middle-income countries, while only about 1% of research studies have represented these countries in the past decade [2]. The situation is especially disappointing for children with autism, who have specific needs and different behavioral patterns in sub-Saharan Africa than in developed countries.

Children with disabilities face challenges when learning in an inclusive environment [3]. These challenges become critical for children with ASD [4]. More than half of children with learning disabilities enrolled in primary education in Rwanda drop out of school at an early age [5].

According to the National Commission for Persons with Disabilities (NCPD) and National Commission for Children (NCC) [6], few centers care for children with disabilities, because the cost of education for children with ASD is still high, making it difficult for them to continue with their studies. Some schools and centers are still facing the problem of accessing information communication technology (ICT) tools to improve the learning of children with ASD. However, a good number of primary education schools are equipped with computers obtained through initiatives such as One Laptop per Child [7] and the Computer on Loan scheme [8]. These initiatives created the opportunity to use technology to enhance students' learning regardless of their level of learning ability. Several studies [9-12] have shown that ICT integration in the classroom supports children with ASD and other developmental disabilities.

The Rwanda Ministry of Education [13] has developed a competence-based curriculum framework with 7 key competencies: literacy, numeracy, ICT, citizenship and national identity, entrepreneurship and business development, science and technology, and communication in the official languages. Literacy and numeracy are basics to accessing learning in other subjects. The Rwanda competence-based curriculum states that all children must be equipped with skills in computing at the end of the basic education program. They should also be able to use numerical patterns and relations to solve problems related to everyday activities, for example, in commercial contexts and financial management, and interpret basic statistical data using tables, diagrams, charts, and graphs [14].

Children with ASD may enjoy basic arithmetic if the content is gamified using multimedia [10]. According to Spek et al [11], gamification of content for the learning of children with ASD brings the opportunity to overcome limitations on accessibility by fostering independence and assisting the children in social relationships. Moreover, ICT has made the internet one of the best resources for discovering entertaining activities that teach and excite children with different learning disabilities.

Educational websites such as IXL Worldwide [12] and Khan Academy [15] provide content to support the learning of children with learning disabilities, including dyslexia, dysgraphia, attention deficit and hyperactivity disorder, and visual-motor deficit. The introduction of gamification in Khan Academy videos has proved promising to bring new ways of engaging students with activities and providing valuable data for teachers [15,16]. Moreover, several studies [17-19] have shown that online content for teaching mathematics positively impacted a variety of learning environments. However, there is no research on using digital tools for intervention in children with ASD to improve their learning of numeracy skills.

Studies by Abubakar [20] and Onaolapo [21] indicate that few studies have focused on technology and education in ASD in sub-Saharan Africa. Thus, existing online content has yet to be investigated for its ability to support the learning of children with ASD in Rwanda. This study aimed to analyze how to adapt the available digital content to enhance numeracy learning for children with ASD in inclusive school settings.

Methods

Design and Setting

This thematic content analysis study was conducted from April to November 2019 to examine how to adapt digital content to improve learning for children with ASD in Rwanda by developing a tailor-made educational mobile app [22]. The researchers recruited 70 participants, including 56 teachers and 14 parents, from a total of 7 schools, including 5 in urban areas and 2 in rural areas. Only 2 of the urban schools were in Kigali, the capital of Rwanda: Autisme Rwanda and Groupe Scolaire (GS) Jabana; the remaining 3 were from different urban areas in upcountry provinces: HVP Rwamagana (Eastern Province) and Ubumwe Community Center (Western Province). College des Amies de la Paix du Christ Roix (APAX), Janja (Northern Province), and APAX Muramba were in the country's rural areas.

Within the group of teachers, there were 32 males and 24 females, and in the group of parents, there were 6 males and 8 females. The teachers taught subjects including mathematics, elementary science and technologies, social sciences, and languages.

The researchers selected the schools based on data from the National Commission of Persons with Disabilities [6] and our previous study [9]. The parents were from the Rwanda Parent's Initiative on Autism [23]. The parents were not required to have a child enrolled in any school, but might have at least one child with ASD. The inclusion criterion for the schools was that the school had a program to care for children with any disability.

Data Collection

The researchers used focus group discussions (FGDs) to collect the data for this study. FGDs were conducted in a separate room in each school to provide an environment where parents could give accurate, complete, and sincere answers during the discussion. The questions in the interview guide fell into two categories: (1) general knowledge about teaching children with ASD and (2) online video content for early math (specifically, counting numbers) (Multimedia Appendix 1). The researchers selected online content from Khan Academy, which focused on early counting mathematics, based on its audiovisual features, as these can aid learning for children with ASD. The researchers projected the selected educational video on a screen, and the participants discussed the design layout and possible modifications to adapt it to the education of children with ASD in the Rwandan context. The parents participated only in discussing the online video content for early math. The researchers chose to mix the questions according to the mission of the visited school or center. The interviews lasted for 1 hour; records and related soft copies were kept confidential and were saved on a physical drive. In contrast, hard copies of the signed consent forms were kept safe in the principal investigator's office.

Data Analysis

In the analysis of the data, the researchers followed the six steps of thematic content analysis recommended by Jugder [24] and Caulfield [25]: (1) familiarization, (2) coding, (3) generating themes, (4) reviewing themes, (5) defining and naming themes, and (6) writing up. The researchers used conceptual reliability to ensure the validity and reliability of the data at every step of the analysis of the coded data by involving 18 teachers and 6 parents in coding and reviewing data [26].

We followed inductive, descriptive thematic analysis of the FGD transcripts, following several steps. Firstly, transcripts were read and reread to establish familiarity with the data. Then, initial codes that captured features of interest and importance to the research questions were identified. Next, a coding framework was developed by collating initial codes to create candidate codes that meaningfully described the overall patterns of participant responses in the data. Next, researchers and representatives of the participants reviewed the transcripts to ensure that the coding framework captured participant responses. The team then coded the transcripts line by line to collate all instances of patterns identified in the data. Finally, the team collapsed the candidate codes to produce higher-order themes that were then read for patterns of similarity and divergence within and across each theme.

In the coding and review of themes, the researchers, teachers, and parents answered the following questions: (1) Did we skip any relevant information? (2) Do these themes represent the data? and (3) Could any modifications make the theme better? In this process, we split some themes and created new ones to make them more valuable and accurate. Only researchers wrote up the final manuscript, which was the last step of data analysis. Data management was supported using the qualitative software program NVivo (version 9; QSR International) [27]. Participant demographic data were analyzed using SPSS-23 (IBM) [28].

Ethics Approval

Trained researchers from the University of Rwanda collected the data from the participants. The research project passed through the college ethical committee by following the established process [29]: (1) presentation of the research proposal, (2) submission of the application and tools to be used for ethical research clearance, and (3) review and approval of the application by the ethical research committee (review number 01/P-CE/635/EN/gi/2019) (Multimedia Appendix 2). The researchers also obtained formal informed consent from the parents and teachers using the approval form (Multimedia Appendix 3).

Results

Participant Characteristics and Themes Identified by the Analysis

Table 1 and Table 2 show the demographic characteristics of the parents and teachers, respectively, who participated in the study.

From the thematic analysis of the interview transcripts captured during 7 FGDs, the researchers identified the following themes: (1) awareness of the existence of ASD among children in schools and community, (2) acceptance of children with ASD in an inclusive classroom and in the community, (3) methods and models used when teaching children with ASD, (4) realia used to improve the learning of children with ASD, (5) the design of educational digital content, (6) accessibility of online educational content, (7) quality of the content of the educational multimedia, (8) opportunity of using the translated and re-created content inside and outside the classroom, (9) the relevance of the digital content in the Rwandan education system, and (10) enhancement of the accessibility and quality of the content.

Table 1. Sociodemographic characteristics of the parents by location.

Characteristics	Kigali City	Eastern Province	Southern Province	Northern Province	Western Province
Gender, n					
Male	2	1	1	1	1
Female	4	1	1	1	1
Total	6	2	2	2	2
Age, n					
Less than 20 years	0	0	0	0	0
Between 20 and 25 years	0	0	0	0	0
Between 25 and 30 years	0	0	0	0	0
More than 30 years	6	2	2	2	2
Total	6	2	2	2	2

Table 2. Sociodemographic characteristics of the teachers by school.

Characteristics	HVP–Gatagara Kigali	HVP–Rwamagana	APAX–Janja	APAX–Muramba	Ubumwe Community Center	Autisme Rwanda	GS Jabana
Gender, n							
Male	2	2	4	3	2	5	6
Female	6	6	4	5	6	3	2
Age, n							
Less than 20 years	1	0	1	2	1	0	0
Between 20 and 25 years	5	6	5	4	2	3	3
Between 25 and 30 years	2	1	1	1	1	3	2
More than 30 years	0	1	1	1	4	2	3
Total	8	8	8	8	8	8	8
Years of experience, n							
Less than 5 years	4	4	5	7	8	8	5
Between 5 and 10 years	2	3	3	1	0	0	3
Between 10 and 15 years	1	1	0	0	0	0	0
More than 15 years	1	0	0	0	0	0	0
Total	8	8	8	8	8	8	8
Subjects taught, n							
Mathematics	5	3	4	6	3	4	3
Science related	1	2	1	1	2	2	2
Social sciences	0	2	1	0	0	1	1
Languages	2	1	2	1	3	1	2

Awareness of the Existence of ASD Among Children in Schools and the Community

When asked about the presence of ASD among children registered in their classes, 24 of the 56 teachers (43%) reported that they were not aware of ASD among their students because they manifest different behaviors. ASD is a spectrum, and can range from low to high functioning ASD; not all children with ASD are the same. Unfortunately, this makes educators confuse children with ASD with children with other cognitive disabilities.

ASD is not yet well known in the Rwandan community, which results in various misunderstandings of the behavior of children with ASD in families. Eight teachers suggested that diagnosing children before they begin school could facilitate them in identifying teaching methods and content for children who may have ASD. One parent reported never knowing what ASD was before confirming his child's diagnosis at a hospital. The increase of qualified medical personnel in all hospitals with professional diagnostic tools and collaboration with educators is a solution to raise awareness of ASD in schools to prepare personalized learning materials for children on the autism spectrum.

Acceptance of Children With ASD in an Inclusive Classroom and Community

In the sample of teachers who participated in the FGDs, 32 of 56 (57%) teachers reported difficulties teaching children with ASD in the same classroom as children without cognitive disabilities. The behavior of children with ASD is a factor that creates difficulties in accepting children with ASD in all schools. However, some participants reported including children with ASD in schools when educators received training in special education for children with ASD and personalized teaching material. The teachers also said that trained educators and teaching aids were more effective if they integrated assistive technologies into their teaching and learning activities. The integration of assistive technologies such as with ICT when teaching children with ASD could bring greater success to the inclusion of children with ASD in the Rwandan education system.

Methods Used When Teaching Children With ASD

This theme elucidated models used to support children with ASD in the educational environment. Teachers understood how to apply the applied behavior analysis model in a classroom environment and were trying to implement strategies to motivate children with ASD to stay focused while learning. Eight of 56 teachers (14%) reported that children with ASD needed quick motivations, such as giving them specific objects that they liked that were available in the school environment. Teachers could also provide personalized content to aid learning for children with ASD. For example, when asked what methods helped pupils get to grips with mathematics, teachers reported that games were the best approach to stay focused and engaged in the classroom. The introduction of games brought freedom to the children to work on the given task and join different groups of children. This approach enabled them to socialize with others and enhanced teamwork in inclusive class settings. Most teachers reported that providing personalized content and creating an environment conducive to learner interaction could help reveal the learning abilities of children with ASD and develop their inner talents.

The Use of Realia to Improve the Learning of Children With ASD

This theme involved using various objects to provide contextualization in the education of children with ASD. Providing a connection to the school's physical environment allowed students to better understand the content of the subject they were learning. When asked about the objects used in teaching children with ASD, 40 of 56 teachers (71%) responded that they used real examples of objects available in schools to replace the models in the curriculum syllabus. The strategy of contextualization of content helps children to understand the topic planned in the curriculum.

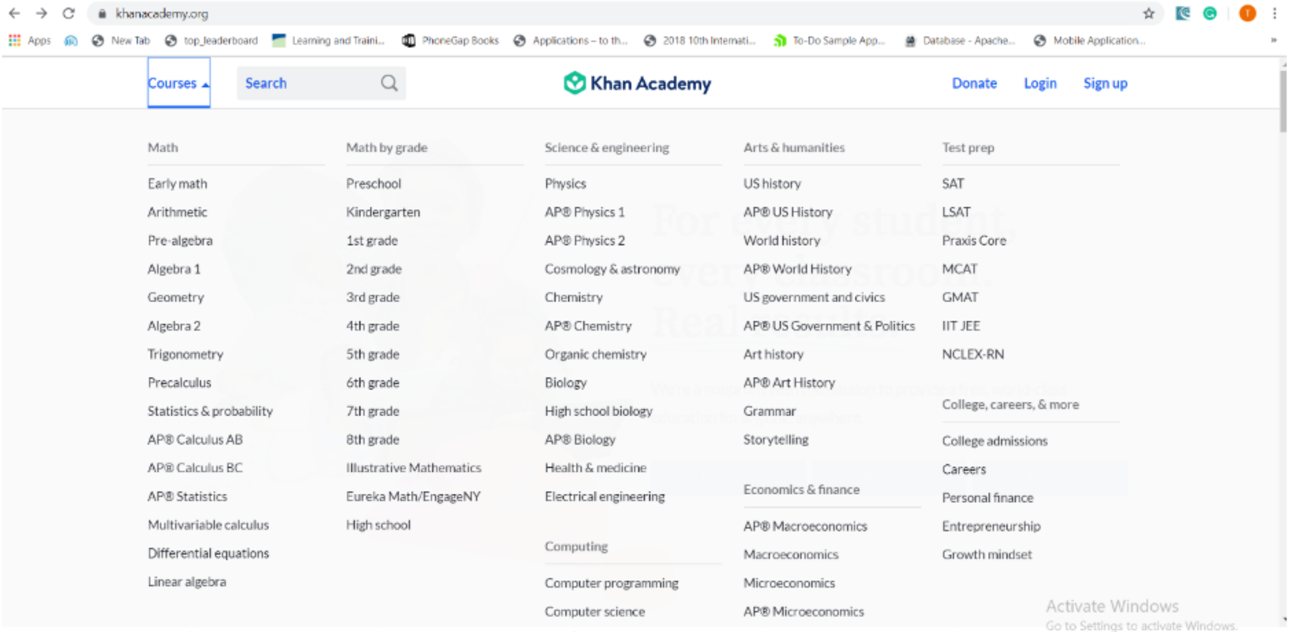
All teachers (100%) reported that educational videos were an alternative to support unconscious learning among children with ASD when they were out of class. Combining traditional teaching methods with ICT-enabled teaching methods in inclusive learning among children with and without ASD is the best approach to provide an equitable education to all children.

The Design of Educational Digital Content

This theme explored the design of the content of the selected website. Teachers reported that it was their first time browsing the Khan Academy website, even those familiar with other websites for finding educational resources. Despite the excellent layout of the website, participants suggested explanations of some abbreviations, such as "SAT," "LSAT," and "MCAT," among others. Figure 1 shows further examples of these abbreviations. All teachers (100%) appreciated the content categorization on the course page. Categorizing the cognitive level of children supports education in Rwanda from preschool to higher education. This might bring an early, positive impact when teaching children with ASD.

All 14 parents (100%) reported that their children could not follow the formal education system and recommended using the examples and images available in their families when designing a web interface for children with ASD. The contextualization of content by creating an intellectual need for information and skills could help children with ASD learn through practice to prevent them from losing attention. In addition, developing the visual and hearing senses of children with ASD could help the designers of digital content to better enable them to learn.

Figure 1. Khan Academy platform [15].



Accessibility of Educational Digital Content

Access to digital content is helpful to support the education of learners with ASD. Educational videos hosted on YouTube channels are made accessible by internet connectivity. Users can download video content from YouTube and use it offline. This option is better for schools that have low internet connectivity. Among the teachers, 16 of 56 (29%) from schools in rural areas reported barriers to accessing online content due to poor internet connectivity and recommended increasing the bandwidth of their internet connection, despite the possibility of downloading the videos and playing them offline. Teachers reported that the reliability of the web interface was a driving factor for adaptation of online content for the education system in Rwanda. The 56 teachers and 14 parents also recommended developing digital content that could be accessed offline as a mobile application.

Content Quality of Educational Multimedia

YouTube is a video-sharing platform used by many online education websites to provide educational content to various users. Professionals from other domains create the videos; some of them are published under a Creative Commons license to allow users to adapt or re-create them. In this study, participants suggested content re-creation to meet the Rwandan context. All 56 teachers (100%) reported that educational digital content could support children with ASD when designers adapt the content to the local context.

Teachers and parents argued that the translation of the content into the local language could bring a positive impact not only to the learning of children with ASD, particularly literacy and numeracy, but also to nondisabled children who are interested in using educational digital content and children who are not registered in schools. This strategy is crucial, because 13 of 14 parents (93%) did not know English, and 42 of 56 teachers (75%) had difficulty using English. The translation and re-creation of content to be adapted for the local context might be an excellent approach to integrate ICT into the education system and improve the learning of children with various cognitive disabilities in Rwanda.

Opportunities to Use Translated and Re-created Content Inside and Outside the Classroom

This theme elucidated the opportunity to use the re-created and translated content to upgrade the current education system to international standards. All 56 teachers and 14 parents reported that Khan Academy's content could help teach basic mathematics to children with ASD because of the availability of interactive videos. In the parent group, 13 of 14 (93%) reported that the content was useful in helping their children be engaged when they were at home. Teaching children with ASD in the local language has greater advantages than teaching in a foreign language, as this creates a deficit in communication. In addition, the translation and adaptation of international content can bring new opportunities to all children to learn the same content prepared by globally recognized experts in different fields.

The Relevance of the Digital Content in the Rwandan Education System

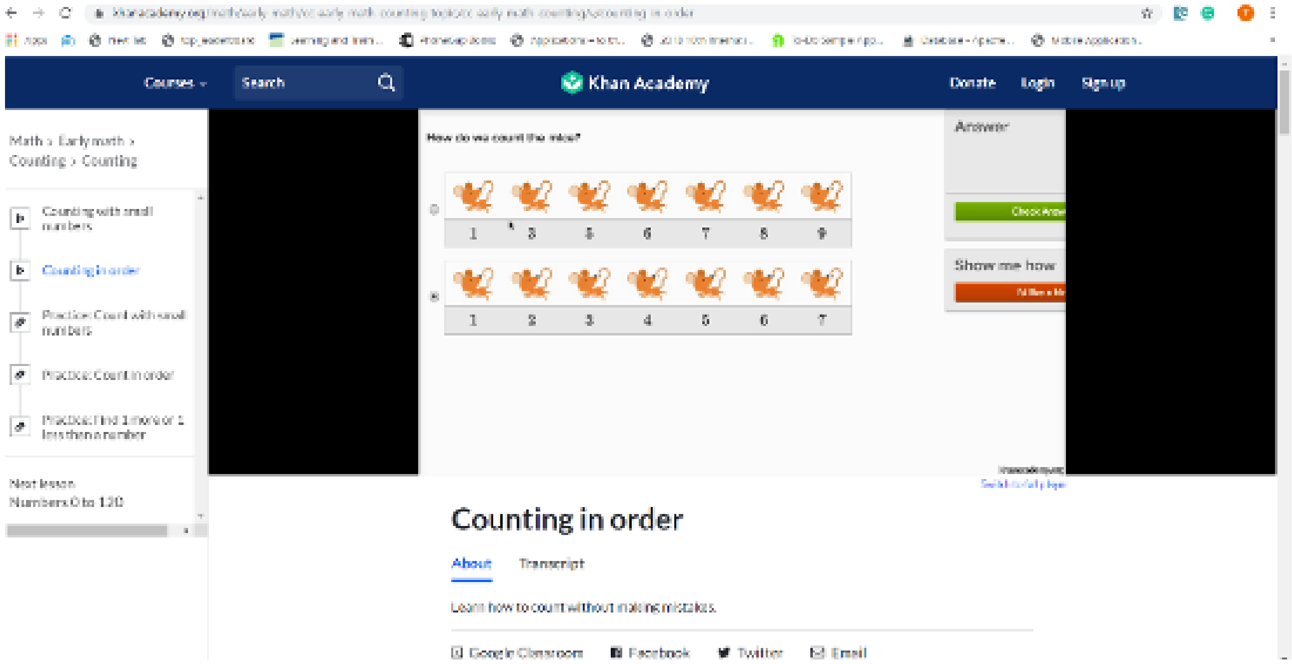
Of the 70 total participants in this study, all the teachers (56/70, 80%) and parents (14/70, 20%) welcomed the implementation of technology-enhanced content in teaching. The teachers reported that digital educational material was essential in preventing cognitive overload and providing content. However, ICT adoption in education needs strategies and financial investment to obtain positive results. Among the teachers, 16 of 56 (29%) suggested that these strategies could include training educators in digital content in teaching and learning, as well as improving infrastructure. The availability of offline content is considered a solution to bridge the gap in internet access between urban and rural areas in Rwanda. However, this is not a perfect solution, as some schools still report a lack of digital devices to play the content offline. Adopting mixed digital and traditional learning methods could enable children to learn despite a lack of digital devices.

Enhancement of the Accessibility and Quality of the Content

In the FGDs, 65 of 70 (93%) participants suggested further online teaching for children with ASD. Teachers criticized the researchers' content during a contextualization exercise for teaching early mathematics [30] in data collection. The narrator's voice in the videos was also not well understood by the teachers. The participants suggested replacing the narrator with a speaker of a local language. The teachers and parents suggested the re-creation of the content to add examples of situations that the children encounter in their everyday lives. Figure 2 shows the interface for early math (counting).

Teachers suggested that re-creating the content of this lesson to better suit the Rwandan teaching environment would provide personalized material to improve the learning of children with ASD in an inclusive environment. Participants also recommended a subscription model for the re-created content to improve its accessibility.

Figure 2. Interface for early math (counting) [17].



Discussion

Principal Results

The findings of this study indicate the possibility of using digital content in the education of children with ASD. The themes reported in the results section show that integrating multimedia content can increase access to education for children with ASD [9]. This is possible if educators are aware of children with ASD in their schools and adopt methodologies that help these learners stay focused [31]. This could increase the acceptance of these children in mainstream schools [18].

The development of various models to support children with ASD [32], as well as computer-assisted educational content [19], is progressing rapidly in various domains. Nevertheless, such research needs to computerize existing methodologies and content in an individualized context. A study by O'Malley et al [33] reported that educational software for children with ASD must respond to specific interests and developmental needs. In addition, the interface for such software must be as interactive as possible to most effectively facilitate learning in children with ASD [34].

Teaching basic numeracy to children who have a deficit in communication, such as those with ASD, is more possible when the content is gamified using multimedia technologies [16]. All participants in this study used games when teaching various subjects in class. Educational videos are essential to improve the learning of children with ASD [35,36] (Multimedia Appendix 4). After watching the Khan Academy content used in this study [17], the teachers reported that integrating content after it was re-created could allow learners with ASD to participate in an inclusive classroom [30].

This study evaluated Khan Academy content [17] on the YouTube platform [37], which has gamified content delivered in educational institutions. A study by Baker et al [38] found that integrating multimedia content improved learning for various subjects [39]. Participants in the present study suggested re-creating the Khan Academy content to meet the school environmental context.

This study was carried out before the COVID-19 pandemic became a crisis in Rwanda [40]. However, our results demonstrate that online content could help educate children with ASD during and after the COVID-19 crisis. A study by Stenhoff et al [41] documented the possibility of supporting children with ASD through distance education during school closures. The availability of digital content for learners with ASD is also crucial to support remote learning in response to COVID-19 prevention measures [42].

The translation of the findings of this study into practice would be helpful for educators and would enable future research to address barriers to education for children with ASD by focusing on the functional abilities of these children, rather than using a deficit model based on specific diagnoses. Furthermore, enabling children with ASD to learn mathematics would contribute to eliminating all causes and obstacles that can lead to disparity in education, such as gender, disability, and geographical or social group. This is an objective of the Rwanda education sector [14].

Strengths and Limitations

This study had many strengths. It reflects the experiences of teachers who serve to support children with ASD; it was performed by experienced researchers in education, information technology, special education, and ASD; and it included schools from both rural and urban areas. Thus, our findings might allow teachers to adopt new, different methods and innovative tools to improve teaching and learning. Nevertheless, this study was limited by including only a small number of teachers and parents, making our results hard to generalize to all education practitioners. Furthermore, we only focused on the subject of basic numeracy, whereas primary education includes many more subjects.

In future studies, all content relating to the existing syllabus should be explored, and participants should be allowed more time to test the interface of the software and provide more detailed opinions. In addition, a longitudinal usability study of the interface might help uncover long-term advantages and disadvantages that teachers may experience and enable adaptation of the curriculum in a way that the first-time experiences described here could not. Finally, future research should further explore the development of personalized ICT solutions for individuals with ASD that respond to their educational needs.

Conclusions

The study documented the process of contextualization of technology to make it a better solution that meets the actual context of its environment. Integrating systems designed by internationally recognized experts and translating these systems into a local context could bring innovation in teaching children with disabilities. This study charts new territory in the investigation of online content and its ability to match the context of primary and secondary schools. We recommend further exploration of methodologies such as applied behavior analysis and verbal behavior therapy, and we also recommend the development of contextualized technologies that respond to the educational needs of children with ASD.

Acknowledgments

Khan Academy did not provide any financial support to influence the research findings of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide question for primary teachers.

[PDF File (Adobe PDF File), 87 KB - games_v10i2e28276_app1.pdf]

Multimedia Appendix 2

Research Ethical Clearance.

[PDF File (Adobe PDF File), 205 KB - games_v10i2e28276_app2.pdf]

Multimedia Appendix 3

Informed Consent Form for teachers.

[PDF File (Adobe PDF File), 151 KB - games_v10i2e28276_app3.pdf]

Multimedia Appendix 4

Sample Form for Observers' Notes at Focus Groups.

[PDF File (Adobe PDF File), 212 KB - games_v10i2e28276_app4.pdf]

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.). Am J Psychiatry. Washington, DC: American Psychiatric Publishing; 2013.
2. Squire LR, Bloom FE, Ghosh A, Berg D, Du Lac S, Spitzer N, editors. Fundamental Neuroscience. Amsterdam: Elsevier; 2014.
3. Karangwa E. Towards Inclusive Education in Rwanda An assessment of the socio-political contributors to Inclusive Education developments. Rwandan Journal of Education 2014;2(1):46-60 [FREE Full text]

4. Hahler E, Elsabbagh M. Autism: A Global Perspective. Curr Dev Disord Rep 2014 Dec 28;2(1):58-64. [doi: 10.1007/s40474-014-0033-3]
5. Rwanda Statistical Yearbook 2019. Geneva: United Nations; 2019.
6. Report on National Assessment of Centres Caring for Children with Disabilities in Rwanda. 2016. URL: http://ncc.gov.rw/fileadmin/templates/document/National_Assessment_of_Centres__Caring_for_CWDs.pdf [accessed 2016-06-16]
7. UN HABITANT. Smart city Rwanda Master plan. 2017 May 01. URL: https://unhabitat.org/smart-city-rwanda-master-plan [accessed 2021-06-21]
8. Ministry of Youth and ICT (MYICT). ICT Sector Profile - 2015. URL: https://www.minict.gov.rw/ [accessed 2021-07-21]
9. Ntalindwa T, Soron TR, Nduwingoma M, Karangwa E, White R. The Use of Information Communication Technologies Among Children With Autism Spectrum Disorders: Descriptive Qualitative Study. JMIR Pediatr Parent 2019 Sep 27;2(2):e12176 [FREE Full text] [doi: 10.2196/12176] [Medline: 31573940]
10. Ern A. The use of gamification and serious games within interventions for children with autism spectrum disorder: a systematic review. University of Twente. 2014 Jan. URL: http://purl.utwente.nl/essays/64780 [accessed 2022-04-04]
11. Spek EVD, Göbel S, Do E. Entertainment Computing and Serious Games. 2019 Presented at: First IFIP TC 14 Joint International Conference; Nov 11-15; Arequipa, Peru. [doi: 10.1007/978-3-030-34644-7]
12. Soutien scolaire ludique et interactif. 2020. URL: https://fr.ixl.com/ [accessed 2022-04-04]
13. Ministry of Education. Republic of Rwanda Ministry of Education Education Sector Strategic Plan 2018/19 to 2023/24. Kigali, Rwanda: Ministry of Education; 2019:2018-2128.
14. Rwanda EB. Competence based Curriculum. 2015. URL: http://reb.rw/ [accessed 2022-04-04]
15. Khan Academy. 2019. URL: https://www.khanacademy.org/ [accessed 2022-04-04]
16. Tenório MM, Lopes RP, Góis LAD, Junior GDS. Influence Of Gamification On Khan Academy In Brazilian High School. PUPIL: International Journal of Teaching, Education and Learning 2018;2(2):51-65 [FREE Full text] [doi: 10.20319/pijtel.2018.22.5165]
17. Khan Academy. Early Math Counting. 2019. URL: https://www.khanacademy.org/math/early-math/cc-early-math-counting-topic/cc-early-math-counting/v/counting-in-order [accessed 2022-04-04]
18. Nolen A, Putten JV, Hastings R. Autism Spectrum Disorder and College Major Choice: Looking Beyond the STEM Majors. In: International Conference on Information Technology and Electrical Engineering. 2015 Presented at: 2015 Annual Meeting of the American Educational Research Association; April 16-20, 2015; Chicago, IL.
19. Fletcher-Watson S. A Targeted Review of Computer-Assisted Learning for People with Autism Spectrum Disorder: Towards a Consistent Methodology. Rev J Autism Dev Disord 2013 Nov 14;1(2):87-100. [doi: 10.1007/s40489-013-0003-4]
20. Abubakar A, Ssewanyana D, de Vries PJ, Newton CR. Autism spectrum disorders in sub-Saharan Africa. Lancet Psychiatry 2016 Sep;3(9):800-802 [FREE Full text] [doi: 10.1016/S2215-0366(16)30138-9] [Medline: 27568261]
21. Onaolapo AY, Onaolapo OJ. Global Data on Autism Spectrum Disorders Prevalence: A Review of Facts, Fallacies and Limitations. Universal Journal of Clinical Medicine 2017 Dec;5(2):14-23. [doi: 10.13189/ujcm.2017.050202]
22. Ntalindwa T, Nduwingoma M, Karangwa E, Rashid Soron T, Uworwabayeho A, Uwineza A. Development of a Mobile App to Improve Numeracy Skills of Children With Autism Spectrum Disorder: Participatory Design and Usability Study. JMIR Pediatr Parent 2021 Aug 31;4(3):e21471 [FREE Full text] [doi: 10.2196/21471] [Medline: 34463629]
23. Rwanda Parent's Initiative on Autism. 2020. URL: http://www.rpia.rw [accessed 2022-04-04]
24. Jugder N. The thematic analysis of interview data: an approach used to examine the influence of the market on curricular provision in Mongolian higher education institutions. URL: https://pdfs.semanticscholar.org/2e0a/746b717fa49b6f4a48a462ee47e602204de9.pdf [accessed 2022-04-04]
25. Caulfield J. How to do thematic analysis. 2019. URL: https://www.scribbr.com/methodology/thematic-analysis/ [accessed 2022-04-04]
26. Nowell LS, Norris JM, White DE, Moules NJ. Thematic Analysis. International Journal of Qualitative Methods 2017 Oct 02;16(1):160940691773384. [doi: 10.1177/1609406917733847]
27. NVivo. 2020. URL: https://www.qsrinternational.com/ [accessed 2022-04-04]
28. IBM. IBM SPSS Statistics. 2020. URL: https://www.ibm.com/products/spss-statistics [accessed 2022-04-04]
29. University of Rwanda. Ethics Operational Guidelines and Procedures. URL: https://ur.ac.rw/IMG/docx/ethics_guidelines_procedures-3.docx [accessed 2022-04-04]
30. Simsek O, Allsopp D. Use of a Game-Based App as a Learning Tool for Students with Mathematics Learning Disabilities to Increase Fraction Knowledge/Skill. 2016. URL: https://digitalcommons.usf.edu/etd/6390/ [accessed 2022-04-04]
31. Alotaibi F. Saudi Teachers? Perceptions of ICT Implementation for Student with Autism Spectrum Disorder at Mainstream Schools. Journal of Education and Practice 2016;7(5):116-124.
32. Virginia Department of Education O of SE and SS. Models of Best Practice in the Education of Students with Autism Spectrum Disorders. 2011. URL: https://www.doe.virginia.gov/special_ed/disabilities/autism/technical_asst_documents/autism_models_of_best_practice.pdf [accessed 2022-04-04]
33. O'Malley P, Lewis M, Donehower C, Stone D. Effectiveness of Using iPads to Increase Academic Task Completion by Students with Autism. Universal Journal of Educational Research 2014 Jan;2(1):90-97. [doi: 10.13189/ujer.2014.020111]

34. Tashnim A, Nowshin S, Akter F, Das A. Interactive interface design for learning numeracy and calculation for children with autism. In: International Conference on Information Technology and Electrical Engineering.: IEEE; 2017 Presented at: 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE); October 12-13, 2017; Phuket, Thailand. [doi: 10.1109/ICITEED.2017.8250507]
35. Semrud-Clikeman M, Fine JG, Bledsoe J. Comparison among children with children with autism spectrum disorder, nonverbal learning disorder and typically developing children on measures of executive functioning. J Autism Dev Disord 2014 Feb;44(2):331-342. [doi: 10.1007/s10803-013-1871-2] [Medline: 23812759]
36. Chatzopoulos D, Bermejo C, Huang Z, Hui P. Mobile Augmented Reality Survey: From Where We Are to Where We Go. IEEE Access 2017;5:6917-6950. [doi: 10.1109/access.2017.2698164]
37. Kong G, LaVallee H, Rams A, Ramamurthi D, Krishnan-Sarin S. Promotion of Vape Tricks on YouTube: Content Analysis. J Med Internet Res 2019 Jun 18;21(6):e12709 [FREE Full text] [doi: 10.2196/12709] [Medline: 31215510]
38. Baker PM, Bujak KR, DeMillo R. The Evolving University: Disruptive Change and Institutional Innovation. Procedia Computer Science 2012;14:330-335. [doi: 10.1016/j.procs.2012.10.037]
39. Ifenthaler S, Sampson D, Yang D, Mukama L. Technology Enhanced Formative Assessment for 21st Century Learning. J Educ Technol Soc 2016;19:58-71.
40. First Case of COVID-19 confirmed in Rwanda. World Health Organization. 2020. URL: https://www.afro.who.int/news/first-case-covid-19-confirmed-rwanda [accessed 2022-04-04]
41. Stenhoff DM, Pennington RC, Tapp MC. Distance Education Support for Students With Autism Spectrum Disorder and Complex Needs During COVID-19 and School Closures. Rural Special Education Quarterly 2020 Oct 04;39(4):211-219. [doi: 10.1177/8756870520959658]
42. Michalak N. Coping with COVID-19: Remote Learning for Students with Autism Spectrum Disorder. 2020. URL: https://www.isbe.net/Documents/Remote-Learning-Students-with-ASD.pdf [accessed 2022-04-04]

Abbreviations

ASD: autism spectrum disorder
FGD: focus group discussion
ICT: information and communication technology

Edited by N Zary; submitted 27.02.21; peer-reviewed by E Toki, Z Reis, M Suppan; comments to author 30.04.21; revised version received 09.08.21; accepted 16.01.22; published 19.04.22.

Please cite as:
Ntalindwa T, Nduwingoma M, Uworwabayeho A, Nyirahabimana P, Karangwa E, Rashid Soron T, Westin T, Karunaratne T, Hansson H
Adapting the Use of Digital Content to Improve the Learning of Numeracy Among Children With Autism Spectrum Disorder in Rwanda: Thematic Content Analysis Study
JMIR Serious Games 2022;10(2):e28276
URL: https://games.jmir.org/2022/2/e28276
doi:10.2196/28276
PMID:35438638

Original Paper

Gamification in Diplomacy Studies as an Effective Tool for Knowledge Transfer: Questionnaire Study

Mihai Ovidiu Cercel[1], MBA, PhD

Department of International Relations and European Integration, National University of Political Studies and Public Administration, Bucharest, Romania

Corresponding Author:
Mihai Ovidiu Cercel, MBA, PhD
Department of International Relations and European Integration
National University of Political Studies and Public Administration
30A, Expozitiei Boulevard
Bucharest, 012104
Romania
Phone: 40 371 445 076
Fax: 40 21 312 25 35
Email: mihai.cercel@dri.snspa.ro

Abstract

Background: Graduate education in modern diplomacy poses several challenges, as it requires several competencies to be developed before diplomatic service is joined. Incorporation of simulation games can have a positive impact on the design of international relations and diplomacy learning process. We have designed a novel role play game (MAEDRI) to simulate part of the activities of a typical Ministry of Foreign Affairs.

Objective: This study aims to evaluate the effectiveness of MAEDRI in transferring knowledge in international relations education programs at the National University of Political Studies and Public Administration, Bucharest, Romania, across a 4-year period.

Methods: The game enrolled master's level graduate students. The data were collected through a voluntary and anonymous questionnaire between 2017 and 2020. At the end of each of the 4 editions we organized debriefing sessions that gave students the opportunity to provide feedback on their experience with this exercise, level of collaboration within the team, lessons learned, and to make suggestions for improvements. Using an online questionnaire, we measured the participants' perception regarding the level of effectiveness in increasing knowledge transfer, motivation, and engagement. Questionnaire data were consolidated in percentages for each item.

Results: A total of 49 participants completed the study. A total of 24 skills (13 professional and 11 social skills) were assessed. We identified a strong positive correlation between stress management and conflict management (r=.86; P<.001) as well as significantly positive correlations between building relations within the team and the ability to dialog and be persuasive (r=.7; P<.001), between procedure compliance and planning and organizing the work (r=.69; P<.001), and between analysis capacity and decision based on data received (r=.68; P<.001). Among social skills, self-control, confidence, and flexibility were the most substantially improved.

Conclusions: We describe several benefits of a novel game, used as an education tool to enhance a series of competencies necessary in international relations studies. Our results demonstrate a significant level of student engagement and motivation while playing MAEDRI, improvement of several essential skills, and enhanced knowledge transfer to real-life situations. While the data are encouraging, further research is needed to evaluate the full impact of role play as an effective experiential learning method.

(JMIR Serious Games 2022;10(2):e32996) doi:10.2196/32996

KEYWORDS

modern diplomacy; international relations; innovations in learning experience; gamification; serious games; role play design; knowledge transfer; competency development

Introduction

International relations and diplomacy are in constant transformation, mostly as a result of the necessary adaptation to the digital revolution that has led to the emergence of new platforms that cover events in real time, while enabling easy access to global communication. Innovations in teaching international relations are critical for the success of graduate programs in these fields, which should remain adaptable and reflective of the new, constantly evolving trends.

The fundamental mission of education is to transfer knowledge in various fields and to assess the competencies acquired, by measuring the ability of students to use in practice what they learnt theoretically [1,2]. Diplomatic skills are difficult to acquire exclusively from theoretical information. In addition, practicing diplomacy is almost impossible without being a career diplomat, given that diplomatic activities are largely based on confidential information. Therefore, one of the challenges when creating successful graduate programs in diplomacy and international negotiations consists of providing learning experiences that are as close to real-life scenarios as possible. To meet some of these challenges, we designed and implemented an innovative simulation game (SG).

SGs combine the features of serious games and simulation and have emerged as powerful tools to improve learning outcomes, by facilitating a better understanding of specific issues, via simulated experience [3-6]. Garris et al [7] argued that both serious game design and the gamification of learning increase learning outcomes, either directly (games) or indirectly through an alteration of contextual learner behavior. Vlachopoulos and Makri [8] see the usage of serious games in higher education as a necessary step forward in pedagogy toward a student-centered environment. Moreover, gamification has the potential to enhance motivation and induce behavioral changes, while fostering team work and promoting friendly competition in different work contexts [9-11].

Using role play learning and personalized learning to support the development of different competencies is not a new educational method, but it is a topic often described in recent years as a powerful tool to enhance learning by challenging students' creativity [12-15]. Dynamic grouping strategies are effective in enhancing students' learning [16,17].

A plethora of research studies have supported the use of SGs as a way to increase motivation, engagement, and learning outcomes [4,5,7,18-24] by enhancing attention, active learning, feedback, and consolidation, the 4 main pillars of learning [5,23].

The class curriculum in international relations courses includes, almost without exception, different diplomatic simulations. Published evidence from studies on experiential learning literature points to several impactful learning aspects of simulations, such as setting up objectives, creating opportunities for interactions, or enhancing teaching notes, but there is little research regarding the students' perception of relevant skill development [8,13,16,19]. As young adults, students are often engaged in cyclical, experiential learning, as described by Kolb's Learning Cycle [25]. As discussed by Duffy [26], individuals differ in their learning process, how they process information, and how they assimilate and use it in future actions. Kim [27] thinks that game theory, together with other behavioral disciplines, may offer a better understanding of political science concepts.

When teaching political science, especially diplomacy, one is confronted with a major challenge—how to develop relevant, applicable skills, taking into consideration the fact that in real-life situations (occurring, for example, in ministries of foreign affairs [MFAs]) there is a tangible need for access to confidential information. The benefits of using a mix of personalized learning and role playing in political and social sciences are well demonstrated in pedagogical literature [27-31]. Hardy and Totman [32] consider that the use of such mixed approaches in international relations requires discipline-specific examination. Analyses of international affairs and case studies are the main tools in the teacher's toolbox. Online diplomatic role play or simulations of debates in International Organizations (such as Model United Nations [UN] or the North Atlantic Treaty Organization [NATO] Conference model) are the most common simulations in all universities.

Nevertheless, such simulations are often inadequate because either they focus on specific situations from the past (and the students try to adopt the same, obsolete approach) or they use imaginary situations, designed to promote an understanding of the internal mechanisms of each diplomatic organization [33-35]. The question is whether role play simulations in international relations (online or face-to-face exercises) generate measurable and meaningful outcomes in developing specific skills [36-38]. Gamification helps students to better understand how theoretical concepts could be applied for solving real-life problems, and whether those skills are retained longer than classical learning methods [39,40].

To increase the impact of graduate-level education on diplomacy, we developed MAEDRI, an innovative SG that simulates parts of the diplomatic activity typically seen in international relations. We hypothesized that the application of game elements in learning diplomacy will enhance knowledge transfer. In addition, we hypothesized that game participants would develop or enhance social skills necessary for practicing international relations and be better prepared for their professional life. Between 2017 and 2020, a total of 150 master's level graduate students have played MAEDRI, which was administered as part of the Diplomacy and International Negotiation Master's Program at the National University of Political Studies and Public Administration, Bucharest, Romania. All student participants were, at the time of this study, enrolled in Master's Programs such as International Relations and European Integration, Conflict Analysis, Diplomacy and International Negotiation, and Security and Diplomacy. The study ran in 4 consecutive annual editions. Our aim was to evaluate the effectiveness of experiential learning through gamification, and to assess the impact of MAEDRI on student preparedness and their ability to develop specific capabilities, typically used by diplomats in their professional life. Ultimately, this study examines how gamification can boost students' competencies in international relations, improve their ability to apply theoretical concepts to practice, motivate them, and enhance their engagement during the educational process.

Methods

Overview

We developed MAEDRI, a new role play game that explores the interactions between real policy, politics, and students' capabilities. The game is primarily designed for participants who are studying international relations or political sciences, at graduate (master's) level. The MAEDRI role play has been organized annually since 2016.

Study Participants

All study participants are graduate students enrolled in master's level programs in International Relations, Diplomacy and International Negotiations, Conflict Analysis, and Security and Diplomacy at the National University of Political Studies and Public Administration, Bucharest, Romania.

All participants provided written consent to participate in the research. A total of 150 students who have played MAEDRI received a message containing information about the study via the MAEDRI Facebook group, which was restricted to students only. The participants were not chosen randomly, and participation in the research was voluntary.

Ethical Considerations

Per institutional guidelines on survey studies, the Department of International Relations and European Integration within the National University of Political Studies and Public Administration in Bucharest, Romania, deemed that this study met criteria for exemption from review by the Quality Assurance Committee.

Study Design

This study evaluated students' perceptions of effectiveness after playing MAEDRI. The data were collected through an anonymous questionnaire consisting of 9 matrix grid questions. The questionnaire was administered online, through the SurveyMonkey platform [41]. Research participants had to meet the inclusion criterion (ie, participation in one of the MAEDRI annual editions between 2017 and 2020). The questionnaire sought feedback on the participants' perception regarding their level on several key learning skills, and the extent the acquired competencies were put in practice in real life during their university studies and within their current work environment. In Romania, most graduate students are employed, and work during their studies. To determine whether the knowledge was transferred effectively, students were invited to self-evaluate on a range of professional and social competencies, before and after the role play simulation.

The questionnaire included response options along a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The online survey was conducted in accordance with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) checklist [42]. The average time to complete the questionnaire was 15 minutes.

Statistical Methods

Questionnaire data were consolidated in percentages for each item and analyzed using the MS Excel (Microsoft) data analysis module.

P-value was determined to validate the results of statistical analysis. A *P*-value <.001 was considered significant. Because of the exploratory nature of this study, we used the Cronbach α value. A high value for Cronbach α (<.8) indicates a good consistency of the items in the scale and helps validate the reliability of the questionnaire (Table 1). A 1-sample 2-tailed *t* test against the neutral value in the 5-point Likert rating was used to assess the responses to the 24 evaluated skills (Textbox 1).

Table 1. Normality and scale reliability results.

Skills	*P*-value	Cronbach α	*t* statistic (*df*)
Professional	<.001	.909	8.63 (48)
Social	<.001	.871	5.71 (48)

Textbox 1. Transferable skills through the MAEDRI game.

Professional skills

- Searching and filtering information in virtual space
- Problem solving
- Task partitioning
- Assertiveness
- Attention to details
- Giving and receiving feedback
- Decision making based on data received
- Analysis capacity
- Change management
- Planning and organizing the work
- Writing reports using a diplomatic language
- Fluency and concision
- Procedures compliance

Social skills

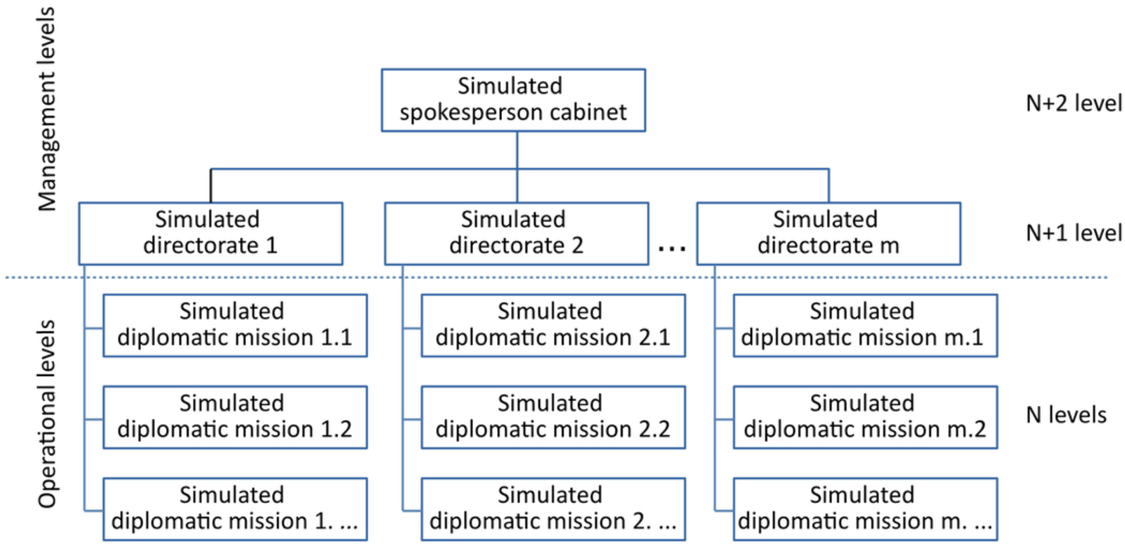
- Ability to dialog and be persuasive
- Building relations within the team
- Team coordination
- Self-control and confidence
- Flexibility
- Team motivation
- Initiative and creativity
- Care for order, quality of work, and accuracy
- Emotional intelligence
- Conflict management
- Stress management

Results

MAEDRI Game Design and Implementation

The preparation of the exercise includes advertising and promoting the game, selecting the "future diplomats" for simulated diplomatic missions, allocating the roles based on their interests and knowledge of different foreign languages, and promoting the activities on social media (Multimedia Appendix 1). The more senior, final-year students assume management roles at simulated headquarters. A system of communication and a chain of command are established to send information and receive feedback (Figure 1). Once roles are assigned, participants will conduct their research online, to identify all the relevant local media in their "host country" and select the most important media websites to generate a balanced portfolio. Google Translate or similar online tools were employed for translation to Romanian.

Figure 1. The organizational chart of the simulated organization - MAEDRI.



Every day during the role play game, the students in simulated diplomatic missions (N level) were asked to monitor all official media webpages selected during the preparation period in the "host" country (Figure 1). They also analyzed, selected, and filtered the information, focusing on several criteria, such as the specifics of bilateral relations between Romania and the country where the simulated diplomatic mission was based; the political, economic, social, and cultural environment in the "allocated" country; and the positions expressed by the representatives of the allocated country regarding the international topics. The participants were required to provide a daily written report (with a limit of 400 words per article) by observing the diplomatic style in terms of language and concision.

The resulting report was subsequently sent to the next decision level—the simulated directorates—organized on geographical or functional criteria by grouping 4-7 simulated diplomatic missions. "Late" reports, that is, those submitted after the agreed upon deadline, were generally not considered. Exceptions were made only for unexpected events occurring in the "simulated" country.

The chain of command simulates the various divisions of an ordinary MFA. In the MAEDRI game (Figure 1), we simulate only 2 management levels: simulated directorates (N+1 level) and simulated spokesperson cabinet (N+ 2 levels). For the chain of command, the tasks focused on analyzing and comparing information obtained from various sources, filtering reports based on the setup criteria, studying the allocated geographic space, and compiling the selected information using a diplomatic language, to finalize the report in a format that is relevant and interesting to the public. The decisions were made collectively in groups of 7-9 students, similar to a newspaper editorial office.

The simulated spokesperson cabinet is the last and the highest hierarchical structure in the simulated MFA. This structure is responsible for selecting the most relevant articles received from the general directorates and publishing them on the MAEDRI simulated MFA Facebook page (Multimedia Appendix 2, in Romanian only).

At the end of each edition debriefing sessions were organized that gave students the opportunity to provide feedback on their experience with this exercise, level of collaboration within the team, lessons learned, and to make suggestions for improvements.

The skills acquired during role play can be separated into 2 categories: professional and social (Textbox 1).

Approximately 10% (5/49) of the respondents participated in 2 successive editions of the role play game (first time as lower-level "diplomats" in the simulated diplomatic missions and second time as "diplomats" in a managerial position), but the online platform used for data collection allowed a single questionnaire to be completed from the same IP address.

Characteristics of Study Participants

The study involved 49 respondents out of 150 students who participated in the MAEDRI simulations. The demographic characteristics of the 49 participants are representative of the entire group of 150 in terms of sex, age, and work status. Table 2 presents their demographic features, their participation in MAEDRI simulations, and employment situation.

Table 2. Descriptive characteristics of participants (n=49).

Characteristics	Value
Sex, n (%)	
Male	11 (22)
Female	38 (78)
Age (years), n (%)	
20-24	17 (35)
25-30	24 (49)
Over 30	6 (12)
Unknown	2 (4)
Participant in, n (%)	
Edition 2017	11 (22)
Edition 2018	8 (16)
Edition 2019	9 (18)
Edition 2020	21 (43)
Employment situation, n (%)	
Working in the international relations field	17 (35)
Working in other sectors	20 (41)
Unemployed	11 (22)
Unknown employment status	1 (2)

Correlations of Measurements

A strong positive correlation (Table 3 and Multimedia Appendix 3) was found between stress management and conflict management (*r*=.86; *P*<.001) as well as significantly positive correlations between building relations within the team and the ability to dialog and be persuasive (*r*=.7; *P*<.001), between procedures compliance and planning and organizing the work (*r*=.69; *P*<.001), and between analysis capacity and decision based on data received (*r*=.68; *P*<.001). No negative correlations were found.

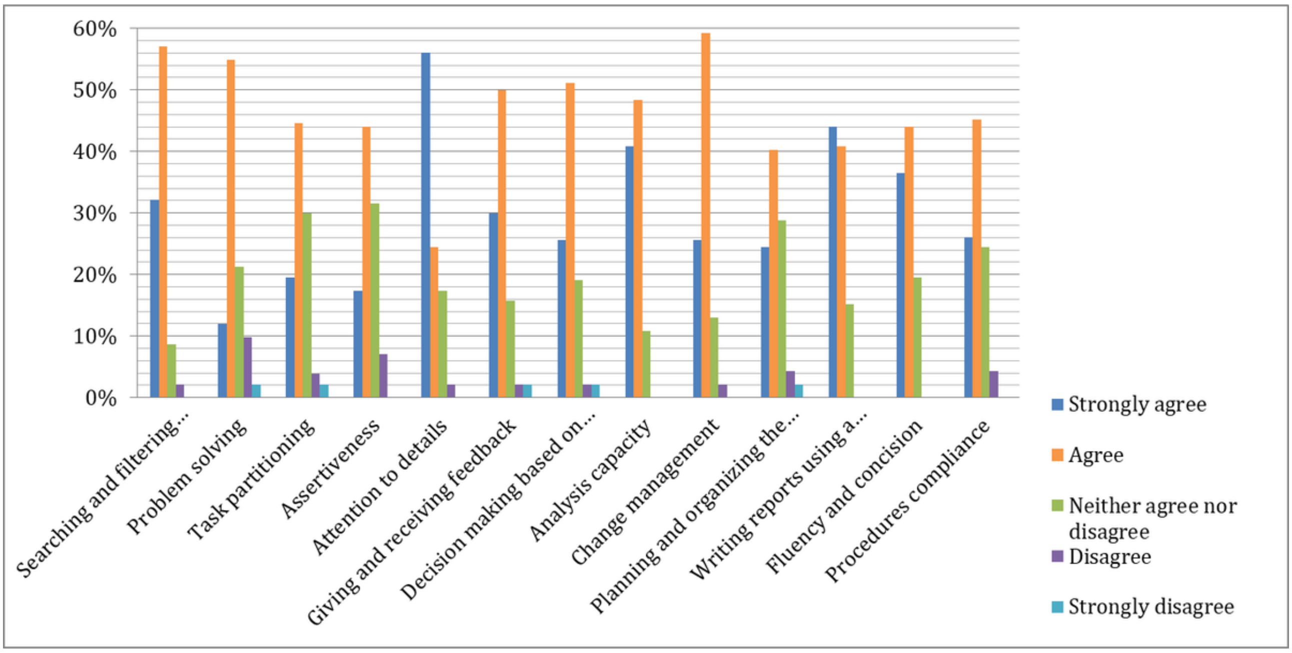
Table 3. Most significant correlations.

Skills	Decision making	Planning and organizing the work	Ability to dialog and be persuasive	Conflict management
Analysis capacity	0.6851	0.3702	0.3794	0.5543
Procedures compliance	0.2638	0.6908	0.3523	0.2237
Building relations within the team	0.4338	0.3283	0.7016	0.2959
Stress management	0.5739	0.4708	0.3339	0.8568

The MAEDRI Simulation Game Enables Knowledge Transfer and Improves Professional Competencies Diplomacy Studies

Most study participants either strongly agreed or agreed that the role play helped them in developing professional skills in the field of international relations (Figure 2). Approximately 39% (19/49) of participants either strongly agreed or agreed that the simulation helped them to better understand the theoretical concepts and to put in practice what they learned. The most remarkable results are observed for searching and filtering information in virtual space, where 88% (43/49) of participants strongly agreed and agreed that the role play simulation improved their competencies. The same tendency for attention to details was observed, where 73% (36/49) of participants strongly agreed and agreed.

Figure 2. Improvement of professional competencies after MAEDRI game.



The MAEDRI Game Enhances Social Skills Necessary for a Professional Career in International Relations

Regarding social skills (Figure 3), we found the same tendency as with professional skills, although the approval rates were not as strong. During the game, several skills were emphasized and participants recognized that self-control, confidence, and flexibility were most substantially improved. Participants also agreed that the ability to dialog and be persuasive, development of emotional intelligence, care for order, attention to quality of work, and accuracy were all important skills that were developed by the simulation. In comparison with professional skills, 53% (26/49) of respondents thought that the exercise did not improve their social skills (or had only a minimal impact). This may be partly explained by the fact that, at least at the master's level, students already have several years of working experience, and those skills may have already been developed in their working environments. In the second part of the questionnaire, respondents were asked if they used the specified competencies in the academic environment or in active life environment. In line with our hypotheses, our results show that the competencies developed by the game are important for preparing the students for the active life (Figure 4).

We found 3 exceptions where the respondents thought that those competencies have been slightly more useful in the academic environment than during their active life. These findings are obvious, and some competencies, such as searching and filtering information, writing reports using a diplomatic language, and analysis capacity, are used intensively during university studies. The same competencies are found useful in the active life only by the participants who are working in international relations–related workplaces.

Regarding social skill usage, we found the same pattern (Figure 5). Although the respondents considered that those skills existed to a certain extent even before the simulation, an important part of respondents indicated that the usage of those competencies was more helpful during their active life. Further research must be performed to investigate this trend.

Figure 3. Improvement of social skills after MAEDRI game.

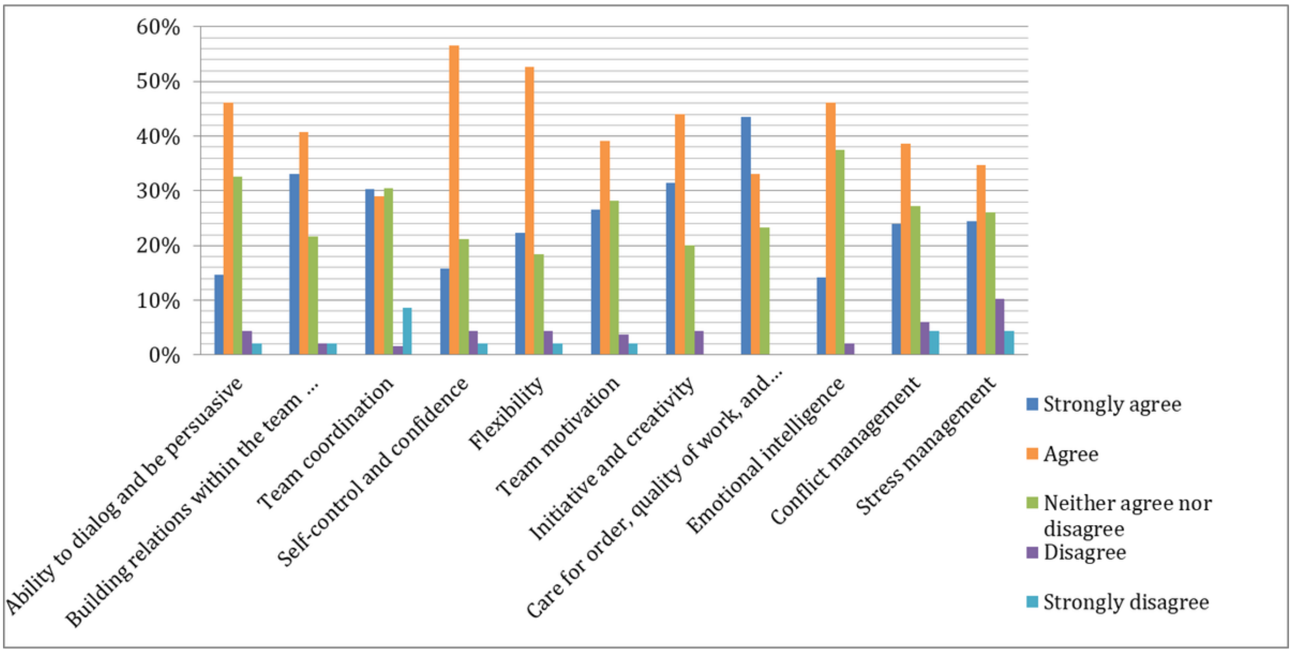


Figure 4. Usage of the acquired professional competencies in the academic environment and in active life.

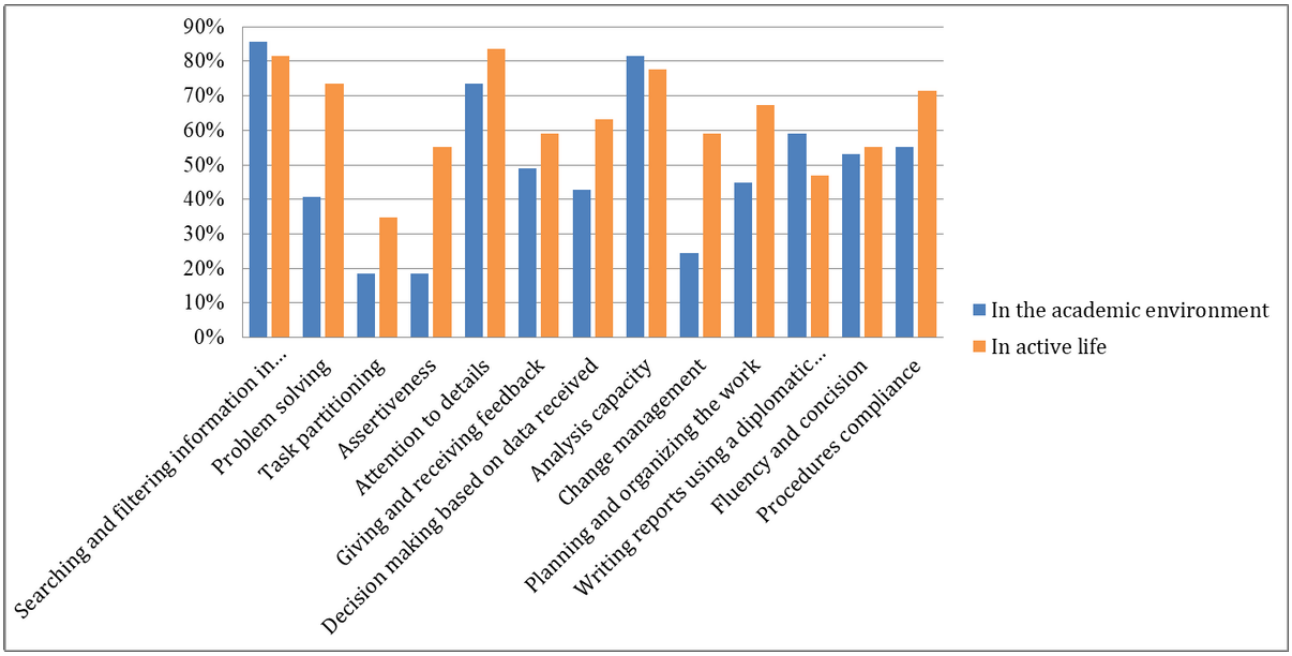
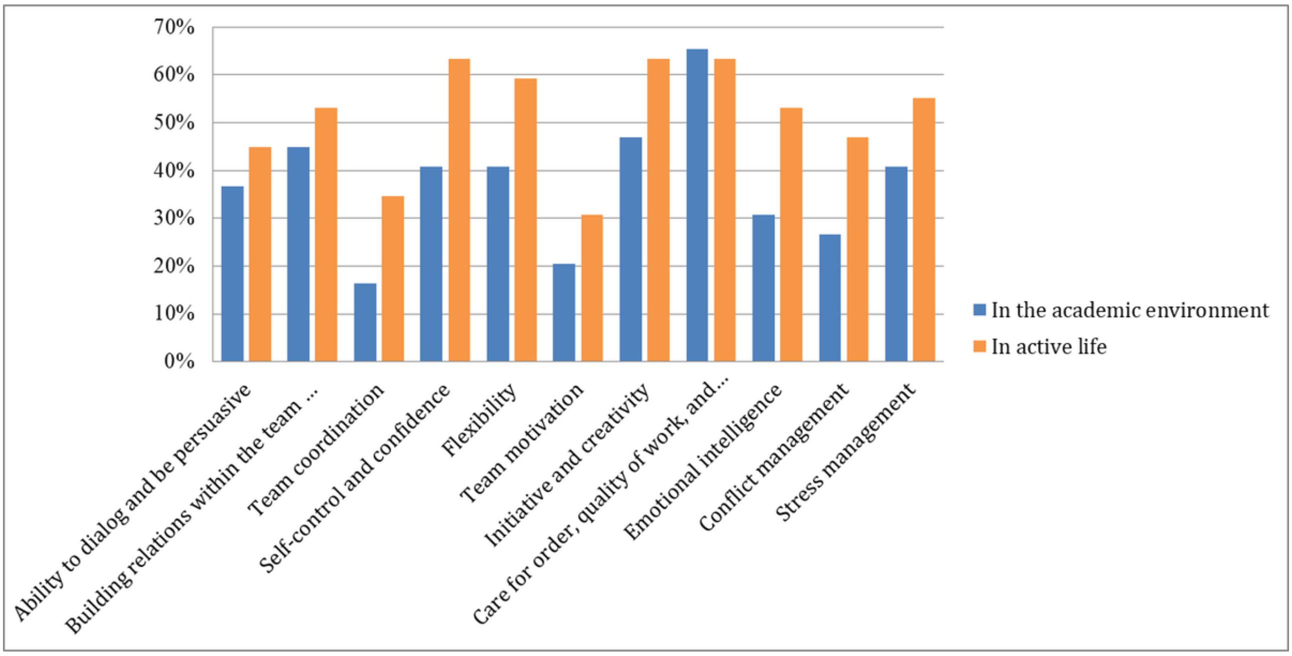


Figure 5. Usage of the acquired social skills in the academic environment and in the active life.



Discussion

Principal Findings

We developed the MAEDRI role play game and tested its impact on experiential learning in international relations and diplomacy studies. Our results reveal that MAEDRI can be easily implemented in the curriculum and can serve as an effective tool to boost knowledge and to form or enhance a series of competencies transferrable to the professional life. Our studies cover 4 years of annual simulations and were designed to analyze mainly the transfer of a number of professional competencies and social skills. In addition to confirming our underlying hypotheses, our results demonstrate that the overall effect of the learning game is in fact more complex and impacts, albeit at a lower extent, several social competencies. It was rewarding to observe that some of those achievements were used not only during but also after the instruction period, pointing to the game's potentially long-lasting impact. It is remarkable that this conclusion is supported even by former students who ended up choosing a different career path and are currently not working in international relation–related fields.

Most participants appreciated the role play experience and provided positive feedback, pointing to their appreciation for opportunities to apply theoretical concepts into an experiential learning exercise, which closely mirrors the diplomatic real life.

The success of the MAEDRI simulation was observed in students' motivation in maintaining a high level of day-to-day active participation in the exercise. In this respect, we observed good competition between different teams, striving to improve the postreach and postengagement rates of their reports, on the MAEDRI Facebook page. Our research is in line with the literature findings [29,43-45], where an increase in engagement by simulation was demonstrated in teaching political sciences. Moreover, for students who cannot participate in an internship, the exercise may generate equivalent competency acquisition.

The MAEDRI design offers a pragmatic pedagogical opportunity to demonstrate through experiential learning how diplomacy works in a day-to-day activity, as well as to engage in teamwork, enhance coordination within a hierarchical organization, increase work accuracy, and master time management. This is in line with how students are guided to discover knowledge through simulated experiences [46]. The MAEDRI role play simulation responds to the elements identified by Wilcox [46]: context (the students learn easier because they recognize the situation), practice (the exercise allows them to experiment in a safe manner with the theoretical concepts learned during their college studies), and experience (the students discover themselves what are the diplomatic day-to-day tasks). The experiential learning exercised during the MAEDRI role play focuses on developing or enhancing a number of capabilities, previously reported to be essential for political and social sciences [11,27-31].

Additionally, this data collection provides an insight into what activities are appreciated by the students and determines what could be done for improvement. The data from 4 consecutive game editions suggest that role play is a welcome alternative to the "classical" teaching approach often considered to be too theoretical by many Romanian students, especially those majoring in political sciences, international relations, or diplomacy.

The hierarchical organization that we created generates interactions between students within each compartment and between compartments in vertical hierarchies. We acknowledge that horizontal interactions between compartments situated on the same level were minimal or even null. This may be explained by the fact that the students are enrolled in different master's programs and even when in the same program, they may be in different years of study. Additionally, for students who are not simultaneously enrolled in the same courses, interactions and collaborations outside the classwork may be more difficult to establish. This exercise confirmed that tendency, revealing that participants interacted almost exclusively within their group or with colleagues they were acquainted with. Next steps focus on improving the exercise by enhancing interactions between compartments of the same rank, to promote the team spirit and create a feeling of belonging to a wider community within our university.

This project was designed to evaluate the capacity of our newly designed MAEDRI role play game to enhance knowledge transfer and acquisition of key skills for graduate students preparing for a career in international relations and diplomacy. Although we used a nonprobabilistic sample, we consider that the results are relevant to validate the research hypotheses.

The students' feedback stressed the fact that the exercise reached its main goal: to allow them to put in practice theoretical concepts. Moreover, it highlighted the role of this game exercise as a motivating factor, providing a better understanding of real diplomats' tasks and offering a snapshot of real professional situations where both their hard and soft skills have been exposed and enhanced. Ultimately, we demonstrate that this role play exercise had a positive effect on knowledge transfer and enabled enhancement of several competencies.

Limitations

We acknowledge several study limitations. Because of the small sample size, this is an exploratory study and further studies with increased sample size (and control arm including) are needed to validate these findings. The number of questionnaire respondents represents approximately 33% (49/150) of the total number of participants and our current focus is on identifying effective strategies to improve uptake.

We also note that the real impact of certain factors is difficult to assess; for example, those that might contribute to skill development in the workplace. Further research will try to evaluate the influence of such factors in the immediate as well as in long-term professional development.

Another limitation refers to a lack of control group. Further research will address this limitation.

Conclusions

In summary, we have designed and implemented a new role play game and acquired data from 4 consecutive years of annual simulations. Our study demonstrates that the incorporation of role play is an effective experiential learning method that helps graduate students in master's programs to better understand and use international relations and diplomacy concepts and enhance professional competencies. Most study participants provided positive feedback and the MAEDRI game was easy to implement. We acknowledge several limitations that must be considered when evaluating and interpreting the results. Further steps should be taken to investigate how the simulation exercise may enhance interactivity between different teams situated on the same level, and further stimulate the development of social skills.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MAEDRI simulation game preparation checklist.
[DOC File , 46 KB - games_v10i2e32996_app1.doc]

Multimedia Appendix 2
MAEDRI Facebook page.
[DOC File , 656 KB - games_v10i2e32996_app2.doc]

Multimedia Appendix 3
Table of correlations.
[DOC File , 283 KB - games_v10i2e32996_app3.doc]

References

1. Kapp KM. The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education. San Francisco, CA: Pfeiffer; 2012:9-18.
2. Epstein RM, Hundert EM. Defining and assessing professional competence. JAMA 2002 Jan 09;287(2):226-235. [doi: 10.1001/jama.287.2.226] [Medline: 11779266]
3. Fonteneau T, Billion E, Abdoul C, Le S, Hadchouel A, Drummond D. Simulation game versus multiple choice questionnaire to assess the clinical competence of medical students: prospective sequential trial. J Med Internet Res 2020 Dec 16;22(12):e23254 [FREE Full text] [doi: 10.2196/23254] [Medline: 33325833]
4. Leung R. Evaluating the benefits of collaboration in simulation games: the case of health care. JMIR Serious Games 2014 Jan 28;2(1):e1 [FREE Full text] [doi: 10.2196/games.3178] [Medline: 25658851]
5. 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 [FREE Full text] [doi: 10.2196/16096] [Medline: 32347811]

6. Omer S, Choi S, Brien S, Parry M. Log in to experiential learning theory: supporting web-based faculty development. JMIR Med Educ 2017 Sep 27;3(2):e16 [FREE Full text] [doi: 10.2196/mededu.7939] [Medline: 28954718]
7. Garris R, Ahlers R, Driskell JE. Games, motivation, and learning: a research and practice model. Simulation & Gaming 2016 Aug 18;33(4):441-467. [doi: 10.1177/1046878102238607]
8. Vlachopoulos D, Makri A. The effect of games and simulations on higher education: a systematic literature review. Int J Educ Technol High Educ 2017 Jul 10;14(1):22. [doi: 10.1186/s41239-017-0062-1]
9. Harrigan P, Wardrip-Fruin N. Second Person: Roleplaying and story in playable media. Cambridge, MA: MIT University Press; 2007:0262514187.
10. Chen C, Liu G, Hwang G. Interaction between gaming and multistage guiding strategies on students' field trip mobile learning performance and motivation. Br J Educ Technol 2015 Mar 16;47(6):1032-1050. [doi: 10.1111/bjet.12270]
11. Giannakas F, Kambourakis G, Papasalouros A, Gritzalis S. A critical review of 13 years of mobile game-based learning. Education Tech Research Dev 2017 Dec 12;66(2):341-384. [doi: 10.1007/s11423-017-9552-z]
12. David RM, Chen S. Serious Games: Games that Educate, Train, and Inform. Boston, MA: Taxmann Publications Private Limited; Oct 11, 2005:1-312.
13. Mustata IC, Alexe CG, Alexe CM. Developing competencies with the General Management II business simulation game. Int. J. Simul. Model 2017 Sep 15;16(3):412-421. [doi: 10.2507/ijsimm16(3)4.383]
14. Pivec M. Editorial: Play and learn: potentials of game-based learning. Br J Educ Technol 2007 May;38(3):387-393. [doi: 10.1111/j.1467-8535.2007.00722.x]
15. Zhang L, Basham JD, Yang S. Understanding the implementation of personalized learning: A research synthesis. Educational Research Review 2020 Nov;31:100339. [doi: 10.1016/j.edurev.2020.100339]
16. Chen B, Hwang G, Lin T. Impacts of a dynamic grouping strategy on students' learning effectiveness and experience value in an item bank‐based collaborative practice system. Br J Educ Technol 2019 Apr 29;51(1):36-52. [doi: 10.1111/bjet.12794]
17. Johnson DW, Johnson RT. An educational psychology success story: social interdependence theory and cooperative learning. Educational Researcher 2009 Jun 01;38(5):365-379. [doi: 10.3102/0013189x09339057]
18. Landers RN. Developing a theory of gamified learning. Simulation & Gaming 2015 Jan 15;45(6):752-768. [doi: 10.1177/1046878114563660]
19. Smiderle R, Rigo SJ, Marques LB, Peçanha de Miranda Coelho JA, Jaques PA. The impact of gamification on students' learning, engagement and behavior based on their personality traits. Smart Learn. Environ 2020 Jan 09;7(1):3. [doi: 10.1186/s40561-019-0098-x]
20. Charsky D. From edutainment to serious games: a change in the use of game characteristics. Games and Culture 2010 Feb 11;5(2):177-198. [doi: 10.1177/1555412009354727]
21. 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]
22. 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 Jan 10;10(1):e33282 [FREE Full text] [doi: 10.2196/33282] [Medline: 35006080]
23. 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]
24. van Benthem J. What Logic Games Are Trying to Tell Us. Amsterdam, The Netherlands: ILLC Publications; 2003. URL: https://eprints.illc.uva.nl/id/eprint/89/1/PP-2003-05.text.pdf [accessed 2021-06-05]
25. Kolb DA. Experiential Learningxperience as the Source of Learning and Development. 2nd Edition. Upper Saddle River, NJ: Pearson Education Inc; 2015:E-416.
26. Duffy SP. Teaching with a mallet: conveying an understanding of systemic perspectives on international relations intuitively-croquet as experiential learning. Int Studies Perspectives 2001 Nov;2(4):384-400. [doi: 10.1111/1528-3577.00067]
27. Kim S. Behavioral learning game for socio-physical IoT connections. J Wireless Com Network 2016 Jan 19;2016(1):1-11. [doi: 10.1186/s13638-016-0521-8]
28. O'Donell AM, Hmelo-Silver CE, Erkens G. (Eds.). Collaborative learning, reasoning, and technology. New York: Routledge; 2006:0805847789.
29. Rivera SW, Simons JT. Engaging students through extended simulations. Journal of Political Science Education 2008 Jul 29;4(3):298-316. [doi: 10.1080/15512160802202805]
30. Sommerhoff D, Szameitat A, Vogel F, Chernikova O, Loderer K, Fischer F. What do we teach when we teach the learning sciences? A document analysis of 75 graduate programs. Journal of the Learning Sciences 2018 Feb 16;27(2):319-351. [doi: 10.1080/10508406.2018.1440353]
31. Swennen A. Experiential learning as the 'new normal' in teacher education. European Journal of Teacher Education 2020 Oct 14;43(5):657-659. [doi: 10.1080/02619768.2020.1836599]
32. Hardy M, Totman S. Teaching an old game new tricks: Long-term feedback on a re-designed online role play. Br J Educ Technol 2016 Aug 03;48(6):1260-1272. [doi: 10.1111/bjet.12498]
33. Brynen R. (Ending) civil war in the classroom: a peacebuilding simulation. APSC 2010 Jan 15;43(01):145-149. [doi: 10.1017/s1049096510990719]

34. Chasek PS. Power politics, diplomacy and role playing: simulating the UN Security Council's response to terrorism. Int Studies Perspectives 2005 Feb;6(1):1-19. [doi: 10.1111/j.1528-3577.2005.00190.x]
35. Dougherty BK. Byzantine politics: using simulations to make sense of the Middle East. APSC 2003 Apr 15;36(02):239-244. [doi: 10.1017/s1049096503002154]
36. Blum A, Scherer A. What creates engagement? An analysis of student participation in ICONS simulations. 2007 Feb 11 Presented at: Proceedings of the APSA Teaching and Learning Conference; February 9-11, 2007; Charlotte, NC p. 9-11 URL: https://www.icons.umd.edu/papers/blum_scherer.pdf
37. Raymond C. Do role-playing simulations generate measurable and meaningful outcomes? A simulation's effect on exam scores and teaching evaluations. Int Stud Perspect 2010 Feb 2;11(1):51-60. [doi: 10.1111/j.1528-3585.2009.00392.x]
38. Stover WJ. Teaching and learning empathy: an interactive, online diplomatic simulation of Middle East conflict. Journal of Political Science Education 2005 May;1(2):207-219. [doi: 10.1080/15512160590961801]
39. Dadheech A. The importance of game based learning in modern education. The Knowledge Review. 2018. URL: https://theknowledgereview.com/importance-game-based-learning-modern-education/ [accessed 2021-06-21]
40. Billett S. Recasting transfer as a socio-personal process of adaptable learning. Educational Research Review 2013 Jan;8:5-13. [doi: 10.1016/j.edurev.2012.05.004]
41. SurveyMonkey. URL: https://www.surveymonkey.com/ [accessed 2022-04-13]
42. 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]
43. Giovanello SP, Kirk JA, Kromer MK. Student perceptions of a role-playing simulation in an introductory International Relations course. Journal of Political Science Education 2013 Apr;9(2):197-208. [doi: 10.1080/15512169.2013.770989]
44. Mariani M, Glenn BJ. Simulations build efficacy: empirical results from a four-week congressional simulation. Journal of Political Science Education 2014 Jul 30;10(3):284-301. [doi: 10.1080/15512169.2014.921623]
45. Owston R. Blended learning policy and implementation: Introduction to the special issue. The Internet and Higher Education 2013 Jul;18:1-3. [doi: 10.1016/j.iheduc.2013.03.002]
46. Wilcox S. Praxis games a design philosophy for mobilizing knowledge through play. Am J Play 2019;11(2):156-182 [FREE Full text] [doi: 10.1057/9781137307675.0007]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
MFA: Ministry of Foreign Affairs
NATO: North Atlantic Treaty Organization
SGs: simulation games
UN: United Nations

Edited by N Zary; submitted 18.08.21; peer-reviewed by D Paraschiv, A Ramazan, V Verma; comments to author 23.12.21; revised version received 13.03.22; accepted 16.03.22; published 25.04.22.

Please cite as:
Cercel MO
Gamification in Diplomacy Studies as an Effective Tool for Knowledge Transfer: Questionnaire Study
JMIR Serious Games 2022;10(2):e32996
URL: https://games.jmir.org/2022/2/e32996
doi:10.2196/32996
PMID:35468081

Original Paper

Effectiveness and Utility of Virtual Reality Infection Control Simulation for Children With COVID-19: Quasi-Experimental Study

Mi Yu[1], RN, PhD; Mi Ran Yang[2], RN, PhD

[1]College of Nursing, Institute of Health Sciences, Gyeongsang National University, Jinju, Republic of Korea

[2]Department of Nursing, Kwangju Health University, Kwangju, Republic of Korea

Corresponding Author:
Mi Yu, RN, PhD
College of Nursing
Institute of Health Sciences
Gyeongsang National University
816 beongil-15
Jinju-daero
Jinju, 52727
Republic of Korea
Phone: 82 10 5258 7075
Fax: 82 55 772 8222
Email: yumi825@gnu.ac.kr

Abstract

Background: It is essential that nurses quickly learn the proper methods for preventing and controlling nosocomial infection and managing intensive care patients during the COVID-19 pandemic, including the donning and doffing of personal protective equipment (PPE). Virtual reality (VR) simulation offers the advantage of learning in a safe environment with a sense of realism similar to that of an actual clinical setting and has been reported to enhance self-efficacy in infection control, safety performance, and learning satisfaction among students.

Objective: This study aims to develop a virtual reality infection control simulation (VRICS) program regarding donning and doffing of PPE and respiratory care for pediatric patients admitted to an isolation unit for COVID-19 and to identify the effects of the program on PPE knowledge, infection control performance, and self-efficacy for nursing students. Additionally, the realism of the VRICS program and the students' level of satisfaction with the program were assessed.

Methods: This was a quasi-experimental study based on a controlled pretest-posttest design. Third- and fourth-year nursing students were divided into an experimental group (n=25) who participated in a VRICS program and a control group (n=25) with no participation. Data were collected from November 13 to December 10, 2021, and analyzed using descriptive statistics and the *t* test, paired *t* test, Mann-Whitney U test, and Wilcoxon matched-pair signed-rank test. The VRICS program consisted of a prebriefing, including direct practice of donning and doffing PPE, VR simulation, and debriefing. The VR simulation comprised 3 sessions: donning and inspection of PPE in the dressing room before entering the negative-pressure isolation unit; assessing for suction care, nasopharyngeal suctioning, and checking of COVID-19 patients in the negative-pressure isolation unit; and doffing PPE in the dressing room. The total execution time for the program was 180 min.

Results: Compared with the control group, the experimental group showed significantly greater improvements in PPE knowledge (z=–3.28, P<.001), infection control performance (t_{48}=4.89, P<.001), and self-efficacy ($t_{36.2}$=4.93, P<.001). The experimental group's mean scores for realistic immersion and learner satisfaction were 4.49 (SD 0.50) points and 4.75 (SD 0.38) points (on a 5-point Likert scale), respectively.

Conclusions: The VR simulation training program involving pediatric COVID-19 patients combined skills training effectively and enhanced theoretical knowledge, respiratory care skills, and infectious disease preparedness. Thus, it could be applied to training nurses to respond more effectively to public health situations involving infectious diseases, including the COVID-19 pandemic.

(JMIR Serious Games 2022;10(2):e36707) doi:10.2196/36707

KEYWORDS

children; infection control; nursing student; simulation training; virtual reality; digital health; medical education; COVID-19; patient management; pediatrics; nursing education; respiratory care skills; program usability; digital learning

Introduction

Health care professionals, including nurses, who care for COVID-19 patients, are at direct or indirect risk of exposure to the virus or other infectious substances. Failure to comply with the appropriate control guidelines can make them a source of infection or a mediator that can spread COVID-19 [1]. Despite the vital role they play, most clinical nurses believe that they lack experience and knowledge in disaster nursing, such as COVID-19, and rate their personal emergency preparedness low [2]. Thus, it is essential that nurses quickly learn the proper methods, including the donning and doffing of personal protective equipment (PPE) for preventing and controlling nosocomial infection and managing intensive care patients during the COVID-19 pandemic [3]. In line with this, thorough and systematic education and training are needed beginning at the nursing undergraduate level. Nursing students need to experience nursing in actual clinical settings, but they get limited opportunities to personally observe and practice COVID-19-related infection control nursing, as well as acquire sufficient nursing skills. In particular, practicing pediatric nursing, including care of infants, is even more challenging due to the increased vulnerability of the child to infection [4]. Therefore, college instructors need to develop curricula to enable nursing students to cope with infectious diseases. However, a large gathering of students should be avoided in the middle of a pandemic to avoid the risk of viral transmission, and thus, new training methods should be applied. Accordingly, nursing training programs using advanced technologies are being developed [5], including a growing trend in virtual reality (VR) [4,6,7].

VR refers to a technology that creates a simulation of real surroundings with people actually experiencing an environment that may be difficult to experience in daily life [8]. VR simulation-based training deals with virtual patients within VR space instead of actual patients and thus does not threaten the safety of patients, allowing the learners to repeatedly practice comfortably on their own within the virtual space [8]. Moreover, it is a user-centric learning method that does not depend on preparations, personnel, and schedule, while allowing the users to practice various clinical techniques within a safe environment, unlike conventional high-fidelity simulation [9]. A recent meta-analysis [10] and systematic review [11] indicated that the use of VR simulation has the potential to produce educational outcomes similar or superior to those of traditional simulation and reported the positive effects of VR on self-efficacy [4,12] and learning satisfaction [4,12,13] among nursing and medical students. Meanwhile, there are mixed results, too, showing that performance increased in the experimental group [12-14] or there was no difference between groups [15].

Yu et al [4] developed a VR program for high-risk neonatal infection control and applied it to nursing students. The results showed improved confidence in infection control among the students, indicating the applicability of VR simulation as a simulation-based training program for nursing students. The study also reported that VR simulation offers the advantage of learning in a safe environment with a sense of realism similar to that of an actual clinical setting, but without major time constraints, and how the instructor carries out the lesson. Moreover, Birrenbach et al [13] compared an immersive VR simulation with a traditional learning method for a COVID-19-related skill set. This before-and-after training for medical students involved the performance of hand disinfection, nasopharyngeal swabbing, and the donning and doffing of PPE. The results showed that safe performance scores, such as for nasopharyngeal swabs, increased more in the VR simulation group than in the control group. Moreover, VR simulation provided user satisfaction, while remaining as effective as conventional learning methods for medical students. Such findings demonstrated that VR could be a useful tool for acquiring simple and complex clinical skills.

Meanwhile, prebriefing, with its importance in simulation emphasized recently [16], is a structured simulation stage before the scenario stage and is carried out by the instructions of the simulation moderator based on the experience and knowledge of the participants. It also includes activities for preparing how to use equipment or supplies in the simulation processes. It can also be viewed as a stage for creating a safe and reliable learning environment to promote participation and help achieve learning goals [16,17]. Accordingly, this study used prebriefing based on the self-efficacy theory by Bandura [18] and designed the program to enhance self-efficacy [19] through dynamic learning experience by actually performing the actions and proxy experience from observing the actions of others [20]. Individuals who have high self-efficacy will exert effort that, if well-executed, leads to successful performance and outcomes [18], such as clinical competency and simulation performance [21].

However, there are limited VR simulation programs on medical conditions, treatment, and patient care for training nurses or nursing students for responding to respiratory infectious disease epidemics. In particular, infection control in hospitals has become more important due to the COVID-19 pandemic and PPE, such gloves, disposable gowns, N95 masks, protective goggles, and shoe covers, can be a primary physical barrier against infection. Despite this, it is difficult to find virtual reality infection control simulation (VRICS) programs applied to pediatric patients for donning PPE, which is not normally used and may appear unfamiliar and complicated. This study thus aims to develop and test a VR simulation program incorporating pediatric COVID-19 cases.

The objective of this study was to develop a VRICS program regarding donning and doffing of PPE and respiratory care for pediatric patients admitted to an isolation unit for respiratory infectious disease and apply the program to nursing students to identify the effects of the program on PPE knowledge, infection control performance, and self-efficacy, as well as the realism of the VRICS program and the level of satisfaction with the program. Accordingly, the study established the following hypotheses:

- Hypothesis 1: The experimental group (participation in the VRICS program) and control group (no participation) will show a difference in PPE knowledge.
- Hypothesis 2: The experimental and control groups will show a difference in infection control performance.
- Hypothesis 3: The experimental and control groups will show a difference in self-efficacy.

Methods

Study Design

This was a quasi-experimental study based on a controlled pretest-posttest design for the development of a VRICS program with the objective of comparing PPE knowledge, infection control performance, and self-efficacy between the participating and the nonparticipating group of nursing students, as well as to identify the level of satisfaction with the program.

Study Population

The target population was all nursing students in Korea, and the study selected third- and fourth-year nursing students from a nursing college in "J" City in Gyeongsangnam Province. The inclusion criteria consisted of nursing students currently enrolled in a nursing college who have clinical practice experience, completed courses on pediatric nursing and a core basic nursing skill course titled "Practice for Donning and Doffing Standard PPE and Managing Medical Waste" in the regular curriculum, and voluntarily consented to participate in the study. The exclusion criterion was the lack of consent to participate in the study. The sample size for the study was calculated using G*Power version 3.1.9.7 [22]. Considering that the effect size was 0.40-0.72 in previous studies, which is similar to that in this study [23-25], a moderate effect size (f) of 0.50, a significance level (α) of .05, and a statistical power (1 – β) of .90 were used, with a suitable sample size (2-tailed test) of 44. Considering a dropout rate of 10%, a total of 50 participants were recruited for the final study population, with 25 (50%) participants each in the experimental and control groups.

Study Instruments

General Characteristics

The general characteristics of the participants included age, sex, previous semester grades, and VR experience.

PPE Knowledge

PPE knowledge was measured using the tool originally developed by Choi [26] for PPE knowledge related to acute respiratory infection, which was modified and supplemented to be suitable for COVID-19. The tool consisted of 20 items: 2 items on the transmission route, 3 items on hand hygiene, and 15 items on donning and doffing PPE. Each correct answer was given 1 point, and a wrong answer was given 0 points. The total score ranged from 0 to 20 points, with higher scores indicating a higher knowledge level. The content validity index of the tool at the time of development was 0.8-1.0 for each item, while the Kuder-Richardson Formula 20 (KR-20) of the knowledge items in the study was 0.68.

Infection Control Performance

Infection control performance was measured based on items related to PPE and pediatric respiratory care. PPE-related items were measured using the tool originally developed by Kwon [27] for measuring PPE use among nurses in specialized infectious disease hospitals, which was modified and supplemented according to the actual performance procedures of this study. The tool consisted of 20 items, and each item was rated on a 5-point Likert score. The total score ranged from 20 to 100 points, with higher scores indicating higher performance. The reliability of the tool (Cronbach α) was .86 in the study by Kwon and .97 in this study.

Self-efficacy

Self-efficacy refers to the personal belief about whether something new that has been learned can be applied [28]. Self-efficacy was measured using the tool originally developed by Ayres [28] and adapted by Park and Kwon [29] for simulation studies. The tool consisted of 10 items, and each item was rated on a 7-point scale (1=not at all to 7=very much so). The total score ranged from 7 to 70 points, with higher scores indicating higher self-efficacy. The reliability of the tool (Cronbach α) was .94 at the time of development, .95 in the study by Park and Kwon [30], and .94 in this study.

Realistic Immersion

Realistic immersion refers to the perception of presence in an environment that provides realistic illusion [30]. Realistic immersion was measured using the items corresponding to realistic immersion within the tool for measuring presence developed by Chung and Yang [31] for 3D video assessment. This study used 3 items: "The appearance of the VR simulation program video seemed to be real," "While learning the VR simulation program, the screen seemed to exist in reality," and "I felt like I was participating in a real field while learning the VR simulation program." Each item was rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). The highest possible score was 15 points, with higher scores indicating higher realistic immersion. The reliability of the tool (Cronbach α) was .75 in the study by Chung and Yang [31] and .76 in this study.

Satisfaction With the Program

Satisfaction with the VR simulation program was measured after providing the program to the participants using 3 items developed by Yu et al [4] and modified for this study, which were as follows: "This program will help me work as a nurse in clinical practice," "I want to recommend this program to other nursing students," and "This training is necessary as part of the nursing college curriculum." Each item was rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). The highest possible score was 15 points, with higher scores indicating higher satisfaction with the program. The reliability of the tool (Cronbach α) was .81 in the study by Yu et al [4] and .81 in this study.

VRICS Program

The VRICS program for patients with COVID-19 was constructed in the order of prebriefing, including prepractice, VR simulation, and debriefing (Table 1). This program was conducted with 6-7 people participating per session, and a total of 4 sessions operated.

Table 1. Program design of the VRICS[a] (overall time expended=180 min).

Procedure	Contents and situation	Time expended (min)
Prebriefing session		
Introduction of scenarios	Simulation scenarios, theory of respiratory care skills for patients with COVID-19	25
Proxy and prepractice	Watching the video, practicing donning and doffing PPE[b]	15
Orientation and precautions for VR[c]	Overview of the VR simulation lab environment, including the use of VR equipment, such as the HMD[d] and leap motion controller as well as disposable eye masks for the headset to prevent cross-contamination	10
VR simulation session		
Donning of PPE, respiratory care, doffing of PPE	Patient status check and nasal-oral suction care: • Patient: 1-year-old infant • Diagnosis: R/O COVID-19 • Symptom: Lung sound with crackle, heart rate 140 beats/min, oxygen saturation (SpO2) 95% • Respiration rate: 20 breaths/min	90-110
Debriefing session		
Discussion	Students reflecting on the simulation experience and exchanging feedback with the instructor	20

[a]VRICS: virtual reality infection control simulation.

[b]PPE: personal protective equipment.

[c]VR: virtual reality.

[d]HMD: head-mounted display.

Prebriefing

Prebriefing in this study consisted of standard prebriefing, along with demonstrations and actual practices. The participants received orientation on the goal of the simulation, introduction, review of scenarios, and the simulation lab environment. Subsequently, knowledge transfer learning regarding COVID-19 infection and suction care for preventing droplet transmission and oxygen therapy in a negative-pressure environment for inpatients with COVID-19 was introduced. The participants were instructed to watch an expert role-modeling video on donning and doffing PPE. Each participant practiced donning and doffing PPE for 15 min. During this process, the instructor corrected any errors made by the participants and encouraged the participants to perform the tasks accurately. Moreover, the participants were allowed to take photographs of each other while wearing PPE to experience success. This was to maximize the effects of VR simulation by carrying out learning using the actual experiences of the students. Users unfamiliar with VR technology may need additional time to become familiar with the controller [32]. Therefore, the instructor gave a personal demonstration for 10 min on VR-related equipment (head-mounted display [HMD] and leap motion controller), environment, and supplies in the VR simulation lesson room to enable imitation learning. The participants were allowed to gain proxy experience by watching a computer virtual screen and other people performing the tasks during the VR simulation lessons.

VR Simulation

The VR simulation consisted of 3 sessions: (1) donning and inspection of PPE in the dressing room before entering the negative-pressure isolation room (Figure 1), (2) nasopharyngeal suctioning and assessment for suction care and checking for pediatric patients with COVID-19 in the negative-pressure isolation room (Figure 2), and (3) doffing PPE in the contaminated area (Figure 3).

To operate the VR simulation program, an instructor's computer, the learner's HMD, and multimedia learning equipment were set up and a VR learning space with a radius of approximately 3 m from the learner was created so that learner movement and depth could be detected, as well as any danger during such movement. In addition, a user interface (UI) was created to induce movement to subsequent steps or to provide brief guidance on some procedural details. Each scenario was carried out for 15 min per participant. The program was designed to allow students who were not taking part in the VR simulation lesson to gain an indirect learning experience by watching the VR screen and other students taking part in the lesson. Students who completed the program were given 20 min to share and write about their opinions about the VR simulation, its strengths and weaknesses, areas of improvement, and when such a program should be operated if it is added to the curriculum. VR simulation took approximately 180 min to complete.

Figure 1. VR simulation session 1: dressing zone (clean area). Checking the PPE, handwashing (HW) > donning PPE (inner gloves > waterproof long-sleeved gown > shoes > N95 mask > goggles > hood > outer gloves) > checking the condition of the PPE by looking in the mirror. PPE: personal protective equipment; VR: virtual reality.

Figure 2. VR simulation session 2: negative-pressure isolation room. Checking patient identification, assessing patient condition > auscultating lung sound > respiratory care (oral and nasal suctioning). VR: virtual reality.

Figure 3. VR simulation session 3: changing zone (contaminated area). Doffing PPE (shoes > handwashing [HW] > outer gloves > HW > gown > HW > goggles > HW > N95 mask > HW > inner gloves > HW). PPE: personal protective equipment; VR: virtual reality.

Debriefing

Upon completion of the VR simulation, the participants and instructor shared their experiences about the VR simulation for approximately 20 min, discussing similarity to reality, immersion, usefulness, and satisfaction. Moreover, feedback about areas of improvement was exchanged, and the students were given time to describe their thoughts.

Data Collection Method

Data were collected from November 13 to December 10, 2021. Students participating in only the questionnaire survey (control group) and students participating in the VRICS program (experimental group) were recruited. Participants in the control and experimental groups had already completed the respiratory care for pediatric patients and the core basic nursing skill practice about donning and doffing standard PPE, not including level D PPE, and managing medical waste before this program. All participants completed an online preintervention survey. The control group completed an online postintervention survey 3 weeks after the preintervention survey. For preventing the diffusion of the experiment, the posttest for the control group was administered first. The experimental group completed a written postintervention survey after the completion of the program, and the completed questionnaire was placed in an envelope, which was retrieved by a research assistant. The survey required approximately 10-15 min to complete.

Data Analysis

Collected data were analyzed using SPSS Statistics version 25.0 (IBM Corp). The general characteristics of the participants were expressed as real numbers and percentages, while homogeneity of the general characteristics between the experimental and control groups was tested using a chi-square test, the Fischer exact test, and an independent *t* test. The normal distribution of PPE knowledge, infection control performance, and self-efficacy scores was tested using the Shapiro-Wilk test. Normality test results indicated that the experimental group showed a normal distribution, but the control group did not show a normal distribution for PPE knowledge scores, and thus, a homogeneity test was performed using the Mann-Whitney test. Both experimental and control groups showed a normal distribution for infection control performance and self-efficacy scores, and thus, a homogeneity test was performed using an independent *t* test.

Differences in PPE knowledge between the experimental and control groups were tested using the Mann-Whitney *U* test and the Wilcoxon matched-pair signed-rank test, while differences in infection control performance and self-efficacy between the experimental and control groups were tested using an independent *t* test and a paired *t* test.

Realism of the program and satisfaction with the program in the experimental group were analyzed by the mean and SD.

Ethical Considerations

After obtaining approval from the institutional review board of the institution affiliated with the researchers (GIRB-A21-Y-0061), an online preintervention survey was used to inform the participants about the objective of the study, their right to refuse participation during the study, and the no negative consequence on their school grade even if they did not participate in the study. Further, informed consent was obtained from the participants. Upon completion of the study, each participant was given a coffee gift certificate worth US $10-20.

Results

General Characteristics of the Participants and the Homogeneity Test

The general characteristics of the experimental and control groups and the PPE knowledge, infection control performance, and self-efficacy scores were homogeneous before the intervention (Table 2).

Table 2. Participants' characteristics and homogeneity of the 2 groups.

Characteristics	Total (N=50)	Control group (N=25)	Experimental group (N=25)	χ^2 (*df*)/*t* (*df*)/*z*	*P* value
College year, n (%)				0	.99
Third	26 (52.0)	13 (52.0)	13 (52.0)	N/A[a]	N/A
Fourth	24 (48.0)	12 (48.0)	12 (48.0)	N/A	N/A
Sex[b], n (%)				N/A	.35
Male	5 (10.0)	1 (4.0)	4 (16.0)	N/A	N/A
Female	45 (90.0)	24 (96.0)	21 (84.0)	N/A	N/A
Age (years), n (%)				2.03 (2)	.36
20-21	18 (36.0)	9 (36.0)	9 (36.0)	N/A	N/A
22-23	18 (36.0)	11 (44.0)	7 (28.0)	N/A	N/A
≥24	14 (28.0)	5 (20.0)	9 (36.0)	N/A	N/A
Age (years), mean (SD)	22.66 (2.08)	22.44 (2.02)	22.88 (2.15)	–0.75 (48)	.46
Previous semester grade[b], n (%)				2.17 (2)	.34
<3.0	5 (10.0)	2 (8.0)	3 (12.0)	N/A	N/A
3.0-3.9	31 (62.0)	18 (72.0)	13 (52.0)	N/A	N/A
≥4.0	14 (28.0)	5 (20.0)	9 (36.0)	N/A	N/A
Experience of VR[c], n (%)				.32 (1)	.57
Yes	24 (48.0)	11 (44.0)	13 (52.0)	N/A	N/A
No	26 (52.0)	14 (56.0)	12 (48.0)	N/A	N/A
PPE[d] knowledge, mean (SD)	14.86 (2.10)	14.52 (1.53)	15.20 (2.53)	–1.81	.07
Infection control performance, mean (SD)	3.40 (0.69)	3.28 (0.65)	3.51 (0.71)	–1.18 (48)	.24
Self-efficacy, mean (SD)	6.08 (0.64)	6.04 (0.60)	6.12 (0.68)	–0.42 (48)	.68

[a]N/A: not applicable.

[b]Fisher exact test.

[c]VR: virtual reality.

[d]PPE: personal protective equipment.

Testing of the Effects of the VRICS Program

The results of testing the effects of the VRICS program are shown in Table 3.

- Hypothesis 1 (difference in PPE knowledge): The experimental group showed a significant increase in the mean score from 14.52 points preintervention to 16.60 points postintervention (z=–3.85, P<.001). In contrast, the control group did not show a significant change in the mean score (z=–0.48, P=.63). The experimental group showed a significant increase in PPE knowledge compared to the control group as a result of the program (z=–3.28, P<.001). Thus, hypothesis 1 was accepted.
- Hypothesis 2 (difference in infection control performance): The experimental group showed a significant increase in the mean score from 3.28 points preintervention to 4.69 points postintervention (t_{24}=6.47, P<.001). In contrast, the control group did not show a significant change in the mean score (t_{24}=1.74, P=.095). The experimental group showed a significant increase in infection control performance compared to the control group as a result of the program (t_{48}=4.89, P<.001). Thus, hypothesis 2 was accepted.
- Hypothesis 3 (difference in self-efficacy): The experimental group showed a significant increase in the mean score from 6.04 points preintervention to 6.64 points postintervention (t_{24}=6.45, P<.001). Meanwhile, the control group showed a decrease in the mean score (t_{24}=–2.20, P=.04). The experimental group showed a significant increase in self-efficacy compared to the control group as a result of the program ($t_{36.2}$=4.93, P<.001). Thus, hypothesis 3 was accepted.

Table 3. Differences in variables between groups (N=50).

Group	Preintervention, mean (SD)	Postintervention, mean (SD)	Difference between time		Program effect		
			t (df)/z	*P* value	Postintervention-preintervention, mean (SD)	*t (df)/z*	*P* value
PPE[a] knowledge						–3.28[b]	<.001
Experimental	14.52 (1.53)	16.60 (1.22)	–3.85[c]	<.001	2.08 (1.75)	N/A[d]	N/A
Control	15.20 (2.53)	15.52 (2.00)	–0.48[c]	.63	0.32 (2.08)	N/A	N/A
Infection control performance						4.89 (48)	<.001
Experimental	3.28 (0.65)	4.69 (0.82)	6.47 (24)	<.001	1.41 (1.09)	N/A	N/A
Control	3.51 (0.71)	3.71 (0.83)	1.74 (24)	.095	0.20 (0.58)	N/A	N/A
Self-efficacy						4.93 (36.2)	<.001
Experimental	6.04 (0.60)	6.64 (0.52)	6.45 (24)	<.001	0.60 (0.47)	N/A	N/A
Control	6.12 (0.68)	5.72 (0.79)	–2.20 (24)	.04	-0.40 (0.90)	N/A	N/A

[a]PPE: personal protective equipment.

[b]Wilcoxon matched-pair signed-rank test.

[c]Mann-Whitney *U* test.

[d]N/A: not applicable.

Realistic Immersion and Learner Satisfaction

The mean score for realistic immersion in the experimental group in the VR simulation experience was 4.49 (SD 0.50) points (out of 5 possible points), and the mean learner satisfaction score was 4.75 (SD 0.38) points (out of 5 possible points); see Table 4.

Table 4. Realistic immersion and learner satisfaction of the experimental group (n=25).

Items	Response, n(%)					Mean (SD)
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	
Realistic immersion						
1. The appearance of the VR simulation program video seemed to be real.	N/A[a]	N/A	N/A	11 (44.0)	14 (56.0)	4.56 (0.51)
2. While learning the VR simulation program, the screen seemed to exist in reality.	N/A	N/A	1 (4.0)	11 (44.0)	13 (52.0)	4.48 (0.59)
3. I felt like I was participating in a real field while learning the VR simulation program.	N/A	N/A	3 (12.0)	8 (32.0)	14 (56.0)	4.44 (0.71)
Overall	N/A	N/A	N/A	N/A	N/A	4.49 (0.50)
Learner satisfaction						
1. This program will help me work as a nurse in clinical practice.	N/A	N/A	N/A	9 (36.0)	16 (64.0)	4.64 (0.49)
2. This training is necessary as part of the nursing college curriculum.	N/A	N/A	N/A	5 (20.0)	20 (80.0)	4.80 (0.41)
3. I want to recommend this program to other nursing students.	N/A	N/A	N/A	5 (20.0)	20 (80.0)	4.80 (0.41)
Overall	N/A	N/A	N/A	N/A	N/A	4.75 (0.38)

[a]N/A: not applicable.

Discussion

Principal Findings

In this quasi-experimental study, the cases of infection control for pediatric patients admitted to a negative-pressure isolation unit for COVID-19 infection were incorporated into a VR simulation program, and the effects of the program were assessed. Students who participated in the training program were third- and fourth-year nursing students with theoretical and clinical practice experience in pediatric nursing.

The findings of this study demonstrated that combining a VR simulation technique and skill training could enhance knowledge and skills for the preparation of respiratory infection care by future nurses. The experimental group, which participated in the VRICS program, showed significantly higher theoretical knowledge about PPE, infection control performance, and self-efficacy than the control group, which did not participate in the program. These findings suggest the need for VR simulation training before future nurses go to the front line to respond to respiratory infectious diseases. This VR simulation training was different than conventional training in that it provided simulated scenes of the isolation unit. Moreover, COVID-19 cases were incorporated into the training program. The order and accuracy of the donning and doffing of PPE and procedures related to respiratory care for pediatric patients were validated by infection control nurses and experts with at least 10 years of clinical experience in a neonatal intensive care unit (ICU) to simulate the actual environment related to COVID-19. Consequently, the participants were able to learn the procedure of donning PPE in the front zone and inspect the donned PPE before entering the isolation unit.

The participants were able to check the patient's name and assess the respiratory status using lung sounds and patient monitoring values for the safety of the patient inside the isolation unit. Subsequently, the participants were able to learn the procedure for suctioning oral and nasal excretions from the pediatric patient, while preventing aerosol transmission. This step included procedures for accurately checking the suction pressure and excretion patterns. The participants also performed procedures for discarding suction tubes and disposable gloves used for respiratory care into the isolation unit medical waste container. They also performed the procedure of exiting the isolation unit and moving to the doffing room before accurately doffing the shoes, gown, gloves, goggles, and mask and discarding them in the waste container inside the contaminated zone. This program simulated the procedures involved in the assessment of clinical symptoms and nursing intervention for patients with COVID-19, contact isolation, and disinfection. The use of PPE can markedly reduce the infection risk associated with caring for patients with COVID-19 [33,34]. Therefore, accurate training for donning and doffing PPE is important.

With no signs of the COVID-19 pandemic subsiding worldwide, training for and acquiring of essential infection control skills are important for not only health care professionals but also for nursing students in coping with infectious diseases under such difficult circumstances. In the immersive VR learning environment using VRICS from this study, the participants can learn theories and skills essential for infection control with the incorporation of respiratory care for pediatric patients with COVID-19. Moreover, the VRICS program can reduce the risk of infection and ensure the safety of the trainees by avoiding direct contact with patients with COVID-19. VR simulation also offers the advantage of allowing individualized learning by overcoming the time and space constraints faced by traditional simulation learning during the COVID-19 pandemic, when gathering large groups is difficult.

A study by Zhang et al [3] developed a VR simulation training program and tested the effects of the program on the response capabilities of emergency reserve nurses facing a public health crisis. The results showed that there was a significant increase in emergency nursing–related knowledge, care ability, and disaster preparedness scores among nurses who participated in the program (P<.01). These results support the findings of our study showing a significant increase in PPE knowledge and infection control performance in the experimental group. However, technical ability increased significantly in the control group, and there was no difference in the postdisaster management score, which was contradictory to the findings of this study. Meanwhile, a study by Yu et al [4] reported that nursing students who participated in VR simulation programs for high-risk neonatal infection control showed greater improvement in self-efficacy for high-risk neonatal infection control performance (t_{48}=–2.16, P=.02) than the control group that received only clinical practice training. These results support the findings of this study showing a significant difference in self-efficacy between the experimental and control groups.

Meanwhile, prebriefing was a crucial component of simulation learning [35,36], and the amount of information provided in prebriefing should be sufficient for the learner to begin problem solving [37]. It is believed that practicing donning and doffing PPE by VR simulation without individual practice or video-based learning would be difficult. Donning and doffing PPE is a complicated process that requires accuracy, and prior learning or practice is needed. Accordingly, this study used the concept of such prebriefing to carry out demonstrations and prior practice with donning and doffing PPE, along with an introduction to the scenarios. Previous studies have reported that improvement in satisfaction with simulation training [38], critical thinking disposition [39], immersion, and confidence [40] was achieved by prebriefing through watching a video that accurately presented the outcomes and goals. Moreover, when video-based prebriefing was applied, the self-efficacy of nursing students increased significantly in a team-based learning effect [41]. These common findings support the fact that video-based prebriefing further enhances the training effects of simulation compared to standard prebriefing that simply delivers information. According to Nayahangan et al [42], simulation-based training for medical staff has been mostly focused on clinical procedures related to practical skills. Ragazzoni et al [43] proposed a VR simulation model combining operational public health skills and hybrid skills training for infection control and Ebola treatment and management. Therefore, when using a VR simulation program, as in this study, diligently using prebriefing or applying a hybrid model could be a method for maximizing the learning effect.

According to Witmer and Singer [30], presence can be divided into 2 dimensions of involvement and immersion. In 3 dimensions, involvement can be further divided into spatial and temporal involvement and immersion can be further divided into dynamic and realistic immersion. Realistic immersion is defined as the perception of the environment that provides a realistic illusion. In this study, the experimental group showed a mean score of 4.49 (SD 0.50) points (out of 5 possible points) for the realistic immersion of VR simulation experience, which was relatively higher than the mean score of 2.75 (SD 0.91) points for realistic immersion reported by Chung [32], who developed the realistic immersion measurement tool. With respect to each item in the tool, “The appearance of the VR simulation program video seemed to be real” had the highest score (mean 4.56, SD 0.51 points), followed by “While learning the VR simulation program, the screen seemed to exist in reality” (mean 4.48, SD 0.59 points), and “I felt like I was participating in a real field while learning the VR simulation program” (mean 4.44, SD 0.71 points). The reason the realistic immersion for the screen seems to exist in reality, while learning and feel like participating in a real field showed relatively lower scores was because errors in sensor detection may have occurred due to differences in the heights of the participants or a change in the standing position may have caused the video to cut off as it was reloading. Moreover, immersion may have been reduced because the participants had to repeat the same motion several times due to poor recognition or had to return to the beginning due to the program being cut off. If the sensitivity of the sensor was increased and measures were taken to reduce program errors, then the participants may have been able to experience even high levels of realistic immersion.

The mean learner satisfaction score in this study was 4.75 (SD 0.38) points (out of 5 possible points), similar to 4.79 (SD 0.35) points reported in a study on high-risk neonatal VR simulation by Yu et al [4] and higher than 4.29 (SD 0.64) points reported in a study on general simulation training by Kim [44]. Among the items for learner satisfaction, “I want to recommend this program to other nursing students” and “This training is necessary as part of the nursing college curriculum” each showed a high mean score of 4.80 (SD 0.41) points (out of 5 possible points). It is believed that aspects of VR simulation that allowed the participants to use VR to indirectly experience the behavior in a neonatal ICU, where clinical training is difficult and can only be observed at the training site, and where practice can be repeated without the burden of committing mistakes, helped increase the level of learner satisfaction.

The participants showed a PPE knowledge score of 14-15 out of 20 points and an infection control performance score of 3 out of 5 points (average performance). Such results indicated that nursing students lacked PPE knowledge. Moreover, the control group, which did not participate in the VRICS program, showed a slight increase in yjr PPE knowledge score and no significant increase in infection control performance after 3 weeks. In the assessment of PPE knowledge using a questionnaire, the scores may have increased due to the participants remembering questions from the preintervention survey or being motivated to correct a wrong answer. No significant change in infection control performance in the control group indicates the need to provide programs or opportunities to teach skills that enhance COVID-19 infection control and allow the students to practice such skills.

Zellmer et al [45] reported that medical professionals did not remove their PPE properly during the Ebola virus outbreak, claiming that there is a need to improve the training that focuses on the protocol for properly wearing and removing PPE. Tabah et al [46] conducted an international study on PPE with 2711 ICU doctors and nurses during the COVID-19 pandemic. The results showed that N95 masks, full-sleeve waterproof gowns, and goggles were being washed or reused in 17%, 11%, and 34% cases, respectively, due to a shortage of PPE supplies. In fact, when conducting a prepractice in this study, level D PPE could not be individually provided to the participants due to the high cost and lack of supply during the COVID-19 pandemic. Accordingly, simulations such as VR offer an economically beneficial alternative to eliminating supply shortage. It is believed that in the future, simulation training can be integrated with a 3D virtual environment to optimize the training model so that the advantages of each feature can be maximized to enhance the overall clinical competency of nurses and nursing students.

Limitations

This study had a few limitations. First, because the study population consisted of nursing students from a single college, caution should be taken when interpreting the results. Each college has different training programs, and students may have different levels of knowledge, performance, and self-efficacy, even if they are in the same grade. Second, the procedures for VR-based skills need to be revised to be more realistic. The VR program used in this study could not reproduce delicate procedures or nursing skills. This is because detailed procedures could not be introduced to each process, since the video run time was set to 15 min to prevent dizziness from wearing the HMD for too long, and allocation of the total operating time for the VR program. Moreover, because learning the skills for donning and doffing of PPE and respiratory care was a bigger goal, the program was set up to complete the applicable skill procedure just by hand touching through hand tracking. Therefore, when developing future VR programs, it is necessary to improve the imaging technology to reproduce even delicate motion. If not, learning effects could also be maximized by including the process of actually donning and doffing PPE, as in this study. The third limitation was testing of skills by means of a self-questionnaire in this study. Therefore, it is suggested that actual performance be evaluated using evaluation mode in a VR program.

Conclusion

In this study, a VR simulation training program with the inclusion of infection control cases involving pediatric patients with COVID-19 admitted to an isolation unit was developed, and the effects of the program were tested. For infection control education, combining VR simulation and skills training was more effective than conventional instruction in producing better outcomes. This training program demonstrated more benefits for enhancing theoretical knowledge, respiratory care skills, and infectious disease preparedness, and thus, it could be applied to training nurses to respond even better to public health situations involving infectious diseases, including the COVID-19 pandemic.

Acknowledgments

MY contributed to the conception, design, and statistical analysis of this study; MRY and MY drafted the manuscript; and MY critically reviewed the manuscript and supervised the study process. All authors have read and approved the final manuscript.

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, and Technology (NRF-2018R1D1A3B07045408).

Conflicts of Interest

None declared.

References

1. Kim J, Yu M, Yu S, Park J. Information and general guidance for healthcare professionals in the fourth wave of COVID-19. J Korean Acad Nurs 2021 Aug;51(4):395-407. [doi: 10.4040/jkan.21137] [Medline: 34497249]
2. Jung H, Kim N, Lee Y, Kim M, Kim M. The effect of a disaster nursing convergence education program on disaster nursing knowledge, preparedness and self-confidence of nursing students. J Korea Convergence Society 2018;9(1):377-386. [doi: 10.15207/JKCS.2018.9.1.377]
3. Zhang D, Liao H, Jia Y, Yang W, He P, Wang D, et al. Effect of virtual reality simulation training on the response capability of public health emergency reserve nurses in China: a quasiexperimental study. BMJ Open 2021 Sep 22;11(9):e048611 [FREE Full text] [doi: 10.1136/bmjopen-2021-048611] [Medline: 34551944]
4. Yu M, Yang M, Ku B, Mann JS. Effects of virtual reality simulation program regarding high-risk neonatal infection control on nursing students. Asian Nurs Res (Korean Soc Nurs Sci) 2021 Aug;15(3):189-196 [FREE Full text] [doi: 10.1016/j.anr.2021.03.002] [Medline: 33894407]
5. Wolf AB. The impact of web-based video lectures on learning in nursing education: an integrative review. Nurs Educ Perspect 2018;39(6):E16-E20. [doi: 10.1097/01.NEP.0000000000000389] [Medline: 30335709]
6. Wang H, Kang K, Gao Y, Yang B, Li J, Wang L, et al. Remote nursing training model combined with proceduralization in the intensive care unit dealing with patients with COVID-19. World J Clin Cases 2021 Feb 16;9(5):999-1004 [FREE Full text] [doi: 10.12998/wjcc.v9.i5.999] [Medline: 33644163]

7. Hwang WJ, Kim JA. Development and evaluation of a home-visit simulation scenario for elderly people with diabetes mellitus who live alone. J Community Health Nurs 2020;37(2):89-102. [doi: 10.1080/07370016.2020.1736399] [Medline: 32233948]
8. Chang TP, Weiner D. Screen-based simulation and virtual reality for pediatric emergency medicine. Clin Pediatr Emerg Med 2016 Sep;17(3):224-230. [doi: 10.1016/j.cpem.2016.05.002]
9. Williams J, Jones D, Walker R. Consideration of using virtual reality for teaching neonatal resuscitation to midwifery students. Nurse Educ Pract 2018 Jul;31:126-129. [doi: 10.1016/j.nepr.2018.05.016] [Medline: 29879636]
10. Rourke S. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. Int J Nurs Stud 2020 Feb;102:103466. [doi: 10.1016/j.ijnurstu.2019.103466] [Medline: 31783192]
11. Kim S, Eom M, Park MH. Effects of nursing education using virtual reality: a systematic review. J Korea Contents Assoc 2019;19(2):661-670. [doi: 10.5392/JKCA.2019.19.02.661]
12. Verkuyl M, Romaniuk D, Atack L, Mastrilli P. Virtual gaming simulation for nursing education: an experiment. Clin Simul Nurs 2017 May;13(5):238-244. [doi: 10.1016/j.ecns.2017.02.004]
13. Birrenbach T, Zbinden J, Papagiannakis G, Exadaktylos AK, Müller M, Hautz WE, 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 Oct 08;9(4):e29586 [FREE Full text] [doi: 10.2196/29586] [Medline: 34623315]
14. Vidal VL, Ohaeri BM, John P, Helen D. Virtual reality and the traditional method for phlebotomy training among college of nursing students in Kuwait: implications for nursing education and practice. J Infus Nurs 2013;36(5):349-355. [doi: 10.1097/NAN.0b013e318243172f] [Medline: 24006114]
15. Smith PC, Hamilton BK. The effects of virtual reality simulation as a teaching strategy for skills preparation in nursing students. Clin Simul Nurs 2015 Jan;11(1):52-58. [doi: 10.1016/j.ecns.2014.10.001]
16. International Nursing Association for Clinical Simulation and Learning. INACSL Standards of Best Practice: SimulationSM Simulation Design. URL: https://www.nursingsimulation.org/article/S1876-1399(16)30126-8/fulltext#secsectitle0150 [accessed 2022-05-25]
17. Brennan BA. Prebriefing in healthcare simulation: a concept analysis. Clin Simul Nurs 2021 Jul;56:155-162. [doi: 10.1016/j.ecns.2021.04.015]
18. Bandura A. Self-efficacy: The Exercise of Control 1st Edition. New York, NY: W. H. Freeman; 1997.
19. Kardong-Edgren S, Butt A, Macy R, Harding S, Roberts CJ, McPherson S, et al. Expert modeling, expert/self-modeling versus lecture: a comparison of learning, retention, and transfer of rescue skills in health professions students. J Nurs Educ 2015 Apr;54(4):185-190. [doi: 10.3928/01484834-20150318-01] [Medline: 25826758]
20. McDermott DS. Prebriefing: a historical perspective and evolution of a model and strategy (know: do: teach). Clin Simul Nurs 2020 Dec;49:40-49. [doi: 10.1016/j.ecns.2020.05.005]
21. Brennan BA. The impact of self-efficacy based prebriefing on nursing student clinical competency and self-efficacy in simulation: an experimental study. Nurse Educ Today 2022 Feb;109:105260. [doi: 10.1016/j.nedt.2021.105260] [Medline: 34973554]
22. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
23. Eun Y, Bang S. Effects of the Lasater's clinical rubric of debriefing in advanced cardiovascular life support training. J Korea Contents Assoc 2016 Apr 28;16(4):516-527. [doi: 10.5392/jkca.2016.16.04.516]
24. Shin S, Park J, Kim J. Effectiveness of patient simulation in nursing education: meta-analysis. Nurse Educ Today 2015 Jan;35(1):176-182 [FREE Full text] [doi: 10.1016/j.nedt.2014.09.009] [Medline: 25459172]
25. Yu M, Kang H, Park J, Yang M. Development of a high-risk neonatal infection control competency scale. J Child Health Care 2021 Sep;25(3):393-411. [doi: 10.1177/1367493520943772] [Medline: 32741206]
26. Choi J. Nurse's Knowledge, Attitude and Use of Personal Protective Equipment Related to Acute Respiratory Infections (master's thesis). Seoul: University of Yonsei; 2016.
27. Kwon M. Knowledge, Perception, and Performance Related to Personal Protective Equipment among Nurses in Specialized Infectious Disease Hospitals (master's thesis). Pusan: Catholic University of Pusan; 2021.
28. Ayres H. Factors Related to Motivation to Learn and Motivation to Transfer Learning in a Nursing Population (dissertation). North Carolina: North Carolina State University; 2005.
29. Park SY, Kweon YR. The effect of using standardized patients in psychiatric nursing practical training for nursing college students. J Korean Acad Psychiatr Ment Health Nurs 2012;21(1):79-88. [doi: 10.12934/jkpmhn.2012.21.1.79]
30. Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence Teleop Virt Environ 1998 Jun;7(3):225-240. [doi: 10.1162/105474698565686]
31. Chung D, Yang H. Reliability and validity assessment in 3D video measurement. J Broadcast Eng 2012 Jan 30;17(1):49-59. [doi: 10.5909/jeb.2012.17.1.49]
32. Hassandra M, Galanis E, Hatzigeorgiadis A, Goudas M, Mouzakidis C, Karathanasi EM, et al. A virtual reality app for physical and cognitive training of older people with mild cognitive impairment: mixed methods feasibility study. JMIR Serious Games 2021 Mar 24;9(1):e24170 [FREE Full text] [doi: 10.2196/24170] [Medline: 33759797]

33. Alhazzani W, Møller MH, Arabi YM, Loeb M, Gong MN, Fan E, et al. Surviving Sepsis Campaign: guidelines on the management of critically ill adults with coronavirus disease 2019 (COVID-19). Intensive Care Med 2020 May;46(5):854-887 [FREE Full text] [doi: 10.1007/s00134-020-06022-5] [Medline: 32222812]
34. Cook TM. Personal protective equipment during the coronavirus disease (COVID) 2019 pandemic: a narrative review. Anaesthesia 2020 Jul;75(7):920-927 [FREE Full text] [doi: 10.1111/anae.15071] [Medline: 32246849]
35. Chamberlain J. Prebriefing in nursing simulation: a concept analysis using Rodger's methodology. Clin Simul Nurs 2015 Jul;11(7):318-322. [doi: 10.1016/j.ecns.2015.05.003]
36. Page-Cutrara K. Prebriefing in nursing simulation: a concept analysis. Clin Simul Nurs 2015 Jul;11(7):335-340. [doi: 10.1016/j.ecns.2015.05.001]
37. McDermott DS. The prebriefing concept: a Delphi study of CHSE experts. Clin Simul Nurs 2016 Jun;12(6):219-227. [doi: 10.1016/j.ecns.2016.02.001]
38. Kim MY, Park S. Effects of pre-briefing in simulation-based learning on nursing students' satisfaction with simulation. JKSSN 2018 Dec 31;6(2):27-34. [doi: 10.17333/jkssn.6.2.27]
39. Choi E. Effects on simulation education of pre-briefing using nursing process with video. JKSSN 2019 Jun 30;7(1):1-11. [doi: 10.17333/JKSSN.2019.7.1.1]
40. Park S, Kwak Y. Impact of pre-briefing using structured video in the pre-hospital trauma care simulation-based education. JKEIA 2020 Oct 31;14(7):421-427. [doi: 10.21184/jkeia.2020.10.14.7.421]
41. Kim H. The effects of pre-briefing team-based learning in standardized patients simulation. J Digit Converg 2020;18(8):271-279. [doi: 10.14400/JDC.2020.18.8.271]
42. Nayahangan LJ, Konge L, Russell L, Andersen S. Training and education of healthcare workers during viral epidemics: a systematic review. BMJ Open 2021 May 28;11(5):e044111 [FREE Full text] [doi: 10.1136/bmjopen-2020-044111] [Medline: 34049907]
43. Ragazzoni L, Ingrassia PL, Echeverri L, Maccapani F, Berryman L, Burkle FM, et al. Virtual reality simulation training for Ebola deployment. Disaster Med Public Health Prep 2015 Oct;9(5):543-546. [doi: 10.1017/dmp.2015.36] [Medline: 25782591]
44. Kim S. Effects of a simulation-based high-risk neonatal care education on learning satisfaction, class participation, learning motivation and clinical competency in nursing students. J Korea Acad-Ind Coop Soc 2015 Oct 31;16(10):6807-6815. [doi: 10.5762/kais.2015.16.10.6807]
45. Zellmer C, Van Hoof S, Safdar N. Variation in health care worker removal of personal protective equipment. Am J Infect Control 2015 Jul 01;43(7):750-751 [FREE Full text] [doi: 10.1016/j.ajic.2015.02.005] [Medline: 26138659]
46. Tabah A, Ramanan M, Laupland KB, Buetti N, Cortegiani A, Mellinghoff J, PPE-SAFE contributors. Personal protective equipment and intensive care unit healthcare worker safety in the COVID-19 era (PPE-SAFE): an international survey. J Crit Care 2020 Oct 13;59:70-75 [FREE Full text] [doi: 10.1016/j.jcrc.2020.06.005] [Medline: 32570052]

Abbreviations

HMD: head-mounted display
ICU: intensive care unit
PPE: personal protective equipment
VR: virtual reality
VRICS: virtual reality infection control simulation

Edited by N Zary; submitted 21.01.22; peer-reviewed by HJ Kang, T Sauter; comments to author 16.03.22; revised version received 26.03.22; accepted 13.05.22; published 27.05.22.

Please cite as:
Yu M, Yang MR
Effectiveness and Utility of Virtual Reality Infection Control Simulation for Children With COVID-19: Quasi-Experimental Study
JMIR Serious Games 2022;10(2):e36707
URL: https://games.jmir.org/2022/2/e36707
doi:10.2196/36707
PMID:35575818

Original Paper

Instructor Development Workshops for Advanced Life Support Training Courses Held in a Fully Virtual Space: Observational Study

Tetsuro Kiyozumi[1], MD, PhD; Norio Ishigami[1], MD, PhD; Daisuke Tatsushima[1], MD; Yoshiyuki Araki[1], MD, PhD; Yuya Yoshimura[2], MD, PhD; Daizoh Saitoh[3], MD, PhD

[1]Department of Defense Medicine, National Defense Medical College, Tokorozawa, Japan

[2]Department of Emergency and Critical Care Medicine, Hachinohe City Hospital, Hchinohe, Japan

[3]Division of Traumatology, Research Institute, Department of Traumatology and Critical Care, National Defense Medical College, Tokorozawa, Japan

Corresponding Author:
Tetsuro Kiyozumi, MD, PhD
Department of Defense Medicine
National Defense Medical College
Namiki 3-2
Tokorozawa, 3598513
Japan
Phone: 81 429951211
Email: kiyosan@ka2.so-net.ne.jp

Abstract

Background: Various face-to-face training opportunities have been lost due to the COVID-19 pandemic. Instructor development workshops for advanced resuscitation (ie, advanced life support) training courses are no exception. Virtual reality (VR) is an attractive strategy for remote training. However, to our knowledge, there are no reports of resuscitation instructor training programs being held in a virtual space.

Objective: This study aimed to investigate the learning effects of an instructor development workshop that was conducted in a virtual space.

Methods: In this observational study, we created a virtual workshop space by using NEUTRANS (Synamon Inc)—a commercial VR collaboration service. The instructor development workshop for the advanced life support training course was held in a virtual space (ie, termed *the VR course*) as a certified workshop by the Japanese Association of Acute Medicine. We asked 13 instructor candidates (students) who participated in the VR course to provide a workshop report (VR group). Reports from a previously held face-to-face workshop (ie, the face-to-face course and group) were likewise prepared for comparison. A total of 5 certified instructor trainers viewed and scored the reports on a 5-point Likert scale.

Results: All students completed the VR course without any problems and received certificates of completion. The scores for the VR group and the face-to-face group did not differ at the level of statistical significance (median 3.8, IQR 3.8-4.0 and median 4.2, IQR 3.9-4.2, respectively; P=.41).

Conclusions: We successfully conducted an instructor development workshop in a virtual space. The degree of learning in the virtual workshop was the same as that in the face-to-face workshop.

(JMIR Serious Games 2022;10(2):e38952) doi:10.2196/38952

KEYWORDS

virtual reality; virtual space; instructor development workshop; resuscitation training course; advanced life support; resuscitation training; digital training; virtual learning; digital education; medical education

Introduction

Background

Excellent instructors are indispensable for conducting advanced resuscitation (ie, advanced life support [ALS]) training courses, and there are existing instructor training programs for ALS training courses that are accredited by academic societies and other organizations. However, due to the COVID-19 pandemic, various face-to-face training opportunities have been lost [1-3], and instructor training programs are no exception. The use of remote meeting systems that use 2D screens is a common method of remote training. However, it is difficult to apply this methodology to a hands-on training course in which tools are used in a 3D space. Remote training using virtual reality (VR) is an attractive strategy for conducting experiential training without the need to gather people together [4-7]. Real-time, interactive interactions in a variety of situations that unfold in a virtual space can be expected to enhance the learning process [4] and improve communication skills, logical thinking, and decision-making skills [5]. The experiences in a fully immersive environment enhance learning, resulting in high knowledge retention and the development of empathy [6]. However, VR content still does not have enough tactile fidelity [1], skill acquisition via VR training may be inferior to skill acquisition via face-to-face training, and there are technical difficulties with conducting practical skills training by using advanced resuscitation equipment within a completely virtual space [8,9].

Importance

In terms of instructor training, there is no need to faithfully reproduce exact real-life skills in all courses. For example, conventional face-to-face instructor training programs have also been developed with an emphasis on teaching and facilitation techniques, often omitting the details of skills related to advanced resuscitation. Therefore, it is conceivable that the necessary learning effect could be obtained even when instructor training programs are held in a virtual space. However, few studies have evaluated the implementation of nontechnical skill building related to educational techniques in virtual spaces [10]. We did not find any reports of resuscitation instructor training programs that were held in a virtual space.

Goals of This Study

This study aimed to investigate the learning effects of an instructor development workshop for an immediate cardiac life support (ICLS) course—an ALS training course that was approved by the Japanese Association of Acute Medicine (JAAM) [11] and conducted in a virtual space.

Methods

Study Design and Setting

In this observational study, we created a virtual workshop space within NEUTRANS (Synamon Inc) [12]—a VR collaboration service. NEUTRANS is a commercial service that allows users to interact as avatars in a virtual space by using a head-mounted display. Users are able to walk around freely, grab objects, and otherwise navigate within the virtual space. This service also allows users to present presentation materials, view 360° videos, and use whiteboards and memos. It is possible to hold lectures, group discussions, practical training sessions, and role-plays in this virtual space.

In this study, we did not have expertise in programming or computer graphics, but we designed 3D models of dolls and resuscitation equipment by using the *paint* function in NEUTRANS, Oculus Medium (Adobe Inc), and the Windows 10 (Microsoft Corporation) onboard *paint* function. We placed these models in the virtual space for practical training sessions and role-plays. In addition, the simulated patient monitoring system—EmerSim (Penguin System Co, Ltd)—was reflected in the virtual space by using the *desktop sharing* function. This enabled the operation and display of the simulated patient monitor in virtual space (Figure 1).

We developed the program curriculum in accordance with the standard workshop program, which was accredited by the JAAM (Table 1), and held a trial workshop in March 2021. Instructor trainers and instructor candidates (students) entered the virtual space as avatars by using the Oculus Quest 2 (Meta Platforms Inc) or Oculus Lift (Meta Platforms Inc) and participated in the workshop. The status of the trial workshop was presented to the JAAM ICLS committee. After December 25, 2021, the workshop in the virtual space was accredited as an official course.

Figure 1. Scenery in the virtual space. We simulated an advanced resuscitation training course in a virtual space. The training mannequin and resuscitation equipment were created by using 3D models. The participants acted as avatars in the virtual space, and the patient monitor could be operated in a virtual section of the space.

Table 1. Program for the instructor development workshop.

Time	Agenda	Type	Notes
8:30 AM to 9 AM	Opening	Presentation and PT[a]	Included VR[b] equipment operation practice
9 AM to 9:40 AM	What is instruction?	SGD[c]	After discussing ideal instruction methodology, each group will present their ideas
9:40 AM to 10 AM	Education techniques	Presentation	Instructional procedures, nonverbal communication, feedback, facilitation, and coaching
10:15 AM to 10:45 AM	Explaining the use of supplies	RP[d]	As an instructor, describe the use of supplies (bag masks, scissors, etc) to 1 student
11 AM to noon	Teaching skills	RP	As an instructor, teach skills (ventilation, use of an AED[e], etc) to several students
1 PM to 1:30 PM	Setting up a simulation booth	PT	Trial an ideal arrangement of teaching materials (mannequins and resuscitation equipment)
1:45 PM to 2:45 PM	Performance evaluation and feedback	RP	Watching a 360° video of a resuscitation team in action, scoring performance by using a checklist, and providing feedback
3 PM to 4:30 PM	Leading a scenario exercise	RP	Act as an instructor to facilitate a scenario exercise until the time of posttraining feedback
4:30 PM to 5 PM	Summary	SGD and presentation	Discuss and present ideal instruction methods and award certificates of completion

[a]PT: practical training.

[b]VR: virtual reality.

[c]SGD: small group discussion.

[d]RP: role-play.

[e]AED: automated external defibrillator.

Intervention and Outcomes

We held the evaluated workshop in a virtual space (ie, the VR course) as a certified course in December 2021 and February 2022. Invited, voluntary participants were enrolled in this workshop. A total of 13 instructor candidates (students) who participated in the VR course (ie, the VR group) were asked to provide a workshop report. The reports from a face-to-face workshop (ie, the face-to-face course) that was held in June 2021 (ie, the face-to-face group) were prepared for comparative evaluations.

A total of 5 certified ICLS workshop instructor trainers viewed and scored each element of the reports on a 5-point Likert scale (poor=1; excellent=5) to evaluate the overall evaluation score. The instructor trainers also rated the reports on a 5-point Likert scale regarding whether they thought a report was a face-to-face course report or a VR course report (ie, the “VR-like score”). We analyzed the characteristics of the workshop reports by using text mining techniques.

After the VR course, we asked 13 students and 7 instructor trainers who participated in the VR course to rate their satisfaction with the course, the operability of the VR equipment, the occurrence of VR sickness, and whether each agenda item was judged to be suitable for implementation in VR. These items were included in a questionnaire, presented to the participants, and evaluated by using a 5-point Likert scale.

Statistical Analysis

We examined the scores’ interrater agreement by using the Kendall agreement coefficient. Following this, we used the average of the five raters' scores as the score for each student. We performed between-group comparisons of scores by using the Mann-Whitney *U* test.

We analyzed the questionnaire findings by using the Mann-Whitney *U* test, Friedman test, or Bonferroni multiple comparison test, as appropriate. The primary end point of this study was the overall evaluation score. *P* values of <.05 were considered statistically significant. All statistical analyses were conducted by using R statistical software (R Foundation for Statistical Computing) [13]. We performed a textual analysis by using KH Coder (Iknow LLC) [14,15].

Ethics Approval

This study was approved by the Ethics Committee of the National Defense Medical College (reference number: 4488) and was conducted with the consent of all participants. Moreover, we conducted this study in accordance with the principles of the Declaration of Helsinki and its later amendments.

Results

Study Participant Characteristics

Participants completed the VR course without any problems, and all students received certificates of completion. Participants in the December 2021 course logged in from 2 laboratories on the campus of the National Defense Medical College.

In the February 2022 course, participants logged in from 5 locations—3 laboratories on the campus of the National Defense Medical College, a meeting room in Hachinohe City Hospital (located approximately 600 km from the National Defense Medical College), and a private home. More specifically, the participant who signed in from a private home was affiliated with one of the laboratories and was identified as a person who had contact with a patient with COVID-19 (ie, 2 days prior to the course); thus, we placed this participant on home standby, which resulted in hastily arranged participation.

Main Results

We collected workshop reports from 13 students who were enrolled in the VR course (the VR group) and 10 students enrolled in the face-to-face course (the face-to-face group) to conduct comparative evaluations (report recovery rate: 23/23, 100%).

The students’ characteristics are listed in Table 2. With regard to evaluations of the workshop reports, the interrater agreement for the five certified instructor trainers was as follows: overall evaluation score=0.71 and VR-like score=0.77 for the Kendall agreement coefficient. The overall evaluation scores for the VR and face-to-face groups did not differ at the level of statistical significance (median 3.8, IQR 3.8-4.0 and median 4.2, IQR 3.9-4.2, respectively; *P*=.41; Figure 2). The median VR-like scores for the VR group and the face-to-face group were 0 (IQR −0.2 to 0.2) and 0.5 (IQR 0.4-1.1), respectively (Figure 3). We rated the face-to-face group as more “face-to-face–like,” and these findings were statistically significant (*P*=.03).

The results of the textual analysis are shown in Figure 4. The words *feel*, *I*, *participant*, and *think* characterized the VR group, and the words *you*, *points*, and *teach* characterized the face-to-face group.

We received questionnaires from 8 students and 6 instructor trainers who attended the VR course (report recovery rate: 14/20, 70%). The results are presented in Table 3. All respondents (14/14, 100%) were satisfied with the VR course, providing a score of 4 or higher on the 5-point Likert scale. Although some found the operation of the VR equipment to be confusing, this did not interfere with the progress of the course. Everyone participated until the end of the course, although some participants experienced VR sickness.

Table 2. Descriptive characteristics of the enrolled students.

Characteristics	Virtual reality group (n=13)	Face-to-face group (n=10)
Residents, n	1	1
Medical students, n	0	1
Nursing students, n	5	4
Nurses, n	5	3
Technicians (EMTs[a] and RTs[b]), n	2	1
Age (years), mean (SD)	29.1 (7.6)	26.1 (6.4)
Ratio of men to women	4:9	3:7

[a]EMT: emergency medical technician.

[b]RT: radiology technician.

Figure 2. Overall evaluation scores. The reports were scored on a 5-point Likert scale (1=poor; 5=excellent), and the averages from the evaluations conducted by the five instructor trainers were analyzed as the overall evaluation score. The black line indicates the median, the gray box indicates the range between the 25th and 75th percentiles, and the minimum and maximum values are indicated by lines or white circles. The overall evaluation scores for the VR and FTF groups did not differ at the level of statistical significance (median 3.8, IQR 3.8-4.0 and median 4.2, IQR 3.9-4.2, respectively). No statistically significant differences were observed within the Mann-Whitney U test results (*P*=.41). FTF: face-to-face; VR: virtual reality.

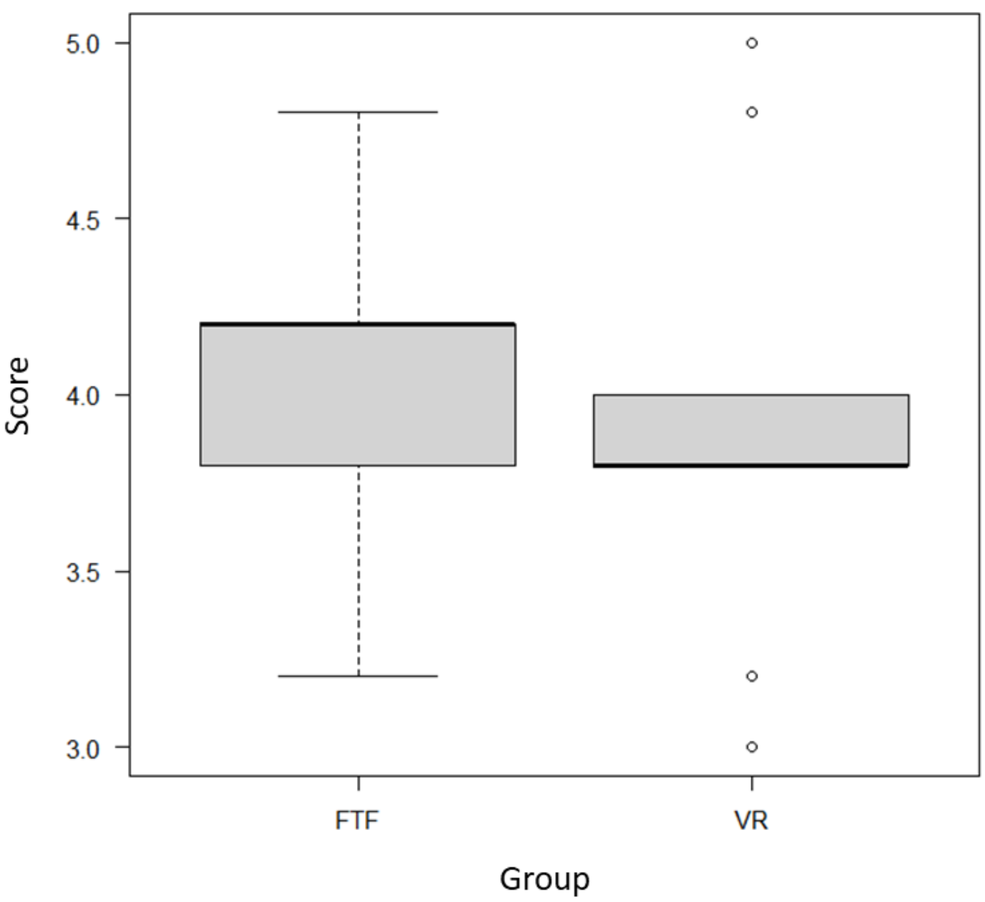


Figure 3. VR-like scores. A total of 5 instructor trainers rated the reports on a 5-point Likert scale regarding whether they thought a was an FTF course report or a VR course report (VR=–2; FTF=2), and the averages were analyzed as "VR-like scores." The black line indicates the median, the gray box indicates the range between the 25th and 75th percentiles, and the minimum and maximum values are indicated by lines or white circles. The median VR-like scores for the VR group and the FTF group were 0.0 (–0.2 to 0.2) and 0.5 (0.4-1.1), respectively. The results for the FTF group were rated as more "FTF-like" according to the findings of the Mann-Whitney U test (*P*=.03). FTF: face-to-face workshop; VR: virtual reality.

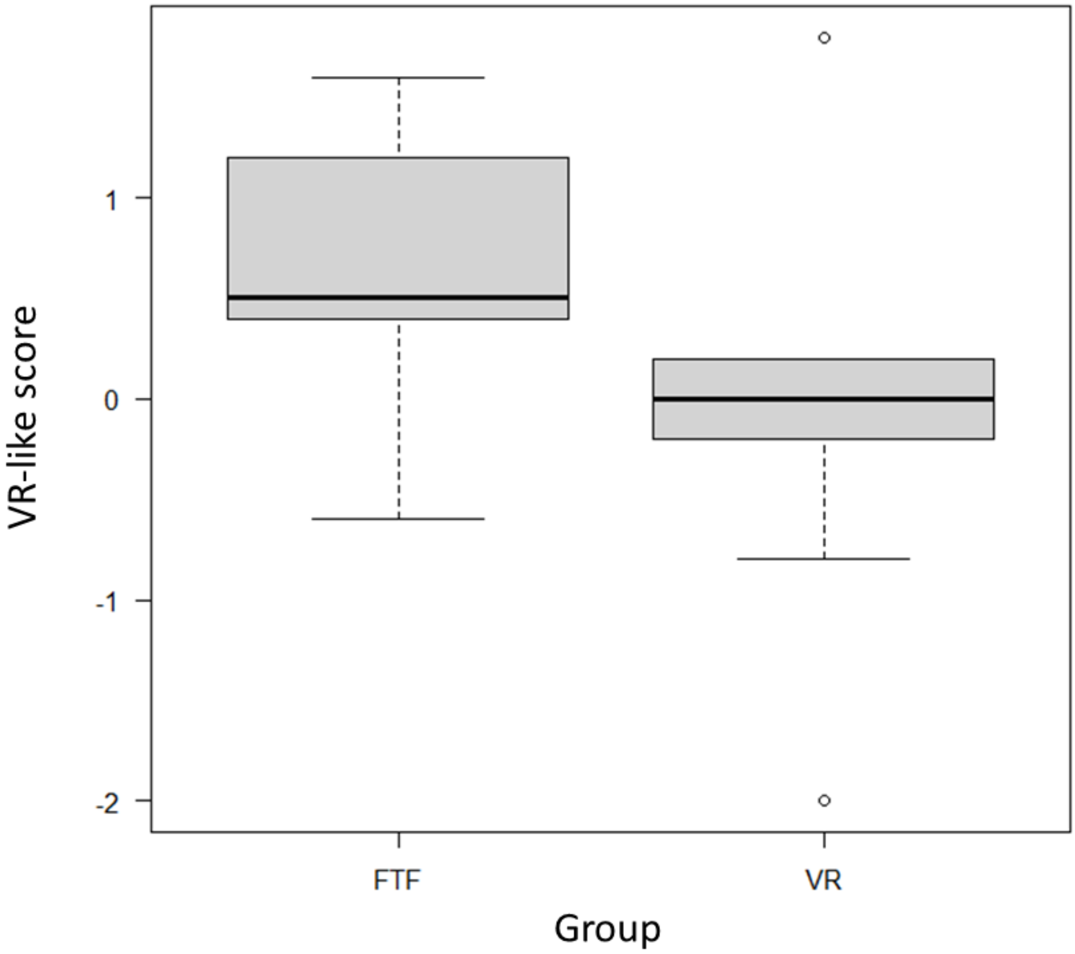


Figure 4. Textual analysis of the workshop reports. The "V" in the square represents the coordinates of the virtual reality (VR) group, and the "R" in the square represents the coordinates of the face-to-face (FTF) group. The size of the circles indicate the frequency of the words in the reports. The circles close to the reference point (0, 0) represent words that are not characteristic of the group, whereas the circles closer to the coordinates of each group represent words that are more characteristic of that group. The VR group was characterized by the words "feel," "I," "participant," and "think," while the FTF group was characterized by the words "you," "points," and "teach.".

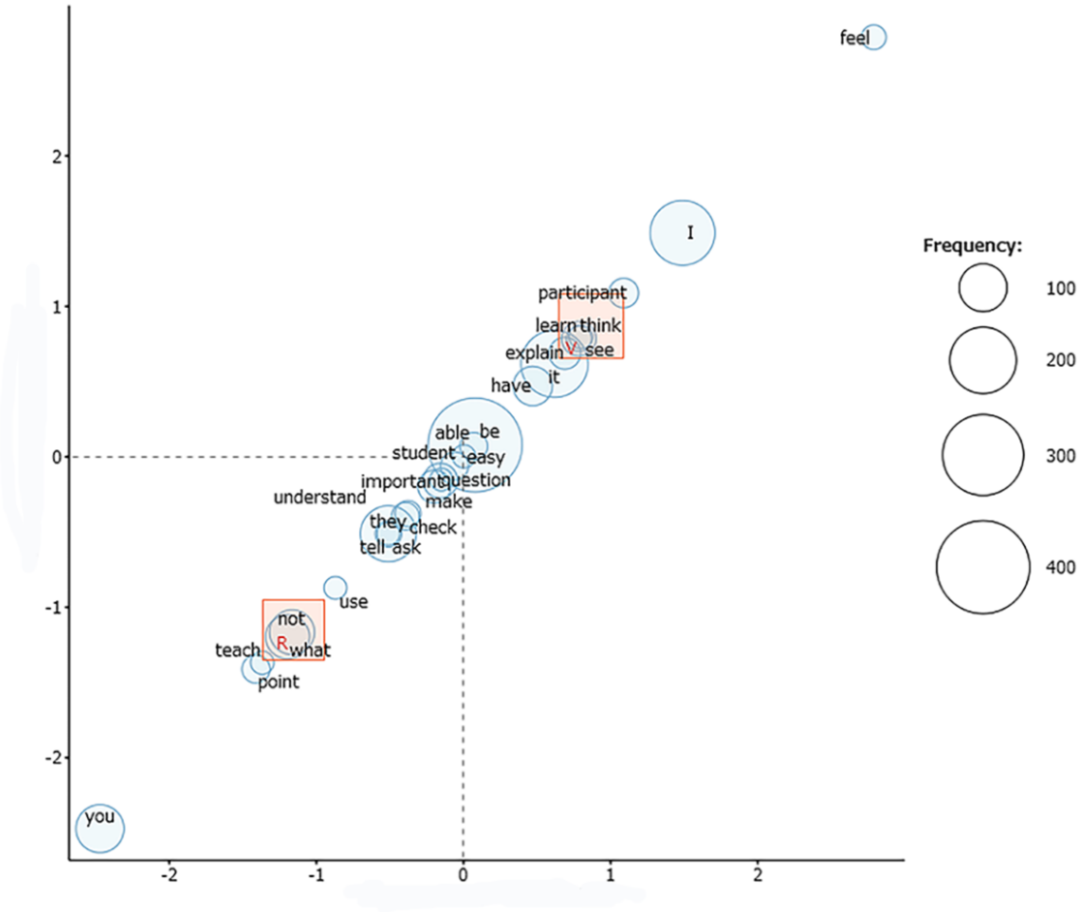


Table 3. Questionnaire results.

	Instructor trainers	Students
Sample size, n	6	8
Course satisfaction[a], score (median [IQR])	5 (4.25-5)	4.5 (4-5)
Operability of virtual reality equipment[a], score (median [IQR])	4 (4-4)	4 (3-4)
Virtual reality sickness[b], score (median [IQR])	5 (4.25-5)	5 (4.50-5)
Suitability for implementation in virtual reality[c], score (median [IQR])		
Discussion: what is instruction?	5 (3.5-5)	5 (3-5)
Presentation: education techniques	3 (3-4.5)	5 (3-5)
Role-play: explaining the use of supplies	5 (3.5-5)	5 (4.5-5)
Role-play: teaching specific skills	3 (3-4.5)	3 (1-3.5)
Practical training: setting up a simulation booth	3 (3-4.5)	3 (3-5)
Role-play: performance evaluation and feedback	5 (5-5)	5 (3-5)
Role-play: leading a scenario exercise	4 (3-5)	5 (1-5)

[a]Likert scale: bad=1; good=5.

[b]Likert scale: severe=1; none=5.

[c]Likert scale (suitable for implementation in virtual reality): no=1; yes=5.

There was a statistically significant difference (*P*=.03) between responses regarding the suitability of VR implementation, depending on the agenda of the workshop. Although we did not detect statistically significant differences in comparative evaluations between each agenda item, many respondents indicated that the teaching of specific discrete skills was not suitable for VR training, whereas the domains of performance evaluation and feedback were deemed suitable for VR training. There were no statistically significant differences in responses between the students and the instructor trainers (what is instruction: *P*=.94; education techniques: *P*=.33; explaining the use of supplies: *P*=.75; teaching specific skills: *P*=.54; setting up a simulation booth: *P*=.89; performance evaluation and feedback: *P*=0.12; leading a scenario exercise: *P*>.99).

Discussion

Principal Findings

To our knowledge, this study is the first to report on the development of instructor development workshops for ALS training courses that were conducted in a virtual space with the goal of measuring learning effectiveness. We performed this study by using a commercially available VR collaboration service in combination with several publicly available applications. This workshop could be held safely and remotely during the COVID-19 pandemic.

A total of 5 certified instructor trainers evaluated the workshop trainees' reports. Since a certain degree of agreement among the raters was observed, a simple average was used as the score. Workshop trainee report scores did not differ between the VR and face-to-face courses.

For the VR group, the median VR-like score was 0, demonstrating that the instructor trainers could not determine whether the VR group was in fact receiving face-to-face instruction or VR instruction. However, the instructor trainers rated the face-to-face group as more "face-to-face–like" than the VR group, and this finding was statistically significant (*P*=.03). This result may have been due to including words specific to the face-to-face course in the reports, such as words related to the performance of defibrillation and the use of a simulator. In other words, the VR course does not adequately reproduce the procedure; therefore, the evaluation may not have been conducted accurately. In the instructor development workshop, the subject matter was how to teach practical skills by using relatively simple skills that can be reproduced in a VR environment, and conducting instructor training in a virtual space was found to have the same learning effect as that of the face-to-face course. However, it is necessary to fully examine whether a VR environment can be used to teach more complex techniques in the future.

The textual analysis of the reports showed that the face-to-face group was characterized by words related to specific actions, while the VR group was characterized by words related to images and creative thinking. In a virtual space, a sense of immersion drives active engagement [16] and increases empathy [17]. It is possible that certain aspects, such as initiative, cohabitation, and imagination, also made a strong impression on the participants in this study. These and similar findings could lead to the discovery of new strengths within VR education.

Based on the questionnaire results of this study, we received many positive responses regarding VR training. However, we believe that many participants found the skills instruction component difficult to implement because the practical considerations for using equipment and tools in a virtual space differed from those in reality [1,8,9]. On the other hand, many participants considered performance evaluation and feedback to be well suited to VR because they were able to repeatedly view the 360° video in order to check evaluation points and provide feedback [18,19]. This finding is supported by the fact that all instructor trainers who were familiar with the conventional face-to-face course responded with a full score of 5.

There may be differences between an instructor development workshop that is conducted in a virtual space and a face-to-face workshop, such as differences in the reproducibility of practical skills [1,8,9]. However, through this study, we were able to confirm that it is possible to hold an entire instructor development workshop for an ALS training course in a virtual space while maintaining the intended educational effect. Since the competencies to be acquired in the instructor development workshop are not practical skills but are ways of teaching practical skills, we believe that learning effectiveness and satisfaction can be enhanced by simple role-plays. The practical skills that can be reproduced in a VR environment and via avatars, such as chest compression and safe electroshock administration, are appropriate as role-play subjects.

As we held the course in an entirely virtual space, participants did not need to gather in person at all. This is an advantage not only in special situations such as the COVID-19 pandemic but also in encouraging the participation of a wide range of students and instructor trainers who have traditionally had difficulties with attending courses due to geographical constraints and other limitations. Thus, workshops that are held in virtual spaces will undoubtedly contribute to the development and recruitment of high-quality instructors.

Since there are no physical constraints in a VR space, there is no need to consider the limitations on the number of participants when securing a venue, which is an issue in face-to-face training. However, in order for a large number of participants to gather in a VR space, each participant must prepare VR equipment, and it is necessary to consider the initial investment in the necessary equipment and acceptable costs. Various VR equipment and applications exist, but they often lack functions for mutual communication and compatibility. A common platform that allows for interactions in a VR space through the use of multiple pieces of equipment and applications may be a solution to this problem.

Students who enroll in a VR course may be positively predisposed toward VR when they are able to indicate their participation after knowing in advance that they are enrolling in a VR course. However, due to the COVID-19 pandemic, a workshop that we offered was rescheduled as a VR course, and the previously enrolled participants were only made aware of this after receiving a call from our administration. Moreover, there was no choice provided to the prospective face-to-face workshop participants. Therefore, bias was considered minimal in this regard.

Conclusion

This study evaluated instructor development workshops for ALS training courses that were held in a fully virtual space. The learning effect of the VR workshop was the same as that of the face-to-face training workshop. Thus, our findings encourage the implementation of virtual training courses for the effective teaching of ALS skills and may be generalizable to other VR contexts, such as VR medical education, VR training during the COVID-19 pandemic, and other settings that may be amenable to VR-based instruction.

Acknowledgments

We would like to thank Ms Azumi Takahashi for her assistance with regard to data input. This research study was funded by a grant from the Defense Medical Advanced Research Fund (Japan). This funder had no role in the design, conduct, or reporting of this research.

Conflicts of Interest

None declared.

References

1. Kassutto SM, Baston C, Clancy C. Virtual, augmented, and alternate reality in medical education: Socially distanced but fully immersed. ATS Sch 2021 Oct 18;2(4):651-664 [FREE Full text] [doi: 10.34197/ats-scholar.2021-0002RE] [Medline: 35079743]
2. Jiang H, Vimalesvaran S, Wang JK, Lim KB, Mogali SR, Car LT. Virtual reality in medical students' education: Scoping review. JMIR Med Educ 2022 Feb 02;8(1):e34860 [FREE Full text] [doi: 10.2196/34860] [Medline: 35107421]
3. Haroon Z, Azad AA, Sharif M, Aslam A, Arshad K, Rafiq S. COVID-19 era: Challenges and solutions in dental education. J Coll Physicians Surg Pak 2020 Oct;30(10):129-131. [doi: 10.29271/jcpsp.2020.supp2.129] [Medline: 33115587]
4. Yin W. An artificial intelligent virtual reality interactive model for distance education. Journal of Mathematics 2022 Feb 14;2022:1-7 [FREE Full text] [doi: 10.1155/2022/7099963]
5. 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]
6. Peisachovich E, Appel L, Sinclair D, Luchnikov V, Da Silva C. CVRriculum program faculty development workshop: Outcomes and suggestions for improving the way we guide instructors to embed virtual reality into course curriculum. Cureus 2021 Mar 04;13(3):e13692 [FREE Full text] [doi: 10.7759/cureus.13692] [Medline: 33824834]
7. Creutzfeldt J, Hedman L, Medin C, Heinrichs WL, Felländer-Tsai L. Exploring virtual worlds for scenario-based repeated team training of cardiopulmonary resuscitation in medical students. J Med Internet Res 2010 Sep 03;12(3):e38 [FREE Full text] [doi: 10.2196/jmir.1426] [Medline: 20813717]
8. Nas J, Thannhauser J, Vart P, van Geuns RJ, Muijsers HEC, Mol JQ, et al. Effect of face-to-face vs virtual reality training on cardiopulmonary resuscitation quality: A randomized clinical trial. JAMA Cardiol 2020 Mar 01;5(3):328-335 [FREE Full text] [doi: 10.1001/jamacardio.2019.4992] [Medline: 31734702]
9. 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 (Lond) 2022 Jan 05;73:103241 [FREE Full text] [doi: 10.1016/j.amsu.2022.103241] [Medline: 35079374]
10. Bracq MS, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: A systematic review. Simul Healthc 2019 Jun;14(3):188-194. [doi: 10.1097/SIH.0000000000000347] [Medline: 30601464]
11. ICLS. Immediate Cardiac Life Support. URL: http://www.icls-web.com/index.html [accessed 2022-06-17]
12. NEUTRANS. Synamon Inc. URL: https://neutrans.space/ [accessed 2022-06-17]
13. Ihaka R, Gentleman R. R: A language for data analysis and graphics. J Comput Graph Stat 1996;5(3):299-314. [doi: 10.1080/10618600.1996.10474713]
14. Higuchi K. A two-step approach to quantitative content analysis: KH Coder tutorial using Anne of Green Gables (Part I). Ritsumeikan Social Science Review 2016;52(3):77-91 [FREE Full text]
15. Higuchi K. A two-step approach to quantitative content analysis: KH Coder tutorial Using Anne of Green Gables (Part II). Ritsumeikan Social Science Review 2017;53(1):137-147 [FREE Full text]
16. 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 Jan 14;15(1):e9 [FREE Full text] [doi: 10.2196/jmir.1715] [Medline: 23318253]
17. Guetterman TC, Sakakibara R, Baireddy S, Kron FW, Scerbo MW, Cleary JF, et al. Medical students' experiences and outcomes using a virtual human simulation to improve communication skills: Mixed methods study. J Med Internet Res 2019 Nov 27;21(11):e15459 [FREE Full text] [doi: 10.2196/15459] [Medline: 31774400]

18. Snelson C, Hsu YC. Educational 360-degree videos in virtual reality: a scoping review of the emerging research. TechTrends 2019 Dec 20;64:404-412. [doi: 10.1007/s11528-019-00474-3]
19. Barbot B, Kaufman JC. What makes immersive virtual reality the ultimate empathy machine? Discerning the underlying mechanisms of change. Comput Human Behav 2020 Oct;111:106431. [doi: 10.1016/j.chb.2020.106431]

Abbreviations

ALS: advanced life support
ICLS: immediate cardiac life support
JAAM: Japanese Association of Acute Medicine
VR: virtual reality

Edited by N Zary; submitted 25.04.22; peer-reviewed by N Moore, J Afonso; comments to author 12.05.22; revised version received 18.05.22; accepted 12.06.22; published 29.06.22.

Please cite as:
Kiyozumi T, Ishigami N, Tatsushima D, Araki Y, Yoshimura Y, Saitoh D
Instructor Development Workshops for Advanced Life Support Training Courses Held in a Fully Virtual Space: Observational Study
JMIR Serious Games 2022;10(2):e38952
URL: https://games.jmir.org/2022/2/e38952
doi:10.2196/38952
PMID:35767318

Original Paper

A Photography-based, Social Media Walking Intervention Targeting Autonomous Motivations for Physical Activity: Semistructured Interviews With Older Women

Michael C Robertson[1], PhD, MPH; Maria Chang Swartz[2], PhD, MPH, RD, LD; Ursela Christopherson[1], PhD, MSOT, OTR/L; Jason R Bentley[1,3], PhD, CSCS; Karen M Basen-Engquist[4], PhD, MPH; Debbe Thompson[5], PhD, RD; Elena Volpi[6], MD, PhD; Elizabeth J Lyons[1], PhD, MPH

[1]Department of Nutrition, Metabolism & Rehabilitation Sciences, School of Health Professions, University of Texas Medical Branch at Galveston, Galveston, TX, United States

[2]Department of Pediatrics-Research, University of Texas MD Anderson Cancer Center, Houston, TX, United States

[3]Department of Clinical Health and Applied Sciences, University of Houston – Clear Lake, Houston, TX, United States

[4]Department of Behavioral Science, University of Texas MD Anderson Cancer Center, Houston, TX, United States

[5]US Department of Agriculture/Agricultural Research Service Children’s Nutrition Research Center, Baylor College of Medicine, Houston, TX, United States

[6]Sealy Center on Aging, University of Texas Medical Branch, Galveston, TX, United States

Corresponding Author:
Elizabeth J Lyons, PhD, MPH
Department of Nutrition, Metabolism & Rehabilitation Sciences
School of Health Professions
University of Texas Medical Branch at Galveston
301 University Blvd
Galveston, TX, 77555
United States
Phone: 1 409 772 2575
Fax: 1 409 772 2577
Email: ellyons@utmb.edu

Abstract

Background: Older adult women are at risk for negative health outcomes that engaging in sustained physical activity can help prevent. However, promoting long-term maintenance of physical activity in this population has proven to be a challenge. Increasing autonomous motivations (ie, intrinsic, integrated, and identified regulations) for physical activity may facilitate enduring behavior change. Digitally delivered games for health that take a celebratory technology approach, that is, using technology to create new ways to experience valued behaviors and express valued beliefs, may be a useful way to target autonomous motivations for physical activity. Formative research with the target population is needed to design compelling intervention content.

Objective: The objective of this study is to investigate older adult women’s reactions to and thoughts about a photography-based, social media walking game targeting autonomous motivations for physical activity.

Methods: During an individual semistructured interview, a moderator solicited feedback from 20 older adult women (age range 65-74 years) as part of formative research to develop a social media game featuring weekly walking challenges. The challenges were designed to target autonomous motivations for physical activity. Interviews were audio-recorded and transcribed verbatim. Two reviewers conducted thematic content analysis on interview transcripts.

Results: We identified 3 overarching themes in qualitative data analysis. These reflected the playful experiences, value, and acceptability associated with the intervention challenges. Generally, participants understood what the challenges were asking them to do, proffered appropriate example responses, and indicated that the challenges would be enjoyable. Participants reported that the intervention content afforded many and varied playful experiences (eg, competition, discovery, exploration, expression, fellowship, humor, nurture, sensation). Further, participants indicated that the intervention increased their motivation for physical activity, occasioned meaningful shifts in perspective, increased their knowledge of various topics of interest, provided an opportunity to create valued connection with others, and provided health-related benefits. Participants suggested the intervention emphasize local history, nature, and cultural events.

Conclusions: The photography-based, social media walking game with relatively simple game mechanics was well received and judged to be apt to bring about a wide variety of emotive experiences. A clear, geographically specific identity emerged as a key driver of interest for intervention content. Taking a celebratory technology approach holds promise for targeting autonomous motivations for physical activity in older adult women.

(JMIR Serious Games 2022;10(2):e35511) doi:10.2196/35511

KEYWORDS

physical activity; walking; exercise; fitness; social media; health; intervention; behavior; behavior mechanism; psychological theory; serious games; gamification; older women; older adults; behavior change; behavioral interventions; mobile phone; photography; patient perspective; patient attitude

Introduction

Older adult women are at risk for negative health outcomes, including chronic health conditions, cognitive decline, mental health challenges, fall-related injuries, decreased physical functioning, and decreased quality of life [1-8]. Physical activity may help protect against these negative health outcomes and is recommended for older adults by authoritative sources [9-12]. However, most older adult women do not engage in recommended levels of physical activity [13,14].

Behavioral interventions to increase physical activity in older adult women have been relatively successful in promoting physical activity in the short term [15]. However, to maximize the health-protective benefits of physical activity, long-term physical activity maintenance is necessary. Consistently achieving long-term behavior change in this population has proven to be an elusive target [16-18]. This may be attributable in part to the theoretical underpinnings of most physical activity interventions. Many interventions have been based on social cognitive theory or notions of self-regulation and have placed great emphasis on the importance of skill building [18]. Although this approach has proven useful for promoting behavior change *initiation*, to achieve changes in long-term *maintenance* of physical activity, behavioral interventions may need to additionally target disparate processes.

The Maintain Identify Transformation (IT) model posits that *centered identity transformation*, the process of integrating a behavior with one's internal self-representations, may facilitate the activation of efficient cognitive processing and, in turn, facilitate enduring behavioral changes [19]. Conversely, it holds that behavior change approaches centered on executive functioning, including those based primarily on social cognitive theory and promoting active self-regulation, are unlikely to lead to long-term maintenance of behavior change. This is in part because executive functioning is relatively slow, error prone, and impaired by adverse physiological and psychological states [19]. Targeting motivational constructs may be 1 way to help elicit centered identity transformation and the activation of some of the more efficient cognitive processes needed to support robust, sustained changes in health-related behavior [19].

Self-determination theory (SDT) provides a nuanced understanding of motivation for physical activity [20,21]. It discriminates intrinsic motivation (ie, finding the behavior to be inherently enjoyable) from extrinsic motivation (ie, engaging in a behavior as a means to an end). Extrinsic motivations are further parsed conceptually according to the degree to which the motivation is external (eg, rewards/punishment) versus internal (eg, valuing the behavior's outcomes or identifying as a person who engages in the behavior) in nature. Increasing one's autonomous motivations, which include both intrinsic motivations as well internal forms of extrinsic motivation (ie, *integrated* and *identified* regulations, which reflect valuing and identifying with the behavior and its consequences), may facilitate centered identity transformation [19]. SDT suggests that integrated and identified regulations for health-related behaviors are modifiable constructs that may increase as one's core psychological needs are satisfied (ie, *autonomy*, *relatedness*, and *competence*).

Digitally delivered interventions hold promise for physical activity promotion in older adults [22]. Most digitally delivered interventions for physical activity take a *corrective approach* to health promotion—they are typically centered on redressing problematic aspects of one's behavior. Such interventions tend to use technology to promote self-regulation and increase external forms of motivation (eg, the provision of wearable devices to provide feedback, data visualization schemes to foster comparison of behavioral performance with goals, and the provision of reward for performance) [23]. Taking a *celebratory technology approach* instead, that is, using technology to create new ways to experience valued behaviors and express valued beliefs, may help frame personal informatics more toward promoting reflection [24], context [25], and storytelling [26]. This may be more conducive to increasing autonomous motivation for physical activity, facilitate centered identity transformation, and lead to more robust, long-lasting behavior change.

Games for health, which use game design elements to achieve nongame objectives, can be used to target SDT constructs in a celebratory manner [27-29]. Games for health oriented toward health promotion emphasize the importance of pairing enjoyable experiences with the promotion of health-related behaviors [30]. They are a natural fit for taking a celebratory technology approach (vs a corrective approach). Rather than attempting to make it *easier* to be physically active, this approach to physical activity promotion may be centered on making it more *interesting* to be active. This represents a noteworthy departure from contemporary health promotion frameworks, and formative research with priority populations is needed to design compelling game content. We previously used the Mechanics, Experiences, Change (MECHA) model, a step-by-step program-planning model for designing games for health, to develop Challenges for Healthy Aging: Leveraging Limits for Engaging Networked Game-based Exercise (CHALLENGE) [31]. CHALLENGE is an autonomy-focused game for health that implements weekly walking challenges using social media. Despite systematic program design, making physical activity more interesting is an inherently challenging endeavor, and collaboration with the target population is needed [32]. It is likely that individual preferences and interests contribute to what people find interesting, so in-depth formative work is critical to the success of such a project. Thus, the purpose of this study is to interview older adult women to investigate the acceptability of this unique approach to physical activity promotion and how we might improve upon CHALLENGE.

Methods

Study Design and Recruitment

This study was a qualitative, formative study conducted prior to a randomized controlled trial. We recruited 20 participants to expand and refine CHALLENGE (presented later). Eligibility criteria included that participants be female adults between 65 and 85 years of age who were able to read and understand English. We excluded the "oldest old" population due to differences between this population and younger-old populations who tend to have fewer functional limitations [33] and because this younger-old population has demonstrated general acceptance and enjoyment of mobile and wearable devices [34,35], games [36,37], and mobile health (mHealth) interventions [38,39]. During the COVID-19 pandemic, we altered study protocols to allow for online interviews. These changes further necessitated that participants have access to teleconference capability from home (eg, Skype, FaceTime, or Zoom). We recruited participants via in-person recruitment at University of Texas Medical Branch (UTMB) gerontology and primary care clinics, online recruitment using the UTMB's newsletters, and flyers and brochures available at locations frequented by older adults. Originally, 20 participants were recruited for this study. However, 1 (5%) participant's audio data failed during her videoconference and thus could not be used. Due to this issue, the research team recruited an additional participant to reach the target sample size.

Ethics Approval

This project was approved by the UTMB at Galveston's Institutional Review Board (protocol no. 19-0158). Verbal informed consent was obtained from all participants prior to participation in this study.

Intervention

We developed CHALLENGE using the MECHA model [31]. The details of this game for health physical activity intervention are presented elsewhere [31]. Briefly, CHALLENGE featured weekly walking challenges to be delivered via a private Facebook group. Rather than promoting performance-based achievements (eg, obtaining a particular step count), the weekly challenges were designed to elicit playful experiences (PLEXs) and target SDT psychological needs (ie, autonomy, competence, and relatedness) and autonomous motivations (ie, intrinsic, integrated, and identified regulations) to encourage walking behaviors [40,41]. For example, 1 weekly challenge invited participants to present a photograph of an object that might serve as a symbol of their community.

In-Depth Interviews

From January to December 2020, we conducted an in-depth, semistructured interview with each participant to investigate perceptions of the feasibility and acceptability of the weekly walking challenges. We conducted interviews in 2 rounds, which were separated by study staff revising CHALLENGE content based on participant feedback. The first round of interviews was conducted in person. Due to closures to research in March 2020 associated with the COVID-19 pandemic, the second round of interviews was conducted via videoconference. A trained research assistant took a modified cognitive interview approach, in which she presented all challenges to the participants and asked them to read each challenge aloud and talk through their impressions of the challenge. An example challenge is

> *Party Crasher. Find an animal party. Go on a walk and look for a group of birds, squirrels, cat, dogs, or any animal socializing around your neighborhood. Snap a photo and share it. Crash the biggest animal party you can find!*

The research assistant responded with probing questions concerning (1) what the participant thought the challenge writer meant, (2) what they believed the challenge writer wanted them to do, (3) what photos they might take to fulfill the challenge, (4) what was the likelihood that they would accept the challenge, (5) what was the likelihood they would enjoy the challenge, and (6) what was the likelihood the challenge would lead to discussion. Interviews took about 2 hours to complete. Participants were compensated with a US $25 gift card. All interviews were digitally recorded and professionally transcribed verbatim.

Qualitative Data Analysis

We imported interview transcripts from both rounds into the qualitative data analysis management program Nvivo version 12 (QSR International Pty Ltd). Two independent reviewers (authors MCR and EJL) conducted thematic content analysis on all interview transcripts using a hybrid approach to assign deductive codes and inductive codes to discrete points made by study participants [42]. The coders started with deductive codes, which reflected the conceptual model of the intervention. These codes reflected the PLEXs targeted by CHALLENGE and were derived from the PLEX framework (eg, *exploration*; Table 1) [40,41]. Inductive codes were used to capture other important or related information in the transcripts that were judged to be important in relation to the research questions (eg, *real-life benefits* [of the intervention] or *impacts of where you live* [on completing challenges]). After coding was complete, the inductive codes were examined and grouped into higher-order categories or themes (eg, impacts of geographic location, suggestions for improvement). The coders maintained a codebook and organized codes into recurring themes and subthemes. They regularly met to reach consensus regarding discrepancies in coding via discussion, identify recurring themes and subthemes, establish the overarching structure of themes and subthemes, and identify illustrative quotations.

Table 1. Definition of PLEX[a] targeted by challenges.

PLEX	Definition [41]
Captivation	Forgetting one's surroundings and the passing of time, absorption
Challenge	Testing abilities in a demanding task, developing and testing skills, learning new things
Competition	Contesting with oneself or an opponent, trying to achieve a defined goal
Completion	Finishing a major task, experiencing closure related to a tension, mastery experience, finishing a level or a game
Control	Dominating, commanding, and regulating; a feeling of being powerful
Cruelty	Causing mental or physical pain, feeling malicious or manipulative, bullying or destruction
Discovery	Finding something new or unknown, uncovering something previously hidden, surprise
Exploration	Investigating an object or situation, curiosity and a thirst for knowledge
Expression	Manifesting oneself creatively; designing, constructing, personalizing
Fellowship	Friendship, communality, or intimacy; sharing experiences with others
Humor	Fun, joy, amusement, jokes, gags
Nurture	Taking care of oneself or others, tutoring and helping others in the game
Relaxation	Relief from bodily or mental work, "unwinding" while playing the game
Sensation	Excitement by stimulating senses, pleasure related to senses
Subversion	Breaking social rules and norms or seeing others do so, twisting meaning
Suffering	Experience of loss, frustration, and anger during play
Thrill	Excitement derived from risk and danger, suspense and excitement

[a]PLEX: playful experience.

Results

Participant Characteristics

Of the 20 participants constituting the analytical sample, 15 (75%) identified as non-Hispanic White, 3 (15%) identified as Hispanic, and 2 (10%) identified as Black. The average age was 70.1 years (SD 2.6, range 65-74).

Thematic Content Analysis

Two rounds of data collection were performed. The first round of interviews included 9 (45%) of 20 participants. Although we originally planned to update the challenges after the 10th participant, closures to research during March 2020 led to a natural stopping point after the 9th participant. We revised the challenges after the first 9 interviews based on participant feedback. We then proceeded to conduct interviews with 11 (55%) additional participants and the revised challenges. The average age and racial/ethnic makeup of the participants were nearly identical in rounds 1 and 2. We identified 3 overarching themes in our qualitative data analysis. These were (1) PLEXs, (2) value derived from the challenges, and (3) acceptability of the challenges.

Playful Experiences

All but 1 of the PLEXs outlined in the PLEX framework were mentioned by participants (*control* was never mentioned by the participants). Table 2 presents illustrative quotes for each of the PLEXs observed. The most commonly occurring PLEXs were competition, discovery, exploration, expression, fellowship, humor, nurture, and sensation.

Competition was the most frequently occurring code. However, discussion of interindividual competition was rare. Discussions of competition were largely centered on an individual-level, low-stake pursuit of trying to achieve a defined goal (eg, finding something specific while walking). Overlapping with experiences of competition, participants commonly articulated experiences of *exploration* and *discovery*. This cluster of codes often reflected a straightforward response to the intervention prompts, typically centered on walking in novel environments with openness and curiosity or physically finding specific items participants felt would constitute suitable responses to the challenges. However, notions of exploration and discovery also sometimes included more abstract discussion of encountering emotive circumstances or discovering new perspectives toward familiar objects. For example, 1 (5%) participant described her pleasant experiences walking through antique stores.

> *I love to walk through antique stores, and it does make me walk. And I will go to a historic district that has a lot of antique stores, and I don't even have to buy anything anymore, because I've got enough junk, but it's so warming to walk through it.* [P11, 69 years old]

Participants' experience of *expression* often reflected their perceptions of their own artistic inclinations and the implications of this for their enjoyment (or lack thereof) of certain challenges. Much discussion linked to experiences of expression was centered on engaging in the creative process or communicating one's identity and values to others in the social group. Numerous participants, for example, expressed pride in their Texan identity. One participant stated,

> *I've got, for instance, on the front of my house, I've got a decoration that has the state of Texas shape on it. I could take a photo of that and explain why I bought that.* [P17, 74 years old]

Table 2. Illustrative quotes for each PLEX[a].

PLEX	Illustrative quote
Captivation	"You have to clear your mind and just put aside all your preconceived notions of what you've seen and done before, and then look at it like somebody that really never has been there." [P15, 65 years old], where "P" stands for "participant."
Challenge	"It would be interesting because my tree knowledge is basic. I did not enjoy my forestry class, back in the sixties, but I keep wanting to revitalize some of my tree knowledge, so I would find this one interesting. I'd probably do it more than twice." [P12, 73 years old]
Competition	"I think this one's more about individual competition. How many did you find and tell us. It's a count thing, so if I don't find them but I went out and walked, am I a loser?" [P19, 65 years old]
Completion	"It's something that I would like to do because I would want to get them all and say, 'See, I got them all.' That's just me." [P6, 68 years old]
Cruelty	"Go to the yacht basin and watch the tourists try to launch their boats. That's always funny." [P15, 65 years old]
Discovery	"Find something that piques our curiosity. And there's a lot of things. Certain buildings. What was the history of that building? Who lives there? Or just different things. I think it'd be interesting." [P10, 70 years old]
Exploration	"To take different routes…to walk in a different way…to go in different areas to walk. Don't walk in the same way all the time. Don't be a creature of habit. Take a chance. Do something different." [P5, 72 years old]
Expression	"I think that's such a personal thing, what people appreciate. One person can say, 'I'm thankful for the house that I live in.' Another person could say, 'I'm thankful for living in Houston or this environment.' You can take pictures of your family…anything like that. It would really depend on the person, I think." [P21, 71 years old]
Fellowship	"…it would initiate conversation. And I need that as much as I do physical activity." [P11, 69 years old]
Humor	"I would just be taking a walk and take a picture of something that made me smile or chuckle, which could be…sometimes, I see a beautiful sunflower, and it just brightens and lightens and [makes me] smile." [P15, 65 years old]
Mindfulness[b]	"Oh, I think it's instead of just walking swiftly through to get to your designation, it's like you notice that some of these streets have bricks on the street to walk on or some of them are higher up where you have to step down." [P4, 72 years old]
Nurture	"And in the spirit of the program, everybody should be encouraging everybody to go see this or try it or do this or do that." [P9, 68 years old]
Relaxation	"So, I think whatever you do should be relaxing and nurturing and calming. And it could be just watching waves come in and out. It could be walking through a forest or like a nature park somewhere here in Greater Houston." [P16, 71 years old]
Sensation	"I'm thinking of the rose garden out at the UTMB[c], which is a nice little, short walk, and the discussion around that could range from, yeah, some of them have no smell at all, but they're gorgeous, and the really old tea flowers, the tea roses, have [a] really nice scent, so there could be discussion about size, color, and scent." [P12, 73 years old]
Subversion	"My husband found a picture on the internet of a cloud that was shooting somebody, the bird. Now, if you don't think I wouldn't put that picture on, you're sadly mistaken, because I would." [P5, 72 years old]
Suffering	"I could end up getting stressed out after 10 or 20 of these if trying to find things that I just don't usually see." [P16, 71 years old]
Thrill	"We've got action scenes. We've got—the police are always out on the seawall, so there's always a cop-type chase." [P10, 70 years old]

[a]PLEX: playful experience.

[b]This experience was not originally featured in the PLEX framework but emerged as an important experience in this study.

[c]UTMB: University of Texas Medical Branch.

Discussions of *fellowship* included sharing or bonding with other participants in the group and having real-world interactions with others (eg, physically meeting with other participants, involving existing walking partners or acquaintances in the walking game). In this context, participants also commonly discussed experiences of *nurture*. These discussions were characterized by prosocial intentions to provide emotional nutriment (eg, positive emotional supportive) or instrumental support (eg, hurricane preparedness tips) to others. Participants also talked about intraindividual experiences of nurture. Such comments were generally centered on engaging in personal reflection or experiencing catharsis that could include feelings of appreciation or nostalgia.

Experiences of *suffering* were relatively rare and usually in reference to participant burden or uncomfortable internal sensations (both physical and emotional, such as reflecting on painful memories of powerful storms). Experiences of *subversion* typically reflected a breach of social norms and encountering potentially uncomfortable or potentially dangerous consequences (eg, reproach from individuals after photographing their likeness or identifying information, such as license plates). Although it is not outlined in the PLEX framework, we added *mindfulness* as a PLEX evoked by the challenges that was related to captivation conceptually but was ultimately judged to be distinct. Rather than forgetting the passing of time and one's surroundings due to absorption in something else, mindfulness involved captivation by and absorption in one's surroundings and the present moment. This code occurred frequently, and such comments typically reflected participants making efforts to purposively focus their attention on the present moment or become more aware of their environment while engaging in the walking challenges.

Value Derived From the Challenges

The second overarching theme we identified was centered on the value participants derived from the challenges. In total, 5 subthemes were reflective of this overarching theme. These were (1) increasing motivation for physical activity, (2) fostering a shift in perspective, (3) increasing knowledge, (4) making meaningful connections with others, and (5) providing health-related benefits.

Participants indicated that by pairing novel, game-like objectives to their walks, the challenges would imbue walks with a sense of directed purpose that could be motivating. Once participant said,

> *To look for things to make it more interesting to walk, because sometimes I think people get bored with just walking and them not having a mission or a purpose.* [P5, 72 years old]

Participants added that the other valued aspects of the challenges (ie, bringing about a shift in perspective, increasing knowledge, making meaningful connections with others, and providing health-related benefits) could further increase motivation for walking. For example, participants indicated that the *shift in perspective* brought about by some of the challenges could improve their autonomous motivation for walking behaviors. One participant stated,

> *And I always see new sunsets or—new sunsets or sunrises—so those are always new beginnings and new renewals. I mean, it seems like if I'm looking at new things or renewals or new beginnings, I would be more involved in my walk. I would have a purpose with my walk.* [P10, 70 years old]

Participants valued many aspects of *knowledge* and learning that the CHALLENGE study intervention featured. This included learning about natural phenomena (eg, types of cloud formations, local flora and fauna). One participant shared,

> *Well, if there's a list of places to go for bird activity, I'd love to find that because my daughter-in-law is a big bird fan, and so she probably doesn't know some places that you've got on your list. But that would be interesting because that would kind of force you out to go to a park or a sanctuary and look for the birds.* [P19, 65 years old]

Participants were also particularly interested in learning about the local attractions that the challenges celebrated (eg, Mardi Gras, Featherfest, Galveston Art Walk, Harvest Moon Regatta boat race, Dickens on the Strand festival), and frequently expressed interest in learning about or visiting specific locations related to the challenges (eg, the historic tall ship *Elissa*, the Seawall, the Strand). Another element of the challenges that participants valued learning about was South Texas history and culture. Many participants were interested in learning about Texas history (eg, historic buildings in Galveston, historic hurricanes) and lore (eg, purportedly haunted locations). These challenges stoked expression of some participants' Texas identity. One participant stated,

> *I'm from Texas and I love Texas symbols, so I probably would be really interested to see what kinds of things are out there…and a lot of people are interested in the history of our state.* [P3, 72 years old]

Participants also valued the potential that engaging with game content may have for facilitating *connectionwith others*. First, participants indicated that the pursuit of completing game challenges may stimulate interesting discussion and knowledge sharing with friends and family members (eg, sharing interesting knowledge gleaned from the challenges). One participant indicated that the challenges would lead to "a lot of discussions. When I look back in history and think of the history and share it with my grandchildren" [P11, 69 years old]. Second, participants talked about including friends and family members in gameplay directly (eg, walking with family members and recruiting them to help craft responses to challenges). Third, participants talked about the prospect of interacting with members of the community during gameplay (eg, asking strangers for permission to photograph their property). Finally, participants valued potentially connecting with other participants in the Facebook group. Some participants were particularly enthusiastic about the prospect of contributing to an uplifting and supportive environment.

Participants made comments indicating that they valued the health-related benefits of increased walking. However, these comments were relatively infrequent. Such comments were relatively general in nature (eg, physical activity facilitating longevity) and did not pertain to specific health-related benefits. For example, 1 participant indicated that she appreciated that the intervention would encourage her to "take pictures of moving things so that you get off your butt and on your feet, because being sedentary can only shorten your life" [P5, 72 years old].

Acceptability of the Challenges

Responses indicated that most of the challenges and the proposed intervention delivery modalities (eg, Facebook) were acceptable. Of the 63 challenges included in the first rounds of interviews, we kept 48 (76%) and removed 15 (24%) based on participant feedback. In the second round, we included 59 challenges in the interview guide, 14 (24%) of which were new (some challenges were combined based on the advice of participants from the first round). Generally, participants understood what the challenges were asking them to do, proffered appropriate example responses, and indicated that the challenges would be doable and enjoyable. Participants appreciated challenges that they felt were both interesting and convenient in the context of their regular routines. Table 3 presents some of the challenges that garnered the most enthusiasm. Participants also appreciated ancillary material provided by CHALLENGE, which included checklists (eg, of local birds, types of clouds), badges (ie, graphic images associated with each challenge awarded to respondents), and props given at the beginning of the study to facilitate responding to certain challenges (eg, silly sunglasses to help conceal identity in selfies).

Participants indicated that some of the challenges presented were not acceptable. Examples of challenges with low acceptability are presented in Table 4. These challenges were typically rejected because they were on a topic the participants did not like (eg, tattoos), required behaviors they did not want to do (eg, pick up garbage), or were felt to be confusing or too abstract (eg, coming up with puppy names). We also removed some challenges that were highly rated in a few cases where participants brought up safety concerns (eg, participants were worried that rocks featured in 1 challenge could be used by others to damage property and that taking photos of out-of-state license plates could provoke anger).

Table 3. Examples of challenges with high acceptability (based on most comments).

Challenge topic (description)	Illustrative quote
Arborist (invited participants to share photos of various trees presented in an ancillary checklist)	"Well, just paying attention to the trees and looking at them and figuring out, well, what type of tree is it? I think this would be really interesting. I would enjoy doing this." [P17, 74 years old]
Change perspective (invited participants to view their environment from a different point of view)	"I think there's this interesting fountain thing on the neighbor's house entrance, and so you could see taking different times of day photos of something like that. That would be interesting, and it would force you to be out at different times." [P19, 65 years old]
Detour-er (invited participants to share photos and take a different route from their normal routine)	"I'd take pictures of things I'd never seen before or didn't know were there." [P3, 72 years old]
Lone Star Bike Rally (referenced a Galveston Island tradition; invited participants to take pictures of interesting vehicles)	"This I could do. I have thousands of pictures I've taken over the years. We love that event. My daddy rode motorcycles. I was raised on them, so all my kids ride." [P8, 71 years old]
Safari at home (invited participants share pictures of animals observed in their environment)	"Yeah, I think that would be interesting because it gives you a different way to look at your environment." [P21, 71 years old]

Table 4. Examples of challenges with low acceptability (based on most comments).

Challenge topic	Illustrative quote
First week of winter (invited participants to share photos reflecting the coming of wintertime)	"But as far as our area, we'd be lucky if it's even cold, you know." [P8, 71 years old]
Neat freak (invited participants to share photos of picking up trash to beautify their environment)	"And would it generate a discussion? Just people like me griping about it." [P6, 68 years old]
Punny ideas (invited participants to share photos reflecting plays on words)	"I might have to look up something on the web or something. I'm not very creative like that." [P7, 71 years old]
Rebel rouser (invited participants to share photos of things that would make a good tattoo)	"I do not believe in tattoos. That would not be a good one for me. My grandson has tattoos all over his body. I just don't think tattoos are cool. I don't think you're supposed to mark your body up like that." [P5, 72 years old]
Terra feeling (invited participants to share photos of different terrains they have walked on)	"I don't walk on anything that's not paved. I don't like to walk on grass, because there might be holes in it. I'm very careful where I walk." [P3, 72 years old]

Nearly every participant suggested changes to challenges, and some participants offered suggestions for challenges that we adopted and subsequently proved to be popular; challenges dedicated to hurricane preparation, the Houston rodeo, and holiday traditions were suggested by participants and included in the final set of challenges. Further, participants provided numerous suggestions for revising presented challenges. Suggested revisions included changes in diction to increase clarity of the challenges and taking measures to provide participants with sufficient guidance for how to respond to challenges. Some challenges were open ended and abstract, and many of participants' suggestions were centered on narrowing the scope of the challenges. Participants also recommended providing example responses following each challenge to facilitate better understanding of exactly how participants might respond. One participant stated,

> *I think having some of those colored examples and pictures for each of these challenges might actually [be] a good stimulant for us 65-and-older people who are not as creative.* [P19, 65 years old]

Thus, in the final intervention, we provide detailed example messages for all challenges. Somewhat related to this, some participants also proposed suggestions for minimizing participant burden. One participant stated,

> *Remember, older people, when you give them a whole lot of tasks to do, they get real anxious. So, I wouldn't give them more than 4 or 5.* [P10, 70 years old]

Finally, participants made suggestions for increasing the relevance and accessibility of the challenges. In most cases, these suggestions pertained to geographic location. The 20 participants in this sample lived in various areas of Southeast Texas and had different experiences of their surroundings. Participants emphasized the important role that one's immediate environment would have in influencing their experience of the walking challenges. Some challenge content was centered on attractions related to a specific region (eg, Galveston Island). Although participants living within this area were often particularly excited about some of these challenges, some who lived outside of this specific area were less enthusiastic. One participant stated,

> *Again, if there were really historical markers in the Clear Lake area, yes. There's so much Galveston-specific things here, and Galveston offers a lot. I don't know how Clear Lake compares to it with something like this.* [P17, 74 years old]

Participants mentioned specific locations as places they would go to complete challenges. Most mentioned were Galveston Island proper and local landmark attractions, including the Seawall, the Strand downtown district, the beach, and the tall ship *Elissa*. Participants also commonly commented on how the specific details of their home environment (even within Galveston Island) would have a bearing on their responses to challenges. For example, 1 participant said,

> *Well, I'm going to tell you, that's going to be hard, there again, because in my neighborhood, there's nothing like that. I've seen it in Galveston where the stumps [have been fashioned] into beautiful pieces of art, but I don't think I could do that here where I walk. This is where I limit myself. I feel safe walking here, and I don't plan on walking anywhere else, you know?* [P14, 69 years old]

Discussion

Principal Findings

In this study, we conducted in-depth interviews with older adult women to investigate the acceptability of CHALLENGE, a digitally delivered game for health centered on the use of celebratory technology to target autonomous motivations for physical activity. We identified 3 overarching themes. These pertained to the PLEXs afforded by the behavioral intervention, the value participants derived from the weekly challenges, and the acceptability of the challenges. Overall, the challenges engendered many and varied PLEXs and participants valued the autonomy-supportive aspects of the behavioral intervention. Participants were particularly interested in challenges that targeted local history, nature, and cultural events, and encouraged us to emphasize these points in the final intervention. Participants provided useful, specific recommendations for adding, removing, and changing challenges that might facilitate targeting autonomous motivations for physical activity.

Our findings are concordant with other studies that have suggested that digitally delivered interventions may be useful for targeting autonomous motivations for health-related behaviors. A study conducted by Chen et al [43] found that a social photography intervention that had participants take daily photos that they believed would make themselves or others happy led to positive outcomes, including feelings of connection and reflection. A social media smoking cessation study found that interactive cocreation of health promotion materials improved information seeking and subjective norms as compared to simply viewing materials [44]. Finally, another trial investigated the effects of social media–delivered daily eating challenges that were either nutritionally prescriptive or nonnutritionally prescriptive (ie, less corrective, more celebratory in nature) [45]. The study found that the nonnutritionally prescriptive challenges were feasible, engaging, and associated with changes in learning about food and diet. Further, assignment to the nonnutritionally prescriptive arm was associated with greater increases in mindful eating than assignment to the nutritionally prescriptive arm. These findings support the conclusions of this study that a celebratory technology approach may hold promise for targeting autonomous motivations for health-related behaviors. More research is needed to better specify and evaluate corresponding mechanisms of successful, long-term behavior change.

PLEXs are important determinants of long-term adherence to digitally delivered physical activity interventions [46]. The findings of this study extend the previous literature by investigating the nature of the PLEXs afforded by a physical activity–promoting game for older adults. Participants indicated that intervention challenges elicited many and varied PLEXs. We initially developed CHALLENGE using the MECHA model [31]. The MECHA model was derived from the Behavior Change Wheel [47,48], and its use ensured specific linkages between game design elements and underlying SDT constructs. In the context of this study, it was particularly useful to have integrated the PLEX framework [40] in the design of CHALLENGE. This framework helped guide our qualitative data analysis to yield a more nuanced understanding of the emotive outcomes of the intervention. Games for health can lead to increases in older adults' physical activity levels and mental and social well-being [49,50], and it may be that the emotive, PLEXs that they engender may mediate their outcomes. The PLEX framework has a quantitative measure that operationalizes its PLEXs and may be used to test such hypotheses [41]. In the randomized controlled study to follow this study, we aim to assess to what degree CHALLENGE elicits the desired experiences and whether this is associated with psychosocial and behavioral outcomes.

The previous literature suggests that it is advantageous to tailor physical activity intervention content to ensure its relevance to the target population [51]. Most individually tailored digital physical activity interventions are tailored to participants' physical activity performance or stage of change [52]; however, tailoring to individuals' cultural identities is also critically important [53,54]. Results of this study highlight the importance of tailoring intervention content to the target population's identities and values. The challenges featured in this study were specific to Southeast Texas from the outset and became more so with refinement following participant feedback. Although the challenges could be completed anywhere, the intervention subject matter was often based on local history, nature, and cultural events. Throughout this study, participants suggested the intervention be expanded to include challenges based on specific local topics, such as historic hurricanes, specific locations, and local festivals. It is not clear to what degree such a specific focus would be necessary for similar interventions in other contexts, but a clear, localized identity emerged as a key driver of interest for this study sample. Although tailoring intervention content in this way may be particularly useful for targeting autonomous motivations, resulting interventions may not be as readily disseminated to other populations and contexts. Extensive qualitative research is likely necessary to create resonant behavioral interventions. A sense of identity as Texan was strong in our sample, but such a geographically oriented identity may not be salient for other populations. Nonetheless, it may be that the platform technologies and general intervention framework featured in this study can help facilitate efficient program adaptation. Content could be adapted for groups that share a common interest rather than geographically limited identity (eg, art, gardening, travel). Geographically limited challenges could also be themed to appeal to individuals who are interested in learning about the culture, history, and language of specific places where they do not necessarily live.

Limitations

This study has several limitations. First, the generalizability of this study is limited by our convenience sampling recruitment method. Participants in our sample mostly identified as non-Hispanic White and may have been particularly receptive to behavior change as they opted to participate in research. Although in-depth, individual interviews are ideal for gathering particularly rich and extensive qualitative data, our study was further limited by a small sample size and the fact that we only interviewed participants on 1 occasion. It is possible that participant sentiment on the topics discussed may shift over time. Further, this study occurred in the context of the uncertainty and sweeping stay-at-home orders associated with the COVID-19 pandemic. These circumstances impacted individuals’ attitudes and propensity to participate in research [55]. Related, please note that personal safety was a major theme related to subversion that we identified in this study, but this topic was not covered here, because we judged the nuance involved to warrant more detailed discussion; these findings are to be presented in an upcoming report.

Conclusion

In conclusion, in this study, we investigated older adult women’s reactions to a social media game designed to target autonomous motivations for physical activity. We identified 3 overarching themes, which reflected the experiences, value, and acceptability afforded by the intervention. Participants enumerated many and varied PLEXs associated with the intervention content and suggested that centering the behavioral intervention on a clear, localized identity would facilitate targeting integrated and identified forms of motivation for engaging in physical activity. This formative work may facilitate taking a celebratory technology approach to targeting autonomous motivations for physical activity in older adult women. Further research is needed to evaluate to what degree findings hold using different methods of assessment and in other populations.

Acknowledgments

This study was supported by the National Institute on Aging of the National Institutes of Health (R01AG06409). This work is a publication of the United States Department of Agriculture, Agricultural Research Service (USDA/ARS), the Children's Nutrition Research Center, and the Department of Pediatrics, Baylor College of Medicine, Houston, Texas, and funded in part with federal funds from the USDA/ARS under Cooperative Agreement 58-3092-5-001 (D Thompson). Support also provided, in part, by a Cancer Center Support Grant (CA16672, PI: P Pisters, The University of Texas MD Anderson Cancer Center), from the NCI/NIH, and from the Center for Energy Balance in Cancer Prevention and Survivorship at The University of Texas MD Anderson Cancer Center.

Data Availability

The data that support the findings of this study are available from the corresponding author, EJL, upon reasonable request.

Authors' Contributions

MCR and EJL conceived the project, conducted the qualitative analysis, and wrote the manuscript with support from MCS, UC, JRB, KB-E, DT, and EV. All authors approved of the final version of this paper.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, et al. Geroscience: linking aging to chronic disease. Cell 2014 Nov 06;159(4):709-713 [FREE Full text] [doi: 10.1016/j.cell.2014.10.039] [Medline: 25417146]

2. Rodgers JL, Jones J, Bolleddu SI, Vanthenapalli S, Rodgers LE, Shah K, et al. Cardiovascular risks associated with gender and aging. J Cardiovasc Dev Dis 2019 Apr 27;6(2):19 [FREE Full text] [doi: 10.3390/jcdd6020019] [Medline: 31035613]
3. Salthouse TA. When does age-related cognitive decline begin? Neurobiol Aging 2009 Apr;30(4):507-514 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2008.09.023] [Medline: 19231028]
4. Padayachey U, Ramlall S, Chipps J. Depression in older adults: prevalence and risk factors in a primary health care sample. S Afr Fam Pract 2017 May 11;59(2):33. [doi: 10.4102/safp.v59i2.4536]
5. Wolitzky-Taylor KB, Castriotta N, Lenze EJ, Stanley MA, Craske MG. Anxiety disorders in older adults: a comprehensive review. Depress Anxiety 2010 Feb;27(2):190-211. [doi: 10.1002/da.20653] [Medline: 20099273]
6. Stevens J, Mack K, Paulozzi L, Ballesteros M. Self-reported falls and fall-related injuries among persons aged ≥65 years--United States, 2006. J Safety Res 2008 Jan;39(3):345-349. [doi: 10.1016/j.jsr.2008.05.002] [Medline: 18571577]
7. Cheng Y, Goodin AJ, Pahor M, Manini T, Brown JD. Healthcare utilization and physical functioning in older adults in the United States. J Am Geriatr Soc 2020 Feb 22;68(2):266-271. [doi: 10.1111/jgs.16260] [Medline: 31755551]
8. Baernholdt M, Hinton I, Yan G, Rose K, Mattos M. Factors associated with quality of life in older adults in the United States. Qual Life Res 2011 Jun 25;21(3):527-534. [doi: 10.1007/s11136-011-9954-z]
9. 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. [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
10. Elsawy B, Higgins K. Physical activity guidelines for older adults. Am Fam Physician 2010 Jan 01;81(1):55-59 [FREE Full text] [Medline: 20052963]
11. Taylor A, Cable N, Faulkner G, Hillsdon M, Narici M, Van Der Bij A. Physical activity and older adults: a review of health benefits and the effectiveness of interventions. J Sports Sci 2004 Aug;22(8):703-725. [doi: 10.1080/02640410410001712421] [Medline: 15370483]
12. Vogel T, Brechat P, Leprêtre PM, Kaltenbach G, Berthel M, Lonsdorfer J. Health benefits of physical activity in older patients: a review. Int J Clin Pract 2009 Feb;63(2):303-320. [doi: 10.1111/j.1742-1241.2008.01957.x] [Medline: 19196369]
13. Jefferis BJ, Sartini C, Lee I, Choi M, Amuzu A, Gutierrez C, et al. Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. BMC Public Health 2014 Apr 19;14:382 [FREE Full text] [doi: 10.1186/1471-2458-14-382] [Medline: 24745369]
14. Evenson K, Buchner D, Morland K. Objective measurement of physical activity and sedentary behavior among US adults aged 60 years or older. Prev Chronic Dis 2012;9:E26 [FREE Full text] [Medline: 22172193]
15. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS One 2016 Dec 20;11(12):e0168614 [FREE Full text] [doi: 10.1371/journal.pone.0168614] [Medline: 27997604]
16. Beishuizen CR, Stephan BC, van Gool WA, Brayne C, Peters RJ, Andrieu S, et al. Web-based interventions targeting cardiovascular risk factors in middle-aged and older people: a systematic review and meta-analysis. J Med Internet Res 2016 Mar 11;18(3):e55 [FREE Full text] [doi: 10.2196/jmir.5218] [Medline: 26968879]
17. Hobbs N, Godfrey A, Lara J, Errington L, Meyer TD, Rochester L, et al. Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. BMC Med 2013 Mar 19;11:75 [FREE Full text] [doi: 10.1186/1741-7015-11-75] [Medline: 23506544]
18. Zubala A, MacGillivray S, Frost H, Kroll T, Skelton DA, Gavine A, et al. Promotion of physical activity interventions for community dwelling older adults: a systematic review of reviews. PLoS One 2017 Jul;12(7):e0180902 [FREE Full text] [doi: 10.1371/journal.pone.0180902] [Medline: 28700754]
19. Caldwell AE, Masters KS, Peters JC, Bryan AD, Grigsby J, Hooker SA, et al. Harnessing centred identity transformation to reduce executive function burden for maintenance of health behaviour change: the Maintain IT model. Health Psychol Rev 2018 Sep 19;12(3):231-253 [FREE Full text] [doi: 10.1080/17437199.2018.1437551] [Medline: 29402182]
20. Deci EL, Ryan RM. Self-determination theory: a macrotheory of human motivation, development, and health. Can Psychol 2008;49(3):182-185. [doi: 10.1037/a0012801]
21. 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]
22. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: a systematic review. Prev Med 2018 Mar;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]
23. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiyagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Oct 04;6(10):e012447. [doi: 10.1136/bmjopen-2016-012447]
24. Grimes A, Harper R. Celebratory technology: new directions for food research in HCI. 2008 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 5-10, 2008; Florence p. 467-476 URL: http://chi2008.org/ [doi: 10.1145/1357054.1357130]
25. Odom W, Banks R, Durrant A, Kirk D, Pierce J. Slow technology: critical reflection and future directions. 2012 Presented at: Proceedings of the Designing Interactive Systems Conference; June 11-15, 2012; Newcastle p. 816-817. [doi: 10.1145/2317956.2318088]

26. Kersten-van Dijk ET, IJsselsteijn WA. Design beyond the numbersharing, comparing, storytelling and the need for a quantified us. Int Des Archit J 2016;29:121-135.
27. Nicholson S. A recipe for meaningful gamification. In: Reiners T, Wood L, editors. Gamification in Education and Business. Heidelberg: Springer; 2015:1-20.
28. 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]
29. 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. [doi: 10.2196/11565]
30. Baranowski MT, Lieberman PD, Buday R, Peng W, Zimmerli L, Wiederhold B, et al. Videogame mechanics in games for health. Games Health J 2013 Aug;2(4):194-204. [doi: 10.1089/g4h.2013.0617] [Medline: 26192223]
31. Robertson MC, Baranowski T, Thompson D, Basen-Engquist KM, Swartz MC, Lyons EJ. Using the behaviour change wheel program planning model to design games for health: development study. JMIR Serious Games 2021 Dec 03;9(4):e29964 [FREE Full text] [doi: 10.2196/29964] [Medline: 34870604]
32. Schell J. The Art of Game Design: A Book of Lenses. 3rd Edition. Boca Raton, FL: Taylor & Francis Group; 2020.
33. Bauman A, Merom D, Bull FC, Buchner DM, Fiatarone Singh MA. Updating the evidence for physical activity: summative reviews of the epidemiological evidence, prevalence, and interventions to promote "active aging". Gerontologist 2016 Apr 18;56 Suppl 2(Suppl 2):S268-S280. [doi: 10.1093/geront/gnw031] [Medline: 26994266]
34. Watkins I, Xie B. Older Adults' Perceptions of Using iPads for Improving Fruit and Vegetable Intake: An Exploratory Study. Care Management Journals 2015 Mar 01;16(1):2-13. [doi: 10.1891/1521-0987.16.1.2]
35. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
36. De Schutter B. Never too old to play: the appeal of digital games to an older audience. Games Culture 2010 May 07;6(2):155-170. [doi: 10.1177/1555412010364978]
37. De Schutter B, Malliet S. The older player of digital games: a classification based on perceived need satisfaction. Communications 2014;39(1):67-88. [doi: 10.1515/commun-2014-0005]
38. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
39. Lewis ZH, Ottenbacher KJ, Fisher SR, Jennings K, Brown AF, Swartz MC, et al. The feasibility and RE-AIM evaluation of the TAME health pilot study. Int J Behav Nutr Phys Act 2017 Aug 14;14(1):106 [FREE Full text] [doi: 10.1186/s12966-017-0560-5] [Medline: 28807041]
40. Korhonen H, Montola M, Arrasvuori J. Understanding Playful User Experience Through Digital Games. : International Conference on Designing Pleasurable Products and Interfaces. Vol 2009. Citeseer; 2009 Presented at: Proceedings from the 2009 International Conference on Designing Pleasurable Products and Interfaces; October 13-16, 2009; Compiegne, France.
41. Boberg M, Karapanos E, Holopainen J, Lucero A. PLEXQ: towards a playful experiences questionnaire. 2015 Presented at: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play; October 5-7, 2015; London, UK p. 381-391. [doi: 10.1145/2793107.2793124]
42. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
43. Chen Y, Mark G, Ali S. Promoting positive affect through smartphone photography. Psychol Well Being 2016;6:8 [FREE Full text] [doi: 10.1186/s13612-016-0044-4] [Medline: 27441169]
44. Namkoong K, Nah S, Record RA, Van Stee SK. Communication, reasoning, and planned behaviors: unveiling the effect of interactive communication in an anti-smoking social media campaign. Health Commun 2017 Jan 27;32(1):41-50. [doi: 10.1080/10410236.2015.1099501] [Medline: 27119592]
45. Epstein D, Cordeiro F, Fogarty J, Hsieh G, Munson S. Crumbs: lightweight daily food challenges to promote engagement and mindfulness. 2016 Presented at: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 5632-5644. [doi: 10.1145/2858036.2858044]
46. Yang X, Ma L, Zhao X, Kankanhalli A. Factors influencing user's adherence to physical activity applications: a scoping literature review and future directions. Int J Med Inform 2020 Feb;134:104039. [doi: 10.1016/j.ijmedinf.2019.104039] [Medline: 31865054]
47. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
48. Michie S, Atkins L, Gainforth H. Changing behaviour to improve clinical practice and policy. In: Dias P, Gonçalves A, Azevedo A, Lobo F, editors. Novos Desafios, Novas Competências: Contributos Atuais da Psicologia. Braga, Portugal: Publicações da Faculdade de Filosofia; 2016:41-60.

49. Loos E, Kaufman D. Positive impact of exergaming on older adults' mental and social well-being: in search of evidence. : Springer; 2018 Presented at: International Conference on Human Aspects of IT for the Aged Population; July 15-20, 2018; Las Vegas, Nevada p. 101-112. [doi: 10.1007/978-3-319-92037-5_9]
50. Ehrari H, Larsen RT, Langberg H, Andersen HB. Effects of playful exercise of older adults on balance and physical activity: a randomized controlled trial. J Popul Ageing 2020 Feb 18;13(2):207-222. [doi: 10.1007/s12062-020-09273-8]
51. Lustria MLA, Noar SM, Cortese J, Van SSK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013 Sep;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
52. Lustria MLA, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
53. Kreuter MW, Lukwago SN, Bucholtz RDDC, Clark EM, Sanders-Thompson V. Achieving cultural appropriateness in health promotion programs: targeted and tailored approaches. Health Educ Behav 2003 Apr;30(2):133-146. [doi: 10.1177/1090198102251021] [Medline: 12693519]
54. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. New York, NY: Routledge; 2012.
55. Cardel MI, Manasse S, Krukowski RA, Ross K, Shakour R, Miller DR, et al. COVID-19 impacts mental health outcomes and ability/desire to participate in research among current research participants. Obesity (Silver Spring) 2020 Dec 21;28(12):2272-2281 [FREE Full text] [doi: 10.1002/oby.23016] [Medline: 32845582]

Abbreviations

CHALLENGE: Challenges for Healthy Aging: Leveraging Limits for Engaging Networked Game-based Exercise
MECHA: Mechanics, Experiences, Change
PLEX: playful experience
SDT: self-determination theory
UTMB: University of Texas Medical Branch

Edited by N Zary; submitted 07.12.21; peer-reviewed by K Waters; comments to author 01.02.22; revised version received 03.02.22; accepted 17.02.22; published 14.04.22.

Please cite as:
Robertson MC, Swartz MC, Christopherson U, Bentley JR, Basen-Engquist KM, Thompson D, Volpi E, Lyons EJ
A Photography-based, Social Media Walking Intervention Targeting Autonomous Motivations for Physical Activity: Semistructured Interviews With Older Women
JMIR Serious Games 2022;10(2):e35511
URL: https://games.jmir.org/2022/2/e35511
doi:10.2196/35511
PMID:35436205

Original Paper

Effect of an Active Video Game Intervention Combined With Multicomponent Exercise for Cardiorespiratory Fitness in Children With Overweight and Obesity: Randomized Controlled Trial

Cristina Comeras-Chueca[1,2,3], MSc; Lorena Villalba-Heredia[1,2,3], MSc; Jose Luis Perez-Lasierra[1,2,3], MSc; Gabriel Lozano-Berges[2,3,4,5,6], PhD; Angel Matute-Llorente[2,3,4,5,6], PhD; German Vicente-Rodriguez[2,3,4,5,6], PhD; Jose Antonio Casajus[1,2,3,5,6], Prof Dr; Alex Gonzalez-Aguero[2,3,4,5,6], PhD

[1]Department of Physiatry and Nursing, Faculty of Health Science, University of Zaragoza, Zaragoza, Spain

[2]GENUD (Growth, Exercise, Nutrition and Development) Research Group, Zaragoza, Spain

[3]EXERNET Red de Investigación en Ejercicio Físico y Salud para Poblaciones Especiales, Zaragoza, Spain

[4]Department of Physiatry and Nursing, Faculty of Health and Sport Science, Universidad de Zaragoza, Zaragoza, Spain

[5]Instituto Agroalimentario de Aragón-IA2, Centro de Investigación y Tecnología Agroalimentaria de Aragón, Universidad de Zaragoza, Zaragoza, Spain

[6]Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición, Madrid, Spain

Corresponding Author:
Alex Gonzalez-Aguero, PhD
Department of Physiatry and Nursing
Faculty of Health and Sport Science
Universidad de Zaragoza
San Juan Bosco St., nº 13; 2th Ft.
Zaragoza, 50009
Spain
Phone: 34 876 55 37 56
Email: alexgonz@unizar.es

Abstract

Background: Childhood overweight and obesity have become major global health problems and are negatively related with the cardiorespiratory fitness (CRF) level in school children and adolescents. Exercise, specifically multicomponent training, is effective for CRF improvement, but the main challenge is to ensure adherence to exercise in children with overweight and obesity. Therefore, new ways of exercising that are more attractive and motivational for this population are needed and playing or training with active video games (AVGs) has been proposed as an effective alternative because they require full-body movement and therefore increase energy expenditure.

Objective: The main aim of this study was to investigate the effects of an AVG intervention combined with multicomponent training on CRF at maximal and submaximal intensities in children with overweight or obesity.

Methods: We recruited 28 children (13 girls and 15 boys) aged 9 to 11 years with overweight or obesity from medical centers and divided them into 2 groups, an intervention group (n=20) that participated in a 5-month supervised AVG exercise program combined with multicomponent exercise, and a control group (n=8) that continued daily activities without modification. A maximal stress test to measure CRF using a walking-graded protocol with respiratory gas exchange was performed by the participants.

Results: The AVG group showed a significant decrease in heart rate and oxygen uptake for the same intensities in the submaximal stages of the maximal treadmill test, as well as a lower oxygen uptake percentage according to the individual maximal oxygen uptake, whereas the control group did not show overall changes. No change in the peak oxygen uptake (VO_{2peak}) was found.

Conclusions: A 5-month AVG intervention combined with multicomponent exercise had positive effects on CRF at submaximal intensity, showing a lower heart rate and oxygen uptake at the same intensities and displaying a lower oxygen uptake percentage according to the individual (VO_{2peak}). Greater benefits were found in children with the highest fat percentage.

Trial Registration: ClinicalTrials.gov NCT04418713; https://clinicaltrials.gov/show/NCT04418713

(JMIR Serious Games 2022;10(2):e33782) doi:10.2196/33782

KEYWORDS

active videogames; VO2peak; obesity; prepuberty; cardiorespiratory; fitness; gaming; childhood; intervention

Introduction

Background

Childhood overweight and obesity have become major global health problems [1,2]. The World Health Organization even refers to obesity as a "global pandemic" [3]. The worldwide prevalence of childhood overweight and obesity remains high, but the rising trends have plateaued in many high-income countries [4]. Nevertheless, the prevalence of overweight was over 30% and that of obesity was over 10% in European children and adolescents in 2016 [5]. This high prevalence is still worrisome because childhood obesity has important health implications such as the increased risk of developing cardiovascular and cardiometabolic diseases (eg, type 2 diabetes, hypertension, or metabolic syndrome), psychosocial problems (eg, low self-esteem, low self-confidence, low self-efficacy, low motivation for physical activity, bullying, or difficulties in establishing relationships) [6,7], and an increased risk of becoming overweight or obese adults [3].

Overweight and obesity have been shown to be negatively related with cardiorespiratory fitness (CRF) levels in school children and adolescents [8,9]. The relationship between both factors has also been found in preschoolers [10] and becomes more pronounced as children grow older [11], suggesting that high CRF levels should be promoted as early as possible as a preventive measure because poor CRF is even associated with the development of cardiometabolic risk factors [12,13] and metabolic syndrome [14]. Preschool BMI is also inversely associated with fitness in adolescence [15]. It is noteworthy that a higher proportion of girls than boys have reduced aerobic capacity [16]. Among all physical fitness variables, the peak oxygen uptake (VO_{2peak}) shows the strongest inverse relationship with BMI and fat mass or body fat percentage [17]. Based on the inverse relationship between CRF and body fat [18,19], it seems that the cutoff point for the negative effects of fatness depending on CRF starts above 16%-20% relative body fat content [18].

On the other hand, an inverse reciprocal relationship was observed between motor competence, and VO_{2peak} and body fatness during childhood [20,21]; therefore, improving motor competence will go hand in hand with improving CRF in the objectives of an exercise intervention.

Exercise, specifically multicomponent training, is effective for improving CRF [22,23]. However, the main challenge is to ensure adherence to exercise in children with overweight and obesity [24]. Therefore, new ways of exercising that are more attractive and motivational for this population are needed.

Playing or training with active video games (AVGs) has been proposed as an effective alternative to exercise and can be as effective as moderate exercising; AVGs are being investigated to determine their effectiveness against childhood obesity. AVGs generally require full-body movement and therefore increase energy expenditure [25,26]; nevertheless, the effects of AVG interventions on CRF are unclear due to the lack of evidence. In fact, the few studies investigating the effects of AVGs on CRF use indirect methods to assess CRF such as the shuttle run test, step test, or The Progressive Aerobic Cardiovascular Endurance Run test. This is a limitation in detecting positive effects at submaximal intensities. In addition, it should be noted that AVG interventions need to be supervised and structured to ensure their effectiveness in improving physical fitness or increasing physical activity [27,28].

Objective

The main aim of this study was to investigate the effects of an AVG intervention combined with multicomponent training on CRF at maximal and submaximal intensities in children with overweight or obesity.

Methods

Ethics Approval

The ethical guidelines of the 1964 Declaration of Helsinki (revised in Fortaleza, 2013) [29] and the Declaration of Taipei [30] were followed in the conduct of this study. The protocol was reviewed and approved by the Research Ethics Committee of the Government of Aragón (certificate number 11/2018, CEICA, Spain). Written informed consent was obtained from all participants and their parents or guardians, after being informed of the nature and possible risks of the experimental procedures in the study.

Study Overview

This randomized controlled trial (RCT) is part of a larger cross-over study (trial registration number: NCT04418713). In this RCT, participants were divided into 2 groups, an intervention group that participated in the AVG exercise program combined with multicomponent exercise, and a control group that continued daily activities without modification.

The recruitment process was carried out through pediatricians from the medical centers. Informative talks about the activity were given in medical centers, and it was the pediatricians themselves who proposed the activity to patients with overweight or obesity who could benefit from the activity. Due to the difficulty of this recruitment process, it was planned to extend the study for another year to obtain a bigger sample. Therefore, a 2:1 randomization was carried out in the first year, prioritizing the AVG intervention. SPSS (version 22.0; SPSS Inc) was used to generate the random allocation sequence and it was performed by the researchers who carried out the project. In the second moment, randomization was going to be 2:1, favoring the control group; unfortunately, it was interrupted by the COVID-19 pandemic and the final groups were not uniform. The second part of the study with a new recruitment process was also interrupted by the COVID-19 pandemic.

Participants

The sample consisted of 28 children (13 girls and 15 boys) with overweight or obesity recruited from medical centers through their pediatricians or from the schools of Zaragoza (Spain). Participants met the following inclusion criteria: aged between 9 and 12 years, Tanner I or II (assessed through direct observation by a physician) and not having had menarche, overweight or obesity calculated by BMI and following the cutoff points of Cole and Lobstein [31], without contraindications for the practice of physical exercise, and without pathologies that worsen with physical exercise. In addition, the following were the exclusion criteria: participating in regular high-level or high-intensity extracurricular physical activities, following any special diet regime, and taking any medication that may interfere with the variables evaluated. The parents and pediatricians were informed about the development of the activity, results, and progress of the children through briefings.

Intervention

The participants were requested to attend 3 sessions per week, lasting approximately 60 minutes each, and the intervention lasted 5 months. The sessions were composed of a regime with a 10-minute warm up, including joint mobility; dynamic flexibility; muscle activation; and core, balance, and coordination exercises. This was followed by the main part, which consisted of 45 minutes of exercise with a combination of AVG and multicomponent exercise, followed by a circuit training dynamic where the participants continuously rotated from AVG to exercises, and finally a 5-minute cooldown part to lower the heart rate (HR) and end the session with static flexibility routines. In general, the sessions consisted of 4 AVGs with an average duration of 8 minutes, and the multicomponent exercise was performed between the AVG sessions. The multicomponent exercise lasted 13 minutes on average per session, divided into 2 or 3 activities with different objectives depending on the planning. Several physical activity and sport professionals supervised the sessions.

In the main part, the AVGs included were the following: the Xbox 360 with Kinect using “Kinect Adventures” and “Kinect Sports;” the Nintendo Wii using “Wii Sports;” “Just Dance” and “Mario and Sonic at the Olympic Games;” dance mats using “Dance Revolution” and “Mario and Sonic at the Olympic Games” adapted from the Nintendo Wii to the dance mats; and the BKOOL interactive cycling simulator connected to a HUAWEI MediaPad T5 AGS2-W09 tablet. The intervention was carried out in 2 locations, the University of Zaragoza and the San Braulio public school in Zaragoza. All the AVGs were provided through funding, and each site was equipped with the AVGs necessary to develop the intervention. The sessions were different every day, following a progression in difficulty and intensity and fulfilling the objectives previously established during planning. The participants did not play all the AVGs in each session, so the number of sessions recorded for each AVG was different. The order in which the activities were carried out was different among the participants, as each participant started in an AVG and changed it after playing.

The AVGs were combined with multicomponent exercises focused on enhancing health-related physical fitness, such as CRF, muscular endurance, and muscular strength, along with coordination and balance. This intervention design combining AVGs with traditional exercise was selected due to a potentially greater energy expenditure [32]. The multicomponent exercise performed had a playful background to enhance motivation and enjoyment.

Outcomes

Anthropometry

All the participants underwent anthropometric examination wearing minimal clothing. Height was measured to the nearest 1 mm with a stadiometer (SECA 225, SECA) and weight to the nearest 0.1 kg with an electronic scale (SECA 861, SECA). BMI was calculated as the weight (kg) divided by the square of the height (m^2).

CRF Measurements

A walking-graded protocol was employed to assess cardiovascular fitness. The test was performed on a treadmill (Quasar Med 4.0, h/p/cosmos) with a face mask fitted. The tests were carried out in the laboratory of the GENUD (Growth, Exercise, Nutrition and Development) research group. The test was explained to the participants, who were fitted with electrodes and had their resting HR measured before starting. After fitting the safety harness, the test started at a comfortable walking pace (3.2 km/h), and the speed was increased by 0.8 km/h every 2 minutes until the participants walked quickly (5.6 km/h), which was the maximum speed reached during the test. Then the slope was increased by 4% every minute until exhaustion or up to a maximal slope of 24%. A sports medicine physician supervised the entire test and performed a preclinical examination to determine if the participant was suitable for performing the stress test. The respiratory gas exchange data were measured breath by breath using open-circuit spirometry (Oxycon Pro, Jaeger/Viasys Healthcare).

Peak values of the VO_2 and HR were defined as the highest average values obtained for any continuous 15-second period. The metabolic cart was calibrated daily with a known gas and volume as recommended by the manufacturer.

The HR was continuously recorded using 12-lead electrocardiography (H12+, Mortara Instrument) from the beginning to the end of the stress test. The maximal HR value was the highest HR value recorded during the last stage of exercise. The blood pressure was also measured with a digital monitor (M3, HEM-72OO-E, Omron Healthcare Europe), for health and safety reasons, before the maximal effort test with the participant lying in a tilt, and during the recovery period in the standing position, both on the right arm. The cuffs were adjusted to the circumference of the tested arm, and the measurement was taken twice. The participants had to be at rest 5 min before the pretest measurement.

Statistical Analyses

SPSS (version 22.0, SPSS Inc) was used to perform all the statistical analyses. Statistical significance was set at $P<.05$ in all tests. Data are presented as means and SDs.

Kolmogorov-Smirnov tests were performed to verify the normal distribution of the variables; several variables did not show a normal distribution, and therefore nonparametric tests were performed.

The Mann-Whitney *U* test was conducted to examine differences between the AVG group and control group for descriptive characteristics and CRF parameters before and after intervention, whereas the Wilcoxon test was performed for within-group comparisons among the preintervention and postintervention measures. The biserial correlation coefficient (r) was calculated using the formula of Fritz et al [33] for nonparametric contrasts/comparisons and the following thresholds were considered: small effect (>0.1), medium effect (>0.3), and large effect (>0.5) [34].

In addition, 2 groups were created based on the baseline body fat percentage using the 50th percentage of the sample to investigate the effect of body fat on the participants' response after the AVG intervention and multicomponent exercise. Low and high CRF categories were established using the 50th percentile based on sex published by Johansson et al [35] for children with overweight and obesity. The reference 50th percentile values for the relative maximal oxygen uptake in boys and girls were 30.8 and 30.6 mL/kg/min, respectively.

The percentage of the VO_{2peak} achieved by all the participants at each stage was calculated through the objective data of their VO_{2peak} obtained in the maximal exercise test. The percentage of change from the pretest measurement to the posttest measurement was calculated for the HR, VO_{2peak}, and percentage of the VO_{2peak}.

Results

Participant Characteristics

The descriptive variables of the participants included in this study are shown in Table 1. No differences between the groups at baseline were found.

No differences between groups were observed in preintervention measurements. Age, weight, height, and lean mass significantly increased in the AVG and control groups from the pretest to the posttest ($P<.05$). The AVG group showed a decreased body fat percentage and BMI z-score ($P<.05$). For the control group, the VO_{2peak} (L/min) increased ($P<.05$).

All the participants completed at least 75% of the sessions, showing good adherence to this AVG intervention combined with multicomponent exercise.

Table 1. Characteristics of the subjects in active video game and control groups.

Variable	Total (N=28)	AVG[a] group (n=20)	Control group (n=8)
Age (years), mean (SD)			
Preintervention	10 (0.8)	10.2 (0.8)	9.7 (0.8)
Postintervention	10.7 (0.8)	10.7 (0.8) *	10.5 (0.8)*
Weight (kg) , mean (SD)			
Preintervention	53.3 (9)	55.3 (9)	48.2 (7.2)
Postintervention	56.2 (10)	57.7 (9.7)*	52.5 (10.3)*
Height (cm) , mean (SD)			
Preintervention	144.6 (7.7)	146 (6.9)	141.1 (8.8)
Postintervention	148.9 (7.5)	149.6 (7.3)*	147 (8.3)*
BMI (kg/m^2)>, mean (SD)			
Preintervention	25.3 (2.8)	25.8 (3)	24.1 (1.7)
Postintervention	25.2 (3)	25.7 (3.1)	24.1 (2.8)
BMI z-score , mean (SD)			
Preintervention	1.95 (0.3)	1.98 (0.4)	1.89 (0.2)
Postintervention	1.84 (0.4)	1.88 (0.42)*	1.72 (0.3)
BMI percentile , mean (SD)			
Preintervention	96.8 (2.2)	96.8 (2.6)	96.8 (1.2)
Postintervention	95.6 (4.1)	95.8 (4.6)	95.1 (2.5)
Body fat percentage, mean (SD)			
Preintervention	40.9 (4)	41.3 (3.7)	39.8 (4.7)
Postintervention	40.1 (4.4)	40.1 (4.5)*	39.9 (4.4)
Lean mass (kg), mean (SD)			
Preintervention	29.7 (4.9)	30.6 (4.7)	27.4 (5)
Postintervention	31.8 (5.5)	32.6 (5.1)*	29.7 (6.3)*
VO_{2peak}[b] (mL/kg/min), mean (SD)			
Preintervention	32.85 (5.9)	32.75 (5.9)	33.1 (6.2)
Postintervention	33.3 (5.3)	32.8 (5.5)	34.7 (4.6)
VO_{2peak} (L/min), mean (SD)			
Preintervention	1.72 (0.29)	1.77 (0.21)	1.6 (0.42)
Postintervention	1.85 (0.29)	1.86 (0.27)	1.81 (0.37)*

[a]AVG: active video games.

[b]VO_{2peak}: peak oxygen uptake.

*Significant differences within groups between preintervention and postintervention (*P*<.05).

Effects of the AVG Intervention Combined With Multicomponent Exercise on HR at Submaximal and Maximal Effort

The effects of the intervention using AVGs combined with multicomponent exercise on the HR for maximal and submaximal efforts are detailed in Table 2. No differences between groups were found neither before nor after the intervention for any HR variable (*P*>.05). Nevertheless, the results showed a significant decrease in the maximal HR of the AVG group (*r*=0.535) and at every submaximal stage: 3.2 km/h (*r*=0.505), 4 km/h (*r*=0.577), 4.8 km/h (*r*=0.689), 5.6 km/h (*r*=0.765), and 5.6 km/h with a slope of 4% (*r*=0.480). Lower HR values were observed for the same intensities, whereas the control group did not show any changes.

Table 2. Heart rates in the different submaximal stages of the maximal stress test by group (N=28).

Heart rate at various levels (bpm[a]), mean (SD)	AVG[b] group (n=20)	Control group (n=8)
Speed: 3.2 km/h; slope: 1%		
Preintervention	118.2 (12.44)	120.38 (10.31)
Postintervention	113.7 (10.6)*	116.38 (9.49)
% change	–3.45 (7.19)	–3.08 (7.21)
Speed: 4 km/h; slope: 1%		
Preintervention	127.6 (13.86)	124 (11.51)
Postintervention	121.75 (11.48)*	122.75 (10.22)
% change	–4.27 (6.34)	–0.75 (6.71)
Speed: 4.8 km/h; slope: 1%		
Preintervention	137.75 (15.42)	133.13 (13.04)
Postintervention	129.9 (10.97)*	130.13 (9.6)
% change	–5.28 (6.22)	–1.91 (6.34)
Speed: 5.6 km/h; slope: 1%		
Preintervention	152.3 (17.42)	146 (14.78)
Postintervention	143.40 (14.43)*	142.63 (12.98)
% change	–5.59 (5.4)	–2.03 (7.03)
Speed: 5.6 km/h; slope: 4%		
Preintervention	157.2 (37.76)	155.5 (16.57)
Postintervention	153.95 (15.74)*	152 (12.88)
% change	–5.48 (44.53)	–1.92 (6.05)
Maximum HR[c] (bpm)		
Preintervention	197.1 (12.24)	191.63 (10.51)
Postintervention	187.2 (22.42)*	191.75 (6.94)
% change	–4.86 (11.1)	0.2 (3.75)

[a]bpm: beats per minute.
[b]AVG: active video game.
[c]HR: heart rate.
*Significant differences within the group between preintervention and postintervention (P<.05).

Effects of the AVG Intervention Combined With Multicomponent Exercise on Submaximal and Maximal Effort Oxygen Uptake and Length of the Maximal Treadmill Test

The effects of the intervention using AVGs combined with multicomponent exercise on VO_2 for the maximal and submaximal efforts are detailed in Table 3. No pretest differences were found between groups (P>.05). After the AVG intervention combined with multicomponent exercise, a decrease in the submaximal VO_2 was found, but no effects were reported for the VO_{2peak}. As shown in Figure 1, lower VO_2 values occur for the same intensities after the intervention at every submaximal stage of the test: at 3.2 km/h (r=0.518), 4 km/h (r=0.434), 4.8 km/h (r=0.593), 5.6 km/h (r=0.535), and 5.6 km/h with a slope of 4% (r=0.551). The control group did not show any change in the VO_2 parameters. In addition, significant differences between groups were found in the submaximal VO_2 after the AVG intervention, with significantly lower VO_2 in the AVG group at 3.2 km/h and 5.6 km/h. The duration in minutes of the maximal treadmill test significantly increased in the AVG and control groups, with no differences observed.

Table 3. Oxygen uptake in the different submaximal stages of the maximal stress test by group (N=28).

Variable, mean (SD)	AVG[a] group (n=20)	Control group (n=8)
VO_2[b] (mL/kg/min); speed: 3.2 km/h; slope: 1%		
Preintervention	13.47 (2.81)	14.59 (1.96)
Postintervention	11.76 (2.26)*	14.86 (2.46)**
% change	–8.73 (27.85)	2.8 (17.3)
VO_2 (mL/kg/min); speed: 4 km/h; slope: 1%		
Preintervention	14.95 (3.05)	15.55 (1.43)
Postintervention	13.15 (2.22)	15.34 (2.58)
% change	–8.73 (23.65)	–1.13 (15.43)
VO_2 (mL/kg/min);speed: 4.8 km/h; slope: 1%		
Preintervention	17.02 (2.9)	17.48 (2.25)
Postintervention	14.92 (2.39)*	17.03 (2.73)
% change	–10.41 (17.92)	–2.06 (13.55)
VO_2 (mL/kg/min); speed: 5.6 km/h; slope: 1%		
Preintervention	19.97 (3.09)	21.50 (2.72)
Postintervention	18.16 (2.88)*	21.13 (2.56)**
% change	–7.49 (18.56)	–0.47 (16.13)
VO_2 (mL/kg/min); speed: 5.6 km/h; slope: 4%		
Preintervention	22.05 (3.18)	23.30 (2.78)
Postintervention	20.59 (2.93)*	23.18 (3.03)
% change	–4.54 (21.04)	–0.04 (11.43)
VO_{2peak} (mL/kg/min)		
Preintervention	32.75 (5.9)	33.11 (6.18)
Postintervention	32.82 (5.54)	34.68 (4.65)
% change	1.23 (13.08)	6.6 (15.36)
Length of the maximal treadmill test (min)		
Preintervention	11.38 (0.94)	11.16 (0.95)
Postintervention	12.37 (1.20)*	12.42 (0.58)*
% change	8.76 (5.58)	11.72 (7.13)
Length of the effort phase of the maximal treadmill test (min)		
Preintervention	3.39 (0.92)	3.25 (0.98)
Postintervention	4.38 (1.16)*	4.42 (0.58)*
% change	36.34 (44.57)	44.4 (38.15)

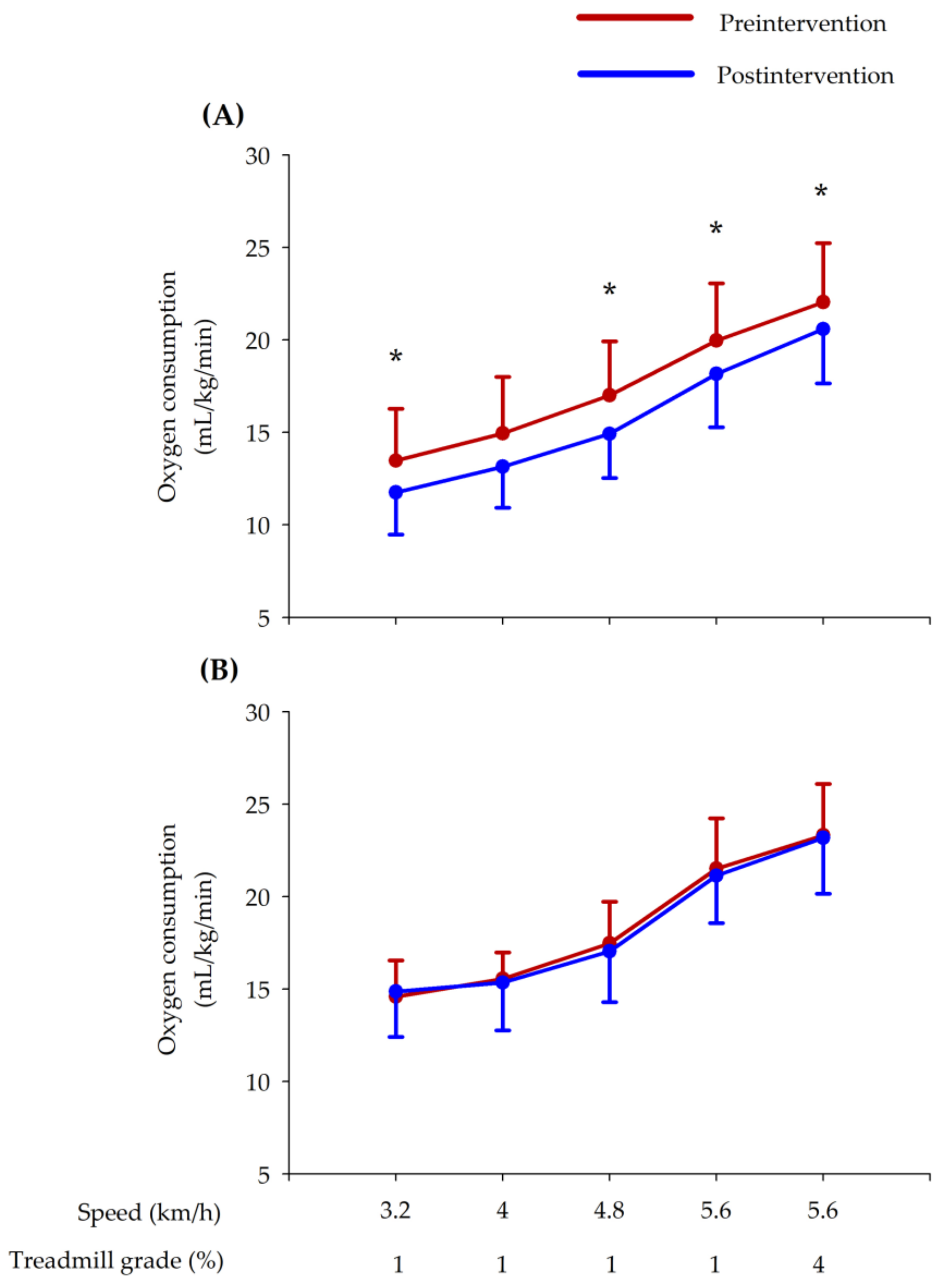
[a]AVG: active video game.

[b]VO_2: oxygen uptake.

* Significant differences within groups between preintervention and postintervention (P<.05).

**Significant differences between AVG and control groups (P<.05).

Figure 1. Changes in VO_2 during the different stages of the maximal stress test observed for the (A) active video game group and (B) control group.



In addition, the VO_2 percentage was calculated for each participant according to the VO_2 peak obtained from the maximal treadmill test. Results reported no differences between groups before the AVG intervention. Based on the VO_2 values in mL/kg/min, the AVG group showed significant decreases in the VO_2 percentage at all submaximal stages of the maximal treadmill test: 3.2 km/h (r=0.609), 4 km/h (r=0.551), 4.8 km/h (r=0.693), 5.6 km/h (r=0.576), and 5.6 km/h with a slope of 4% (r=0.476). On the other hand, the control group showed decreases in the VO_2 at only 4.8 km/h (r=0.792). Significant differences in the VO_2 were found between the AVG and control groups after the AVG intervention at 3.2 km/h.

Effects of the AVG Intervention Combined With Multicomponent Exercise on Submaximal and Maximal Effort Oxygen Uptake Percentage by Groups According to Body Fat Percentage Baseline in the AVG Group

As explained before, the sample was divided into 2 groups according to the 50th percentile of the body fat percentage for the AVG group. The results showed that the participants with a higher body fat percentage showed a significantly decreased HR and VO_2 at the same intensities after the intervention. Specifically, lower HR values were found for the participants in the AVG group with higher body fat percentages at 3.2 km/h (r=0.737), 4 km/h (r=0.662), 4.8 km/h (r=0.751), 5.6 km/h (r=0.886), and 5.6 km/h with a slope of 4% (r=0.864); furthermore, lower HR values were found for the AVG group participants with lower body fat percentages at only 4.8 km/h (r=0.454). On the other hand, lower VO_2 values were found for the AVG group participants with higher body fat percentages in the submaximal stages of the maximal treadmill test at 4 km/h

(r=0.617), 4.8 km/h (r=0.697), and 5.6 km/h (r=0.885); in contrast, the AVG group participants with lower body fat percentages did not show any changes. However, the influence of the body fat percentage on the VO_2 percentage determined from the individual VO_{2peak} of the maximal treadmill test was low, showing that the AVG group participants with higher body fat percentages displayed significant decreases in the VO_2 percentage during the submaximal stage at 4.8 km/h (r=0.590) and 5.6 km/h (r=0.751). However, the AVG group participants with lower body fat percentages displayed significant decreases in the VO_2 percentage during the submaximal stages at 3.2 km/h (r=0.691) and 4.8 km/h (r=0.810). No significant differences between the AVG group participants with higher and lower body fat percentages were found in the baseline and posttest.

Discussion

Principal Findings

The aim of this study was to investigate the effects of an AVG intervention combined with multicomponent training on CRF at maximal and submaximal effort levels in children with overweight and obesity. The main finding of this study was the significant decrease in the HR and VO_2 shown by the AVG group for the same intensities at the submaximal stages of the maximal treadmill test, along with a lower VO_2 percentage according to the individual maximal oxygen uptake. As a reference, the control group did not show overall changes in the HR, VO_2, or VO_2 percentage in the submaximal stages of the maximal treadmill test, except for a significant decrease in the VO_2 percentage at 4.8 km/h. However, no changes in the relative or absolute VO_{2peak} values (in mL/kg/min and L/min) were found for the AVG group, although an increment in the absolute VO_{2peak} value was observed for the control group. This might be due to the weight gain in the control group, mainly fat accumulation. Although we did not expect improvements in the VO_{2peak}, the adequate intensity level required for achieving the desired improvements at the maximum intensities was not reached. This may be explained by the nature of the intervention limiting the intensity to that produced by the AVGs and the inability of the participants to reach such demanding intensities due to excess body fat. However, one of the objectives of the intervention was to improve quality of life by enabling these children to better cope with daily activities, which are often submaximal efforts.

Regarding the CRF endurance level of the participants, mean VO_{2peak} values of 32.75 (SD 5.90) mL/kg/min for the AVG group and 33.11 (SD 6.18) mL/kg/min for the control group were obtained. The children with overweight and obesity who participated in this study showed low CRF levels with an increased risk of health problems due to this low level of CRF in 75% of the participants, according to the cutoff points proposed by Ruiz et al [36]. Moreover, 67.9% of the participants are above the 50th percentile according to Johansson et al [35], which means that the participants in this study had a higher CRF compared to the mean of the normative data of children with overweight or obesity.

In relation to maximal exercise testing, it is difficult for children, especially those with overweight or obesity, to meet all the maximal criteria and demonstrate a plateau in their maximal oxygen uptake [37]. To determine if the exercise test was maximal and the VO_{2peak} data were valid, the percentage of the theoretical maximal HR reached at the end of the test and the respiratory exchange ratio≥1.15 [38] were used. Only 4 of the 28 participants in the preintervention maximal exercise test and 3 of the 28 participants in the postintervention maximal exercise test did not reach 90% of their theoretical maximum HR at the end of the maximal stress test, but they achieved a respiratory exchange ratio very close to (1.13 and 1.14), equal to, or higher than 1.15. Therefore, it can be strongly believed that the tests were maximal.

Thus far, only 7 studies have investigated the effects of AVG interventions on the CRF of children with overweight or obesity, and the results are unclear. The first study that reported the effects of AVGs on the CRF of adolescents with overweight and obesity was that by Adamo et al [39] and the results showed a significant training effect over time with two different interventions: AVG cycling and stationary cycling with music interventions, with sessions of 60 minutes twice per week for 10 weeks. Both interventions produced significant improvements in the peak HR, peak workload, or the time to exhaustion, along with significant reductions in the body fat percentage; however, no significant differences were found between the exercise groups. With this same intervention, Goldfield et al [40] observed that the psychological benefits of these aerobic exercises were related to improved aerobic fitness. These positive effects are in line with our results. Although they show that the positive effects of AVG cycling on CRF are comparable to those of the stationary cycling with music intervention, adherence to this stationary cycling with music intervention was greater. Maddison et al [41] found decreases in body fat percentage with no significant increases in the CRF for the AVG group, measured by the 20-m shuttle test. However, this positive effect of AVGs on body composition in children with overweight or obesity is most likely mediated through improved aerobic fitness [42]. This relationship between body fat percentage and CRF supports the results of this study in which participants with higher body fat percentage had greater improvements in CRF. The interpretation of this result could be that participants with higher body fat percentage have a greater facility for decreasing body fat, which is associated with an increase in CRF. Maloney et al [43] observed no improvements in CRF of the AVG and control groups, assessed by a 3-minute step test after playing Dance Dance Revolution for 12 weeks. This discrepancy may be due to the method used to measure CRF; the participants did not reach adequate intensities and the study did not report the number of sessions per week and the duration of these sessions, which could also explain the lack of positive effects. Christison et al [44] showed that the number of shuttle runs did not change after a 6-month AVG intervention, with 2 sessions per week using several devices. This study has several limitations that could explain the lack of positive effects of the AVG intervention such as the small sample, the difference in the number of participants between the intervention and control groups, the short length of the intervention whose duration was 10 weeks, and the method for measuring the CRF for which the 20-m shuttle run test was used. The most recent study was conducted by Bonney et al [45], who investigated the effect of Wii Fit, in comparison with a task-oriented functional training, on the performance in the shuttle run test and positive effects on CRF in both groups. However, no differences between the AVG and control groups performing the task-oriented functional training were found after a 14-week intervention conducted once a week with each session lasting for 45 minutes. Furthermore, 2 noncontrolled trials studied the effects of AVG on CRF in children with overweight and obesity [46,47]; the limitation of these trials was the lack of a control group, which means that the results should be interpreted with caution. Calcaterra et al [47] observed an improvement in CRF (3.8 mL/kg/min, P<.001) measured by a walking test on a treadmill reaching 85% of the maximal HR after a 12-week intervention using interactive video games. Huang et al [46] showed no effects of AVGs using Nintendo Wii and Xbox Kinect on CRF after 16 sessions, probably due to the inclusion of only 14 participants and the short length of the intervention. A systematic review performed by Zeng and Gao [48] included only 1 RCT [41], which reported positive effects of an AVG intervention in comparison with an exercise group, but these results were unclear due to the inclusion of only 1 study. Given the lack of studies on the usefulness of AVGs to improve CRF, more quality research investigating AVGs as tools to improve CRF is needed. In addition, systematic reviews on AVGs that include CRF are needed, given that CRF is one of the key components of health-related physical fitness and is closely related to quality of life and health.

On the other hand, when the AVG group was divided in 2 groups according to body fat percentage, the results showed that the participants with higher body fat percentage showed significantly decreased HR and VO_2 values at the same intensities after the intervention, which translates into improved efficiency during submaximal efforts. As stated above, participants with higher body fat percentages had a greater range of improvement, and therefore, it is easier for them to achieve improvements, showing significant decreases in the HR and VO_2 at the same submaximal intensities after the AVG intervention combined with multicomponent exercise. However, smaller improvements were observed in participants with lower body fat percentages. These results are supported by previous studies considering the effect of AVGs on CRF, although there are none that determine CRF with breath-by-breath measurements, indirect calorimetry, and the maximal stress test along with gas exchange measurements. As previously stated, there are several studies that support the use of AVGs to improve CRF in children with overweight or obesity. None of these studies compare the outcomes in children based on body fat, and they do not differentiate between overweight and obesity. However, the studies investigating the use of AVGs to improve CRF in children with a healthy weight indicate low effectiveness; among the 6 controlled trials, 3 studies [49-51] report positive effects and the other 3 studies [52-54] report no effects, contrary to the studies involving children with overweight or obesity, most of which found significant improvements in CRF.

Strengths and Limitations

Some limitations must be considered in this study. The number of participants was low, especially in the control group. This could make it more difficult to identify the effects of the AVG intervention and achieve enough statistical power. Another important limitation is the unequal number of participants in the AVG and control groups because the control group could not be completely formed owing the COVID-19 pandemic, which interrupted research activities worldwide.

However, some strengths can be highlighted. The intervention was supervised and structured, with a duration (5 months) and frequency (3 sessions per week and 60 minutes per session) similar to or higher than that of any other previous AVG intervention reporting benefits for this population. Another important aspect is the combination of AVG and multicomponent training focused on CRF, muscular strength, agility, and coordination. In addition, the wide variety of AVGs used should be highlighted. All these devices offer opportunities and possibilities to significantly increase energy expenditure in children with overweight or obesity [26]. Finally, the main strength was the maximal stress test conducted to measure CRF using a walking-graded protocol with respiratory gas exchange measurements that allowed HR and VO_2 values to be recorded at submaximal intensities.

Conclusions

A 5-month intervention of AVGs combined with multicomponent exercise had positive effects on CRF at submaximal intensity, showing lower HR and VO_2 values at the same intensities and lower VO_2 percentages according to the individual VO_{2peak} values. The body fat percentage and the BMI z-score were also reduced after the AVG intervention. In addition, greater improvements were found in children with the highest fat percentage.

Future research could focus on the design and implementation of AVG interventions with higher intensities than those used in this study to produce improvements in CRF at submaximal and maximal effort levels, as this type of intervention appears effective for children with overweight or obesity.

Acknowledgments

We thank all the health centers, pediatricians, and physicians from Zaragoza (Spain) for their involvement. We also thank the University of Zaragoza and the San Braulio Primary School for their engagement and for providing us with the space to carry out the study. Finally, we thank the children who participated in the study and their families for their commitment. This work was funded by the Spanish "Ministerio de Economía y Competitividad" (Project DEP2017-85194-P). CC-C received a grant from “Gobierno de Aragón” (grant DGA IIU/2023/2017), and LV-H received a grant from the Spanish “Ministerio de Economía y Competitividad” (Project DEP2017-85194-P).

Authors' Contributions

All the authors have been actively involved in the planning and execution of the study. JAC and AG-A were the main researchers, and CC-C was the first author. All authors have read and approved the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1112 KB - games_v10i2e33782_app1.pdf]

References

1. Katzmarzyk PT, Chaput J, Fogelholm M, Hu G, Maher C, Maia J, et al. International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE): contributions to understanding the global obesity epidemic. Nutrients 2019 Apr;11(4):1-24 [FREE Full text] [doi: 10.3390/nu11040848] [Medline: 30991687]
2. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug;384(9945):766-781. [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]
3. World Health Organization. Consideration of the evidence on childhood obesity for the Commission on Ending Childhood Obesity: report of the ad hoc working group on science and evidence for ending childhood obesity. World Health Organization. Geneva, Switzerland: World Health Organization; 2016. URL: https://apps.who.int/iris/handle/10665/206549 [accessed 2019-05-03] [WebCite Cache ID 209655]
4. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet 2017 Dec;390(10113):2627-2642 [FREE Full text] [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]
5. Nittari G, Scuri S, Petrelli F, Pirillo I, di Luca NM, Grappasonni I. Fighting obesity in children from European World Health Organization member states. Epidemiological data, medical–social aspects, and prevention programs. Clin Ter 2019 May;170(3):e223-e230 [FREE Full text] [doi: 10.7417/CT.2019.2137] [Medline: 31173054]
6. Gurnani M, Birken C, Hamilton J. Childhood obesity: causes, consequences, and management. Pediatr Clin North Am 2015 Aug;62(4):821-840. [doi: 10.1016/j.pcl.2015.04.001] [Medline: 26210619]
7. McCrindle BW. Cardiovascular consequences of childhood obesity. Can J Cardiol 2015 Feb;31(2):124-130. [doi: 10.1016/j.cjca.2014.08.017]
8. Ceschia A, Giacomini S, Santarossa S, Rugo M, Salvadego D, Da Ponte A, et al. Deleterious effects of obesity on physical fitness in pre-pubertal children. Eur J Sport Sci 2016;16(2):271-278. [doi: 10.1080/17461391.2015.1030454] [Medline: 25866877]
9. Tsiros MD, Coates AM, Howe PRC, Walkley J, Hills AP, Wood RE, et al. Adiposity is related to decrements in cardiorespiratory fitness in obese and normal-weight children. Pediatr Obes 2016 Apr;11(2):144-150. [doi: 10.1111/ijpo.12037] [Medline: 25950151]
10. Tuan SH, Li CH, Sun SF, Li MH, Liou IH, Weng TP, et al. Comparison of cardiorespiratory fitness between preschool children with normal and excess body adipose ~ an observational study. PLoS One 2019 Oct;14(10):e0223907 [FREE Full text] [doi: 10.1371/journal.pone.0223907] [Medline: 31603948]
11. Brunet M, Chaput JP, Tremblay A. The association between low physical fitness and high body mass index or waist circumference is increasing with age in children: the 'Québec en Forme' Project. Int J Obes (Lond) 2007 Apr;31(4):637-643. [doi: 10.1038/sj.ijo.0803448] [Medline: 17006443]
12. Schmidt MD, Magnussen CG, Rees E, Dwyer T, Venn AJ. Childhood fitness reduces the long-term cardiometabolic risks associated with childhood obesity. Int J Obes (Lond) 2016 Jul;40(7):1134-1140. [doi: 10.1038/ijo.2016.61] [Medline: 27102049]
13. Zaqout M, Michels N, Bammann K, Ahrens W, Sprengeler O, Molnar D, et al. Influence of physical fitness on cardio-metabolic risk factors in European children. The IDEFICS study. Int J Obes (Lond) 2016 Jul;40(7):1119-1125. [doi: 10.1038/ijo.2016.22] [Medline: 26857382]
14. Eisenmann JC. Aerobic fitness, fatness and the metabolic syndrome in children and adolescents. Acta Paediatr 2007 Dec;96(12):1723-1729. [doi: 10.1111/j.1651-2227.2007.00534.x] [Medline: 17971189]

15. Pahkala K, Hernelahti M, Heinonen OJ, Raittinen P, Hakanen M, Lagström H, et al. Body mass index, fitness and physical activity from childhood through adolescence. Br J Sports Med 2012 Jul 18;47(2):71-77. [doi: 10.1136/bjsports-2011-090704] [Medline: 22809525]
16. Racette SB, Uhrich ML, White ML, Yu L, Clark BR. Sex differences in FITNESSGRAM® health risk based on aerobic capacity and body composition among urban public elementary school children. Prev Med 2017 Oct;103:56-59. [doi: 10.1016/j.ypmed.2017.07.032] [Medline: 28782561]
17. Ara I, Moreno LA, Leiva MT, Gutin B, Casajús JA. Adiposity, physical activity, and physical fitness among children from Aragón, Spain. Obesity (Silver Spring) 2007 Aug;15(8):1918-1924 [FREE Full text] [doi: 10.1038/oby.2007.228] [Medline: 17712107]
18. Szmodis M, Szmodis I, Farkas A, Mészáros Z, Mészáros J, Kemper HCG. The relationship between body fat percentage and some anthropometric and physical fitness characteristics in pre- and peripubertal boys. Int J Environ Res Public Health 2019 Apr;16(7):1170 [FREE Full text] [doi: 10.3390/ijerph16071170] [Medline: 30939794]
19. García-Hermoso A, Izquierdo M, Alonso-Martínez AM, Faigenbaum A, Olloquequi J, Ramírez-Vélez R. Association between exercise-induced changes in cardiorespiratory fitness and adiposity among overweight and obese youth: a meta-analysis and meta-regression analysis. Children (Basel) 2020 Sep;7(9):147 [FREE Full text] [doi: 10.3390/children7090147] [Medline: 32967205]
20. Lima RA, Pfeiffer KA, Bugge A, Møller NC, Andersen LB, Stodden DF. Motor competence and cardiorespiratory fitness have greater influence on body fatness than physical activity across time. Scand J Med Sci Sports 2017 Dec;27(12):1638-1647. [doi: 10.1111/sms.12850] [Medline: 28164374]
21. Lima RA, Bugge A, Ersbøll AK, Stodden DF, Andersen LB. The longitudinal relationship between motor competence and measures of fatness and fitness from childhood into adolescence. J Pediatr (Rio J) 2019 Jul;95(4):482-488 [FREE Full text] [doi: 10.1016/j.jped.2018.02.010] [Medline: 29782811]
22. García-Hermoso A, Ramírez-Vélez R, Saavedra JM. Exercise, health outcomes, and pædiatric obesity: a systematic review of meta-analyses. J Sci Med Sport 2019 Jan;22(1):76-84. [doi: 10.1016/j.jsams.2018.07.006] [Medline: 30054135]
23. Alves ASR, Venâncio TL, Honório SAA, Martins JMC. Multicomponent training with different frequencies on body composition and physical fitness in obese children. An Acad Bras Cienc 2019 Nov;91(4):e20181264 [FREE Full text] [doi: 10.1590/0001-3765201920181264] [Medline: 31778458]
24. Beedie C, Mann S, Jimenez A, Kennedy L, Lane AM, Domone S, et al. Death by effectiveness: exercise as medicine caught in the efficacy trap!. Br J Sports Med 2015 Feb;50(6):323-324. [doi: 10.1136/bjsports-2014-094389] [Medline: 25807166]
25. Oh Y, Yang S. Meaningful Play 2010 Conference Proceedings. 2010 Oct Presented at: Meaningful Play; October 21-23, 2010; East Lansing, MI, United States URL: https://meaningfulplay.msu.edu/proceedings2010/mp2010_paper_63.pdf
26. Comeras-Chueca C, Villalba-Heredia L, Pérez-Llera M, Lozano-Berges G, Marín-Puyalto J, Vicente-Rodríguez G, 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;17(18):6714 [FREE Full text] [doi: 10.3390/ijerph17186714] [Medline: 32942663]
27. Liang Y, Lau PWC. Effects of active videogames on physical activity and related outcomes among healthy children: a systematic review. Games Health J 2014 Jun;3(3):122-144. [doi: 10.1089/g4h.2013.0070] [Medline: 26196172]
28. De Brito-Gomes JL, Perrier-Melo R, Fernandes Melo de Oliveira S, De Sá Pereira Guimarães FJ, da Cunha Costa M. Physical effort, energy expenditure, and motivation in structured and unstructured active video games: a randomized controlled trial. Human Movement 2016;17(3):190-198. [doi: 10.1515/humo-2016-0021]
29. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
30. World Medical Association. Declaration of Taipei- research on health databases, big data and biobanks. World Medical Association. 2016. URL: https://www.wma.net/what-we-do/medical-ethics/declaration-of-taipei/ [accessed 2020-01-23]
31. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes 2012 Aug;7(4):284-294. [doi: 10.1111/j.2047-6310.2012.00064.x] [Medline: 22715120]
32. 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 May;16(9):783-794. [doi: 10.1111/obr.12287] [Medline: 25943852]
33. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012 Feb;141(1):2-18. [doi: 10.1037/a0024338] [Medline: 21823805]
34. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Routledge; 1988:567.
35. Johansson L, Brissman M, Morinder G, Westerståhl M, Marcus C. Reference values and secular trends for cardiorespiratory fitness in children and adolescents with obesity. Acta Paediatr 2020 Aug;109(8):1665-1671. [doi: 10.1111/apa.15163] [Medline: 31917879]
36. Ruiz JR, Cavero-Redondo I, Ortega FB, Welk GJ, Andersen LB, Martinez-Vizcaino V. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; what level of fitness should raise a red flag? a systematic review and meta-analysis. Br J Sports Med 2016 Dec;50(23):1451-1458. [doi: 10.1136/bjsports-2015-095903] [Medline: 27670254]
37. Armstrong N, Welsman J, Winsley R. Is peak VO2 a maximal index of children's aerobic fitness? Int J Sports Med 1996 Jul;17(5):356-359. [doi: 10.1055/s-2007-972860] [Medline: 8858407]

38. Howley ET, Basseti DR, Welch HG. Criteria for maximal oxygen uptake: review and commentary. Med Sci Sports Exerc 1995 Sep;27(9):1292-1301. [Medline: 8531628]
39. Adamo KB, Rutherford JA, Goldfield GS. Effects of interactive video game cycling on overweight and obese adolescent health. Appl Physiol Nutr Metab 2010 Dec;35(6):805-815. [doi: 10.1139/h10-078] [Medline: 21164552]
40. Goldfield GS, Adamo KB, Rutherford J, Murray M. The effects of aerobic exercise on psychosocial functioning of adolescents who are overweight or obese. J Pediatr Psychol 2012 Nov;37(10):1136-1147. [doi: 10.1093/jpepsy/jss084] [Medline: 23027721]
41. Maddison R, Foley L, Ni MC, Jiang Y, Jull A, Prapavessis H, et al. Effects of active video games on body composition: a randomized controlled trial. Am J Clin Nutr 2011 Jul;94(1):153-163. [doi: 10.3945/ajcn.110.009142] [Medline: 21562081]
42. Maddison R, Mhurchu CN, Jull A, Prapavessis H, Foley LS, Jiang Y. Active video games: the mediating effect of aerobic fitness on body composition. Int J Behav Nutr Phys Act 2012 May;9(1):54. [doi: 10.1186/1479-5868-9-54]
43. Maloney AE, Threlkeld KA, Cook WL. Comparative effectiveness of a 12-week physical activity intervention for overweight and obese youth: exergaming with “Dance Dance Revolution”. Games Health J 2012 Apr;1(2):96-103. [doi: 10.1089/g4h.2011.0009] [Medline: 6193183]
44. Christison AL, Evans TA, Bleess BB, Wang H, Aldag JC, Binns HJ. Exergaming for health: a randomized study of community-based exergaming curriculum in pediatric weight management. Games Health J 2016 Dec;5(6):413-421. [doi: 10.1089/g4h.2015.0097] [Medline: 27911621]
45. Bonney E, Ferguson G, Burgess T, Smits-Engelsman B. Benefits of activity-based interventions among female Adolescents Who are overweight and obese. Pediatr Phys Ther 2019 Oct;31(4):338-345. [doi: 10.1097/PEP.0000000000000636] [Medline: 31568378]
46. Huang HH, Cook L, Harrison J, O’Keefe N, Schultz K, Curtis DA. Exercise by playing active video gaming in overweight and obese children. J Phys Ther Sports Med 2017;01(01):25-31. [doi: 10.35841/physical-therapy.1.1.25-31]
47. Calcaterra V, Larizza D, Codrons E, De Silvestri A, Brambilla P, Abela S, et al. Improved metabolic and cardiorespiratory fitness during a recreational training program in obese children. J Pediatr Endocrinol Metab 2013;26(3-4):271-276. [doi: 10.1515/jpem-2012-0157] [Medline: 23327787]
48. Zeng N, Gao Z. Exergaming and obesity in youth: current perspectives. Int J Gen Med 2016 Aug;9:275-284 [FREE Full text] [doi: 10.2147/IJGM.S99025] [Medline: 27536158]
49. Ye S, Lee JE, Stodden DF, Gao Z. Impact of exergaming on children’s motor skill competence and health-related fitness: a quasi-experimental study. J Clin Med 2018 Sep;7(9):1-11. [doi: 10.3390/jcm7090261] [Medline: 30205483]
50. Fu Y, Burns RD. Effect of an active video gaming classroom curriculum on health-related fitness, school day step counts, and motivation in sixth graders. J Phys Act Health 2018 Sep;15(9):644-650. [doi: 10.1123/jpah.2017-0481] [Medline: 29741441]
51. Lau PWC, Wang JJ, Maddison R. A randomized-controlled trial of school-based active videogame intervention on Chinese children's aerobic fitness, physical activity level, and psychological correlates. Games Health J 2016 Dec;5(6):405-412. [doi: 10.1089/g4h.2016.0057] [Medline: 27855265]
52. Gao Z, Lee JE, Zeng N, Pope ZC, Zhang Y, Li X. Home-based exergaming on preschoolers' energy expenditure, cardiovascular fitness, body mass index and cognitive flexibility: a randomized controlled trial. J Clin Med 2019 Oct;8(10):1745 [FREE Full text] [doi: 10.3390/jcm8101745] [Medline: 31640158]
53. Ye S, Pope Z, Lee J, Gao Z. Effects of school-based exergaming on urban children’s physical activity and cardiorespiratory fitness: a quasi-experimental study. Int J Environ Res Public Health 2019 Oct;16(21):4080. [doi: 10.3390/ijerph16214080] [Medline: 31652857]
54. Azevedo LB, Burges Watson D, Haighton C, Adams J. The effect of dance mat exergaming systems on physical activity and health-related outcomes in secondary schools: results from a natural experiment. BMC Public Health 2014 Sep;14:951 [FREE Full text] [doi: 10.1186/1471-2458-14-951] [Medline: 25217144]

Abbreviations

AVG: active video game
CRF: cardiorespiratory fitness
HR: heart rate
RCT: randomized controlled trial

Edited by N Zary; submitted 23.09.21; peer-reviewed by N Zeng; comments to author 25.10.21; revised version received 12.11.21; accepted 22.04.22; published 24.05.22.

Please cite as:
Comeras-Chueca C, Villalba-Heredia L, Perez-Lasierra JL, Lozano-Berges G, Matute-Llorente A, Vicente-Rodriguez G, Casajus JA, Gonzalez-Aguero A
Effect of an Active Video Game Intervention Combined With Multicomponent Exercise for Cardiorespiratory Fitness in Children With Overweight and Obesity: Randomized Controlled Trial
JMIR Serious Games 2022;10(2):e33782
URL: https://games.jmir.org/2022/2/e33782
doi:10.2196/33782
PMID:35471240

Original Paper

User Experience and Usability of Neumorphism and Gamification User Interface Designs in an HIV Self-Test Referral Program for Men Who Have Sex With Men: Prospective Open-Label Parallel-Group Randomized Controlled Trial

Tsz Ho Kwan[1], PhD; Denise Pui Chung Chan[1], PhD; Shui Shan Lee[1], FRCP

Stanley Ho Centre for Emerging Infectious Diseases, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Shui Shan Lee, FRCP
Stanley Ho Centre for Emerging Infectious Diseases
The Chinese University of Hong Kong
Room 207, Postgraduate Education Centre
Prince of Wales Hospital
Hong Kong
China (Hong Kong)
Phone: 852 22528812
Email: sslee@cuhk.edu.hk

Abstract

Background: Digital interventions have been applied for promoting HIV prevention and care among men who have sex with men (MSM). As user interface (UI) design plays a role in determining usability and user experience (UX), the intervention outcome could be affected.

Objective: In this study, we hypothesized that 2 UI design styles, namely gamification and neumorphism, could impact usability and be differentially preferred by distinct groups of MSM.

Methods: A prospective parallel-group open-label randomized controlled trial was conducted in Hong Kong. Eligible participants were adult MSM recruited by the research team or referred by enrolled participants, who followed instructions for performing an HIV self-test and promoted its use within their social network. Participants were randomized in a 1:1 ratio into either a gamification or neumorphism arm, with primarily visual differences in the UI only. The primary outcome was usability measured by the System Usability Scale (SUS) between the 2 arms. Distinct characteristics of promoters in the 2 arms who gave an SUS score of 80 or above were identified.

Results: Of 463 MSM registered in the study, 232 and 231 were randomized to the gamification and neumorphism arms, respectively. Excluding those who did not request a self-test kit, data from 218 and 216 participants in the gamification and neumorphism arms, respectively, were analyzed (totally 434 participants). With a median SUS score of 80 overall, participants in the neumorphism arm gave a higher score (P<.001), with a higher proportion giving a promoter-level SUS score (P=.002). Promoters used social media for sex networking (P=.02), used pre-exposure prophylaxis in the preceding year (P=.006), had higher satisfaction in UI design (P<.001), and had made a self-test referral (P=.04). In general, higher usability was recorded among participants who were confident in performing the HIV self-test (P<.001), and this was associated with a promoter-level SUS score in both arms. While no other personal characteristics were associated with promoters in the neumorphism arm, those in the gamification arm had higher HIV-related knowledge (P=.01), preferred a specific partner body image type (P=.03), and progressed toward peer referral by completing online training (P=.04).

Conclusions: Both gamified and neumorphic UI designs were well-accepted by MSM. UX and satisfaction of UI were both crucial in influencing the willingness of MSM to promote the application by referring their peers in the community to participate. The simplistic visual design of neumorphism conferred a more general acceptance in the community, whereas gamification was preferred in certain MSM subcommunities. Appropriate UI/UX design should be considered when developing digital interventions targeting the MSM community.

Trial Registration: ClinicalTrials.gov NCT04379206; https://clinicaltrials.gov/ct2/show/NCT04379206

(JMIR Serious Games 2022;10(2):e35869) doi:10.2196/35869

KEYWORDS

HIV; self-test; men who have sex with men; gamification; neumorphism; digital intervention; HIV prevention; user interface; games; digital health

Introduction

Digital technology has been leveraged as an innovative approach to improve health in different settings, especially in HIV care and prevention. It was adopted for promoting HIV testing through the use of specially designed apps targeting the men who have sex with men (MSM) community [1], monitoring adherence of medication intake [2], and providing HIV prevention materials to the community [3]. A qualitative study showed that ease of operation is an important factor for MSM to use and keep using such apps [4]. This is in line with concepts of the theoretical framework in the technology acceptance model stating that perceived ease of use and usefulness contribute to one's intention to use technology [5]. Perception of ease of use is underpinned by user interface (UI) design, such that poor visual design could negatively affect the user experience (UX) and create usage challenges [6]. Therefore, UI design is one of the key elements influencing the usability of a digital intervention [7].

There are many models and styles in UI design, and a popular one is gamification. Gamification is the application of video game elements in nongaming contexts to enhance UX [8]. It has been applied for improving drug adherence in rheumatoid arthritis patients [9] and for supporting clinical decision-making [10]. The effectiveness of gamification has been shown in the promotion of health among MSM. A previous study found that a gamified mobile app could improve HIV knowledge and confidence in adherence to HIV medications among these individuals [11]. The positive effect of gamification was markedly higher in regular users [12]. With rewards like badges, credit points, and actual prizes, and social elements, such as leader boards, the elements in the gamified experience were not just acceptable to MSM, but could also motivate them to engage in promotional activities [13]. On the other hand, a new UI design style, neumorphism or soft UI, emerged in late 2019, which was introduced by Alexander Plyuto [14]. It is a simplistic version of skeuomorphism, which was widely used in the past to imitate a real-world physical object in the UI. Neumorphism simplifies skeuomorphism by making use of light and dark shadows to accentuate the UI element from its background with the same color while capturing the realistic form of the object [15]. The combination of realistic element and the simplistic design can offer an interactive and tactile UX [16].

The 2 UI design styles of gamification and neumorphism were visually distinctive, leading us to hypothesize that different groups of MSM may be in favor of either design. Against these backgrounds, we aimed to assess the differential impact of gamification and neumorphism UI design styles on UX and usability in a web application developed for promoting HIV self-testing among MSM.

Methods

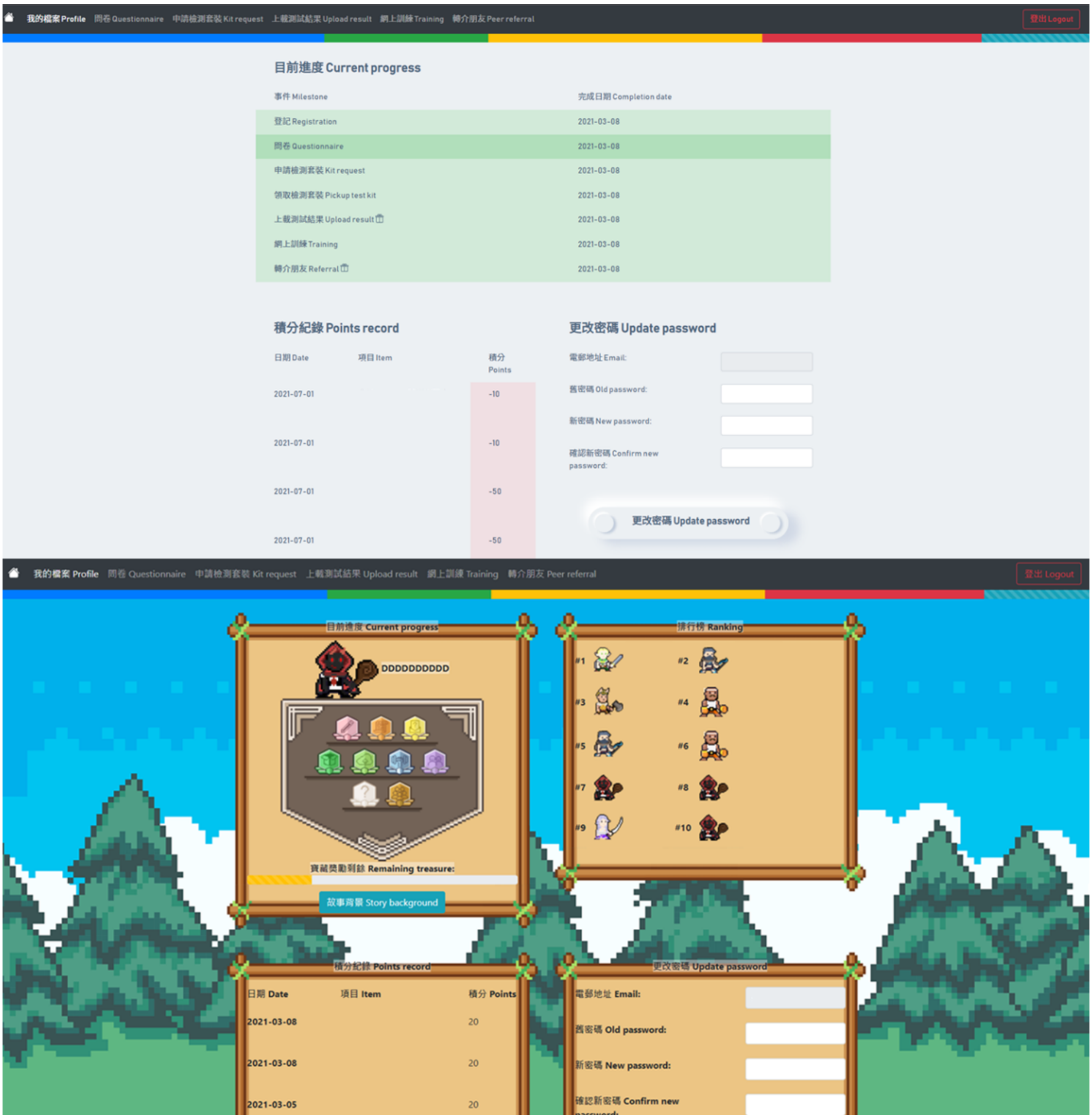
Study Design and Participants

This was a prospective parallel-group open-label randomized controlled study. The study was conducted in Hong Kong, where MSM accounted for a majority of new HIV infections in the past decade [17]. Through a social network approach, enrolled eligible participants were men who ever had sex with other men, who were aged 18 years or above, and who were either referred by another participant in the study or recruited by the research team as seeds. The exclusion criterion was inability to communicate in written English or Chinese. Block randomization was adopted with a block size of 4 and an allocation ratio of 1:1, using a computer-generated list concealed in server-side codes in the web application. With a sample size of 400, a mean difference in the acceptance score of 1% could be detected at an assumed standard deviation of 10% between 2 arms.

Intervention

The participants were randomized into either the neumorphism or gamification UI arm, with the unified objective of undertaking a free HIV self-test using oral fluid or blood at their will. The differences in the 2 arms were related to the UI only (primarily visual). In the neumorphism arm, in terms of form elements, the buttons and frames were in the neumorphism style, while checkboxes and radio buttons remained browser default to balance readability and identification of form elements (Multimedia Appendix 1). A monochromatic color scheme was adopted in compliance with the neumorphism design style. To enrich a gamified UX, the process table in the neumorphic design was replaced by a badge board, and a story background was created (Figure 1). Each of the pages in the gamification arm had a unique backdrop simulating different places in a role-playing game. The profile and questionnaire pages appeared as a forest in daytime and at night, respectively. The test kit request and result upload pages depicted a sword to be picked from the soil and items to be devoted in a shrine, respectively. While training was illustrated as a magical book, the backdrop on the referral page was a tavern. The incentive redemption page was assimilated as a reward bestowed by the "crown." The study procedures remained the same between the 2 arms, namely the questionnaire, test kit request, test result upload, online training, and peer referral.

Figure 1. The profile pages of the neumorphism (upper panel) and gamification (lower panel) designs.



Data Collection

The questionnaire contained questions on demographics; sex networking characteristics, including self-perceived body image types and preferences of sex partners' body image types [18,19], history of sexual behavior, sex networking, and HIV testing; preferences of self-tests; and 17 questions on HIV-related knowledge. The primary outcome of the study was the usability of the web application measured by the System Usability Scale (SUS) [20] and the Single Ease Question (SEQ) [21], as well as the satisfaction of the UI designs of the 2 styles. The SEQ was used after each task throughout the study, whereas the SUS was only displayed after test result upload and peer referral. Participants were free to fill out these scales without obstruction to the main study procedure. Akin to the SUS, an additional question, "I am satisfied with the system interface design," with a scale ranging from 1 (strongly disagree) to 5 (strongly agree), was used to assess participants' acceptance of the UI (UI score). For individual SUS items, a score of at least 4 out of 5 was considered a positive response, whereas for the SEQ, a score of 5 or higher on the scale of 7 was regarded as positive. Overall, the SUS score was calculated according to the standard by inversing responses in even number questions and then multiplying the scores by 2.5. Previous research has shown that an SUS score of 68 is average [22], 71 is acceptable [23] or "good" [24], and 80 indicates a user is willing to promote the product or system to a friend [21]. In the analysis, we adopted the latter 2 thresholds as the definitions of acceptance and participation as a promoter of the system, respectively. Two subscale scores were derived from the SUS by separating items 4 and 10 from the remainder to form the learnability and usability scores, respectively, and the same threshold was used.

Analysis

The scores of the UI, individual SUS items, SUS main scale and subscales, and SEQ at different stages were compared between the 2 arms, and other participant attributes and sex networking characteristics were compared using the Mann-Whitney *U* test. With an overarching aim to examine the willingness of referring peers to undergo self-testing, the SUS scores were dichotomized by the promoter-level threshold, and associated factors were identified per arm by univariable and multivariable logistic regression models. The analyses and reporting of the results followed the CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines [25] (Multimedia Appendix 2). This study has been registered at ClinicalTrials.gov (ID: NCT04379206). All analyses were conducted in R software (R Project for Statistical Computing).

Ethics Approval

This study was approved by the Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee and the Ethics Committee of the Department of Health (CREC reference number 2020.087).

Results

Overall, 463 MSM registered in the study between March 1, 2021, and May 12, 2021, of whom, 232 and 231 were assigned to the gamification and neumorphism arms, respectively (Figure 2). After excluding 21 subjects who had not completed the questionnaire and 8 who failed to request a self-test kit, data from 434 participants were analyzed. Overall, the participants' median age was 28 years (IQR 24-33 years) (Table 1). Moreover, 76.5% (332/434) and 43.3% (188/434) of participants had used location-based MSM social networking apps and social media apps, respectively, to seek male sex partners in the past year, whereas patronizing physical venues, such as bars (33/434, 7.6%), saunas (56/434, 12.9%), and sex parties (26/434, 6.0%), for sex networking amidst the COVID-19 pandemic when social distancing policies were in force was rare. In terms of HIV-related knowledge, 59.0% (256/434) of participants got 11 correct answers and 21.0% (91/434) gave at the most 3 wrong answers out of 17 question items. Overall, 340 (78.3%) had tested for HIV before, of whom, 139 (40.9%) never used an HIV self-test product. More participants were confident to collect sufficient oral fluid (269/434, 62.0%) than blood (197/434, 45.4%) for the self-test. In addition, 66.8% (290/434) were confident in performing self-tests and 70.3% (305/434) were confident in interpreting the test results. A similar proportion (282/434, 65.0%) did not prefer assistance in self-tests.

Figure 2. Study flow.

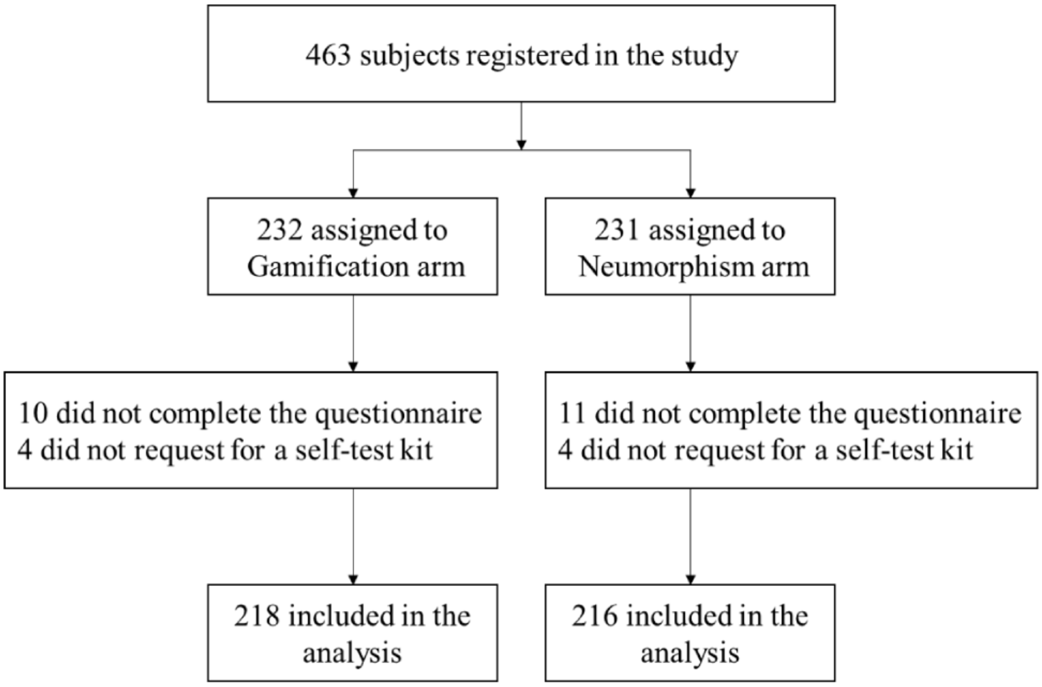


Table 1. Participants' characteristics.

Characteristic	Overall (N=434)	Gamification arm (N=218)	Neumorphism arm (N=216)
Age (years), median (IQR)	28 (24-33)	28 (24-32)	29 (24-34)
Sex networking in the past 1 year, n (%)			
Used location-based MSM[a] social networking apps	332 (76.5)	163 (74.8)	169 (78.2)
Used social media apps	188 (43.3)	90 (41.3)	98 (45.4)
Patronized local bars	33 (7.6)	18 (8.3)	15 (6.9)
Patronized local saunas	56 (12.9)	35 (16.1)	21 (9.7)
Attended local sex parties	26 (6.0)	18 (8.3)	8 (3.7)
Gave a correct answer in at least 11 out of 17 questions on HIV-related knowledge, n (%)	256 (59.0)	130 (59.6)	126 (58.3)
Gave a correct answer in at least 14 out of 17 questions on HIV-related knowledge, n (%)	91 (21.0)	46 (21.1)	45 (20.8)
Ever tested for HIV, n (%)	340 (78.3)	166 (76.1)	174 (80.6)
Ever HIV self-tested (N=340), n (%)	201 (59.1)	101 (60.8)	100 (57.5)
Confidence in self-test procedures, n (%)			
Collecting sufficient oral fluid[b]	269 (62.0)	135 (61.9)	134 (62.0)
Collecting sufficient blood[b]	197 (45.4)	90 (41.3)	107 (49.5)
Correctly performing the HIV self-test[b]	290 (66.8)	142 (65.1)	148 (68.5)
Interpreting the self-test result[b]	305 (70.3)	146 (67.0)	159 (73.6)
Preferring no assistance for the self-test	282 (65.0)	143 (65.6)	139 (64.4)

[a]MSM: men who have sex with men.

[b]Gave a score of 9 or above on a scale of 0 to 10.

The median SUS score given at test result upload was 80.0 (IQR 70.0-90.0), with median learnability and usability scores of 87.5 (IQR 62.5-100.0) and 78.1 (IQR 68.8-90.6), respectively. Higher learnability was scored by MSM who used social media for sex networking (P=.02), preferred sex partners to be of the "cute" body image type (P=.009), never had sex with a female individual (P=.04), and gave more correct answers in HIV-related knowledge (P=.005) (Table 2). They also gave higher confidence scores in sample collection for both oral fluid (P=.002) and blood (P=.004), performance of the self-test (P<.001), and result interpretation (P<.001). A higher usability score was given by participants who used social media for sex networking (P=.002), preferred sex partners with the body image type of "cute" (P=.03) or "decent" (P=.02) but not "mature" (P=.04), were confident in using the self-test (P<.01), and returned the self-test result within 24 hours upon receiving the self-test kit (P=.04) (Table 3).

Table 2. Differential patterns of learnability scores and associated participant characteristics (N=351).

Variable	Learnability score when the variable is false, median (IQR)	Learnability score when the variable is true, median (IQR)	*P* value
Used social media for sex networking in the past 1 year	87.5 (62.5-100.0)	87.5 (75.0-100.0)	.02
Preferred sex partners to be of the “cute” body image type	87.5 (62.5-100.0)	87.5 (75.0-100.0)	.009
Ever had sex with a female individual	87.5 (75.0-100.0)	75.0 (62.5-100.0)	.04
Gave at least 14 correct answers out of 17 questions on HIV-related knowledge	87.5 (62.5-100.0)	87.5 (75.0-100.0)	.004
Confidence in self-test procedures			
Collecting sufficient oral fluid[a]	75.0 (62.5-100.0)	87.5 (75.0-100.0)	.002
Collecting sufficient blood[a]	75.0 (62.5-100.0)	87.5 (75.0-100.0)	<.001
Correctly performing the HIV self-test[a]	75.0 (62.5-100.0)	87.5 (75.0-100.0)	<.001
Interpreting the self-test result[a]	75.0 (62.5-87.5)	87.5 (75.0-100.0)	<.001
Preferring no assistance for the self-test	75.0 (50.0-100.0)	87.5 (75.0-100.0)	<.001

[a]Gave a score of 9 or above on a scale of 0 to 10.

Table 3. Differential patterns of usability scores and associated participant characteristics (N=351).

Variable	Usability score when the variable is false, median (IQR)	Usability score when the variable is true, median (IQR)	*P* value
Used social media for sex networking in the past 1 year	75.0 (68.8-87.5)	84.4 (71.9-93.8)	.002
Preferred body image type of sex partners			
“Cute” type	78.1 (68.8-90.6)	81.3 (75.0-95.3)	.03
“Decent” type	75.0 (68.8-87.5)	78.1 (71.9-93.8)	.02
“Mature” type	81.3 (71.9-93.8)	75.0 (68.8-87.5)	.04
Returned the self-test result within 24 hours upon receiving the self-test kit	78.1 (65.6-89.8)	78.1 (71.9-93.8)	.04
Confidence in self-testing procedures			
Collecting sufficient oral fluid[a]	75.0 (68.8-87.5)	81.3 (71.9-93.8)	.001
Collecting sufficient blood[a]	75.0 (65.6-87.5)	84.4 (75.0-93.8)	<.001
Correctly performing the HIV self-test[a]	75.0 (65.6-84.4)	84.4 (71.9-93.8)	<.001
Interpreting the self-test result[a]	75.0 (65.6-84.4)	81.3 (71.9-93.8)	<.001

[a]Gave a score of 9 or above on a scale of 0 to 10.

Promoters giving at least 80 points in the SUS scale at test result upload (184/351, 52.4%) were more likely to have passed the online training, made a peer referral, used social media for sex networking, used pre-exposure prophylaxis (PrEP) for HIV prevention when having condomless anal intercourse (CLAI) with known male sex partners, and given a higher UI score (*P*<.05). They were also more confident in using the self-test (*P*<.001), while preferring no assistance for the self-test (*P*=.001), and had requested an oral fluid self-test (*P*=.04) (Table 4). While promoters did not give more correct answers in HIV-related knowledge, nonpromoters were more likely to wrongfully consider blood donation an appropriate way for HIV screening in the HIV-related knowledge assessment (odds ratio [OR] 2.57, 95% CI 1.33-4.95; *P*=.01). They also gave a higher SEQ score after the test kit request, result upload, online training, and peer referral (*P*<.01).

Table 4. Characteristics of participants giving a System Usability Scale score of 80 or above (promoters) (N=351).

Variable	Promoters (N=184), n (%)	Nonpromoters (N=167), n (%)	OR[a] (95% CI)	*P* value
Used social media for sex networking in the past 1 year	90 (48.9)	61 (36.5)	1.66 (1.09-2.55)	.02
Used PrEP[b] in the past 1 year (N=165[c])	59 (70.2)	40 (49.4)	2.42 (1.28-4.58)	.006
UI[d] score of 4 or above on a scale of 1-5	174 (94.6)	119 (71.3)	7.02 (3.42-14.42)	<.001
Confidence in self-test procedures				
Collecting sufficient oral fluid[e]	130 (70.7)	89 (53.3)	2.11 (1.36-3.27)	<.001
Collecting sufficient blood[e]	102 (55.4)	53 (31.7)	2.68 (1.73-4.14)	<.001
Correctly performing the HIV self-test[e]	139 (75.5)	90 (53.9)	2.64 (1.68-4.16)	<.001
Interpreting the self-test result[e]	147 (79.9)	97 (58.1)	2.87 (1.79-4.60)	<.001
Preferring no assistance for the self-test	132 (71.7)	92 (55.1)	2.07 (1.33-3.22)	.001
Requested for an oral fluid self-test	118 (64.1)	89 (53.3)	1.57 (1.02-2.40)	.04
SEQ[f] ≥6 on a scale of 1-7 at test kit request (N=348)	147 (80.8)	107 (64.5)	2.32 (1.42-3.77)	<.001
SEQ ≥6 on a scale of 1-7 at result upload	177 (96.2)	123 (73.7)	9.05 (3.94-20.75)	<.001
SEQ ≥6 on a scale of 1-7 at online training (N=213)	53 (44.5)	24 (25.5)	2.34 (1.30-4.22)	.004
SEQ ≥6 on a scale of 1-7 at peer referral (N=119)	43 (61.4)	18 (36.7)	2.74 (1.29-5.83)	.008
Passed online training	115 (62.5)	83 (49.7)	1.69 (1.10-2.58)	.02
Made a peer referral	66 (35.9)	43 (25.8)	1.61 (1.02-2.55)	.04

[a]OR: odds ratio.

[b]PrEP: pre-exposure prophylaxis.

[c]Among participants who had condomless anal intercourse with previously acquired male sex partners.

[d]UI: user interface.

[e]Gave a score of 9 or above on a scale of 0 to 10.

[f]SEQ: Single Ease Question.

Participants in the neumorphism arm gave a higher SUS score than those in the gamification arm (median 82.5 vs 77.5; “Excellent” vs “Good” using the adjective rating scale [24]; *P*<.001), and more participants in the neumorphism arm had an acceptance level (*P*<.001) and a promoter level (*P*=.002) (Table 5). The learnability score was similar between the 2 arms; therefore, the differences in the scores of SUS items 4 and 10 were insignificant. On the other hand, the usability score was higher in the neumorphism arm (*P*<.001). The proportions of participants who were satisfied with the UI (*P*=.01) and agreed to all positively worded statements in the SUS, except for item 9 on confidence in system use, were higher in the neumorphism arm (*P*<.01).

Table 5. Comparison of individual and overall System Usability Scale items, learnability scores, and usability scores between the gamification and neumorphism arms (N=351).

Statement/variable	Gamification arm (N=176), median (IQR)	Neumorphism arm (N=175), median (IQR)	*P* value	Gamification arm (N=176), n (%)[a]	Neumorphism arm (N=175), n (%)[a]	OR[b] (95% CI)	*P* value
I am satisfied with the system interface design	4 (4-5)	4 (4-5)	.01	138 (78.4)	155 (88.6)	0.47 (0.26-0.84)	.01
SUS[c] #1: I think that I would like to use this system frequently	4 (4-5)	4 (4-5)	.007	136 (77.3)	158 (90.3)	0.37 (0.20-0.67)	.01
SUS #2: I found the system unnecessarily complex	2 (1-3)	2 (1-2)	.01	23 (13.1)	17 (9.7)	1.40 (0.72-2.72)	.32
SUS #3: I thought the system was easy to use	4 (4-5)	4 (4-5)	.008	145 (82.4)	162 (92.6)	0.38 (0.19-0.74)	.004
SUS #4: I think that I would need the support of a technical person to be able to use this system	1 (1-2)	1 (1-2)	.40	14 (8.0)	9 (5.1)	1.59 (0.67-3.79)	.29
SUS #5: I found the various functions in this system were well integrated	4 (3-4)	4 (4-5)	<.001	103 (58.5)	145 (82.9)	0.29 (0.18-0.48)	<.001
SUS #6: I thought there was too much inconsistency in this system	2 (1-3)	1 (1-2)	<.001	9 (5.1)	8 (4.6)	1.13 (0.42-2.99)	.81
SUS #7: I would imagine that most people would learn to use this system very quickly	4 (4-5)	4 (4-5)	.005	137 (77.8)	156 (89.1)	0.43 (0.24-0.78)	.004
SUS #8: I found the system very awkward to use	2 (1-2)	1 (1-2)	.002	9 (5.1)	9 (5.1)	0.99 (0.39-2.57)	.99
SUS #9: I felt very confident using the system	4 (4-5)	4 (4-5)	.09	149 (84.7)	159 (90.9)	0.56 (0.29-1.07)	.08
SUS #10: I needed to learn a lot of things before I could get going with this system	2 (1-3)	2 (1-2)	.06	15 (8.5)	14 (8.0)	1.07 (0.50-2.29)	.86
SUS acceptance level (≥71)	N/A[d]	N/A	N/A	109 (61.9)	147 (84.0)	0.31 (0.19-0.51)	<.001
SUS promoter level (≥80)	N/A	N/A	N/A	78 (44.3)	106 (60.6)	0.52 (0.34-0.79)	.002
SUS overall score	77.5 (65.6-89.4)	82.5 (75.0-92.5)	<.001	N/A	N/A	N/A	N/A
SUS learnability score	87.5 (62.5-100.0)	87.5 (75.0-100.0)	.11	N/A	N/A	N/A	N/A
SUS usability score	75.0 (65.6-87.5)	84.4 (75.0-93.8)	<.001	N/A	N/A	N/A	N/A

[a]Percentage refers to the proportion of participants giving a score of 4 or above on a scale of 5 for the respective statement.

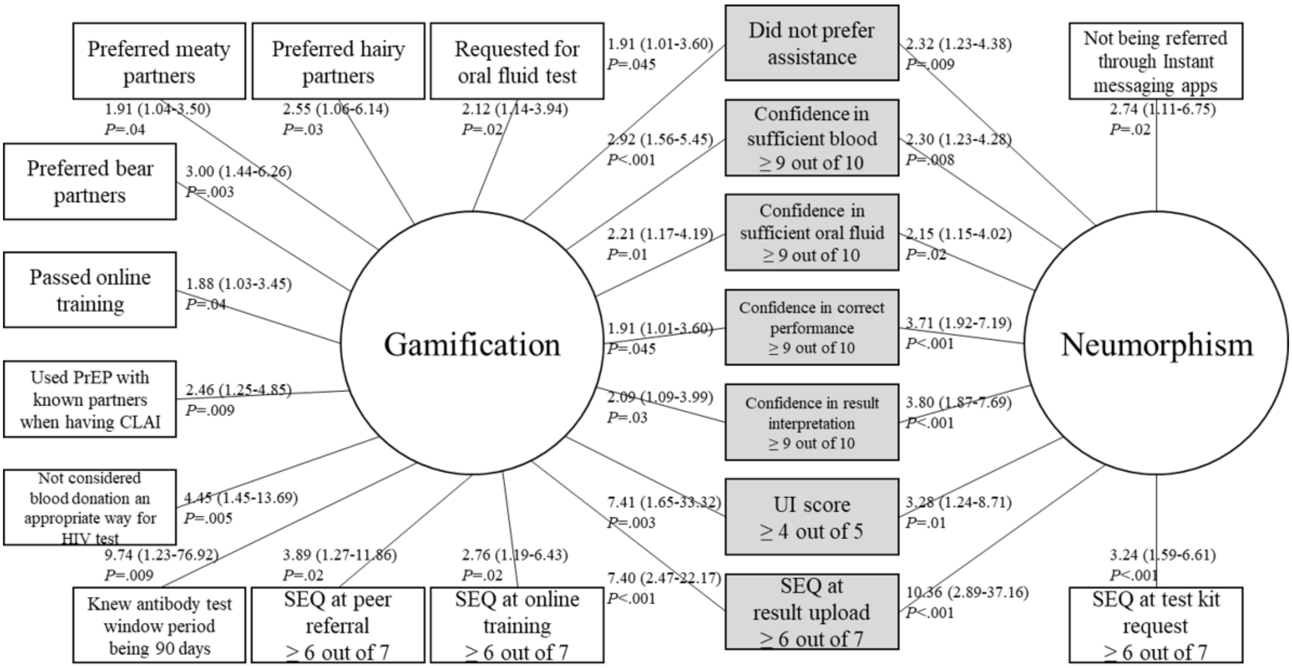
[b]OR: odds ratio.

[c]SUS: System Usability Scale.

[d]N/A: not applicable.

To identify the distinct characteristics associated with a promoter-level SUS score with gamified and neumorphic designs, subgroup analyses were performed (Figure 3). Promoters in the gamification arm were more likely to prefer sex partners to be of “meaty,” “hairy,” or “bear” body image types, request for an oral fluid test, and give a higher SEQ score at peer referral (*P*<.05). They were also less likely to give a wrong answer in some HIV-related knowledge questions (*P*<.05) and more likely to use PrEP when having CLAI with known sex partners (*P*=.009) and pass the online training (*P*=.04). Neumorphism arm promoters were more likely to give a higher SEQ score at test kit request (*P*<.001) and be referred from channels other than instant messaging apps (*P*=.02). The multivariable logistic regression model showed that, other than high UI (*P*<.001) and SEQ (*P*=.02) scores, promoters in the gamification arm were more likely to request for an oral fluid test (*P*=.001), prefer a “bear” body image type partner (*P*=.02), and be confident in collecting blood samples (*P*=.003), while those in the neumorphism arm gave a higher SEQ score at result upload (*P*<.001) and were confident in performing the self-test (*P*<.001).

Figure 3. Factors associated with promoters in the gamification and neumorphism arms. Shaded variables are mutually associated with both arms. Values on the link are odds ratios (95% CIs). CLAI: condomless anal intercourse; PrEP: pre-exposure prophylaxis; SEQ: single ease question; UI: user interface.



Discussion

The social network–based HIV self-test web application was well-accepted by MSM, with an overall median SUS score of 80. Higher system usability was associated with confidence in using the self-test kit and completing the self-test quicker, indicating that this application had linked the user experiences of online activities with offline ones. MSM encountering issues in learning to use the system were less knowledgeable about HIV and more likely to be bisexual. This echoed the results from a previous study showing that bisexual men had lower HIV-related knowledge, who had a higher level of stigma or discrimination associated with the infection [26]. Online social networking tools could be leveraged to promote health behaviors [27], and users gave higher usability and learnability scores. Together with the participants' distinct preferences in sex partners' characteristics, our results informed the need of addressing bisexual MSM and the choice of key opinion leaders of different body image types in future HIV-related promotion campaigns through social media.

The study results confirmed that participants who were satisfied with the system and with a high perceived ease of use throughout the engagement, were more likely to promote it to their peers [21]. Promoters were confident in using the self-test without assistance. They tended to have used PrEP for HIV prevention and possessed adequate HIV knowledge; therefore, they were competent to share their positive self-test experiences and knowledge in HIV, and access to HIV prevention services with their peers, which could in turn invite them to get self-tested. On the other hand, knowledge deficit may have hindered some from referring their peers to engage in prevention programs. This showed that a satisfactory UX, adequate HIV knowledge, and peer referral and support could facilitate HIV testing promotion programs [28].

Participants in the 2 arms were satisfied with the system and the interface design. Compared with the findings in the gamification arm, higher SUS and UI scores were observed among participants assigned to the neumorphism arm, with over half giving a promoter-level score. The impact of UI on UX was only reflected in the usability facet, but not the learnability facet. In reference to the 10 usability heuristics of UI design [29], "esthetic and minimalist design" was found to be of particular importance in contrasting the 2 designs in this study. Neumorphism adopts a simplistic visual design conforming with the heuristics, whereas the colorful graphics in the gamification interface may have distracted users from the primary goal and subsequently weakened its usability. Nevertheless, no matter which arm participants were assigned to, almost all users considered the system consistent and agreed that they did not need to learn a lot of things or ask for support to use the system. These were underpinned by the fourth and sixth heuristics, which were "consistency and standards" and "recognition rather than recall," respectively.

Preferences in UI design style may be related to one's personality, age, gender, and education [30,31]. Shared characteristics among promoters in the 2 arms included higher SEQ and UI scores at the same time when the promoter-level SUS score was given, and confidence in performing the self-test without assistance. No personal attributes were associated with promoters in the neumorphism arm, but those in the gamification arm were featured with partner body image type preference, PrEP use, and HIV-related knowledge. This highlighted that the gamified UI was favored by selected MSM, whereas the neumorphic design was more universally accepted. Our finding provides an insight into the future selection of art styles for designing promotion materials. Although physical appearance traits may differ contextually, a previous study in the United States has suggested that the bear community could be a unique group characterized by body image type in the MSM population, which may be at higher HIV risk than others [32]. If a campaign is launched to target them and their potential partners, a gamified approach could be of use. A previous study also showed that gamification could make HIV and sexually transmitted infection testing less stressful [33]. On the contrary, a neumorphic style would be more suitable to reach the general MSM community.

This study has several limitations. First, the use of sensitive sexual behavioral questions might have given rise to social desirability bias in responders, which we have minimized by adopting a self-administered approach. We also used a short recall period to reduce recall bias. Second, in a randomized parallel-group design, we were unable to capture participants' preferences by presenting them with both UI designs. Instead, we compared the scores between 2 arms and conducted subgroup analyses to identify associating factors with scores reaching the promoter level. The SUS scores used in this study were collected at test result upload; therefore, all participants who did not upload their results were excluded. As the primary goal of the study was to have participants refer their peers for self-testing, uploading a test result was a necessary step, and thus, only those who had uploaded a test result were considered to have completed the study procedure and were included in the analyses. Third, HIV knowledge was not assessed with a conventional validated scale, as such a scale may have been outdated with the recent development of HIV medicine and PrEP. The knowledge questionnaire we designed could however only serve the purpose of identifying a knowledge gap in the community, which was relevant to the theme of the study. Fourth, generalization to other settings with different cultural backgrounds could be limited, particularly for body image types. Finally, neumorphism has an inherent weakness in accessibility due to the lack of contrast in color. We did not enquire about users' color vision; hence, we were unable to consider the impact of color vision deficiency on the usability of the 2 UI designs.

In conclusion, we demonstrated the effective use of a web application that could link users from online activities to offline engagement for the purpose of promoting HIV testing through the collection of self-test kits. The experience of the offline activity could impact subsequent online engagements, which involve peer referral for extending the network of MSM who could go for HIV self-testing as a health promotion strategy. We also identified MSM's preferences in and possible implementation scenarios of gamified and neumorphic UI designs, laying the scientific foundation for future UI and UX designs for internet interventions targeting the MSM community.

Acknowledgments

We thank the Li Ka Shing Institute of Health Sciences for providing technical support. This study was supported by AIDS Trust Fund (MSS 324 R).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the pages in both the gamification and neumorphism arms.

[PDF File (Adobe PDF File), 319 KB - games_v10i2e35869_app1.pdf]

Multimedia Appendix 2

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 1202 KB - games_v10i2e35869_app2.pdf]

References

1. Veronese V, Ryan KE, Hughes C, Lim MS, Pedrana A, Stoové M. Using digital communication technology to increase HIV testing among men who have sex with men and transgender women: Systematic review and meta-analysis. J Med Internet Res 2020 Jul 28;22(7):e14230 [FREE Full text] [doi: 10.2196/14230] [Medline: 32720902]
2. Chai P, Mohamed Y, Bustamante M, Goodman GR, Najarro J, Castillo-Mancilla J, et al. DigiPrEP: A pilot trial to evaluate the feasibility, acceptability, and accuracy of a digital pill system to measure PrEP adherence in men who have sex with men who use substances. J Acquir Immune Defic Syndr 2022 Feb 01;89(2):e5-e15 [FREE Full text] [doi: 10.1097/QAI.0000000000002854] [Medline: 34753871]
3. Simoni JM, Kutner BA, Horvath KJ. Opportunities and challenges of digital technology for HIV treatment and prevention. Curr HIV/AIDS Rep 2015 Dec 28;12(4):437-440 [FREE Full text] [doi: 10.1007/s11904-015-0289-1] [Medline: 26412082]
4. Mitchell JW, Torres MB, Joe J, Danh T, Gass B, Horvath KJ. Formative work to develop a tailored HIV testing smartphone app for diverse, at-risk, HIV-negative men who have sex with men: A focus group study. JMIR Mhealth Uhealth 2016 Nov 16;4(4):e128 [FREE Full text] [doi: 10.2196/mhealth.6178] [Medline: 27852558]
5. 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]

6. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: Descriptive qualitative study. J Med Internet Res 2019 Apr 08;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
7. Hentati A, Forsell E, Ljótsson B, Kaldo V, Lindefors N, Kraepelien M. The effect of user interface on treatment engagement in a self-guided digital problem-solving intervention: A randomized controlled trial. Internet Interv 2021 Dec;26:100448 [FREE Full text] [doi: 10.1016/j.invent.2021.100448] [Medline: 34471610]
8. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification: using game-design elements in non-gaming contexts. In: CHI '11 Extended Abstracts on Human Factors in Computing Systems. 2011 Presented at: Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC, Canada p. 2425-2428. [doi: 10.1145/1979742.1979575]
9. Sciascia S, Radin M, Cecchi I, Di Nunzio P, Buccarano N, Di Gregorio F, et al. Tailoring tofacitinib oral therapy in rheumatoid arthritis: The TuTORApp-A usability study. Int J Environ Res Public Health 2020 May 15;17(10):3469 [FREE Full text] [doi: 10.3390/ijerph17103469] [Medline: 32429245]
10. Melnick ER, Hess EP, Guo G, Breslin M, Lopez K, Pavlo AJ, et al. Patient-centered decision support: Formative usability evaluation of integrated clinical decision support with a patient decision aid for minor head injury in the emergency department. J Med Internet Res 2017 May 19;19(5):e174 [FREE Full text] [doi: 10.2196/jmir.7846] [Medline: 28526667]
11. Hightow-Weidman L, Muessig K, Knudtson K, Srivatsa M, Lawrence E, LeGrand S, et al. A gamified smartphone app to support engagement in care and medication adherence for HIV-positive young men who have sex with men (AllyQuest): Development and pilot study. JMIR Public Health Surveill 2018 Apr 30;4(2):e34 [FREE Full text] [doi: 10.2196/publichealth.8923] [Medline: 29712626]
12. Hightow-Weidman L, Muessig KE, Egger JR, Vecchio A, Platt A. Epic Allies: A gamified mobile app to improve engagement in HIV care and antiretroviral adherence among young men who have sex with men. AIDS Behav 2021 Aug 19;25(8):2599-2617. [doi: 10.1007/s10461-021-03222-y] [Medline: 33740213]
13. Andrade-Romo Z, Chavira-Razo L, Buzdugan R, Bertozzi E, Bautista-Arredondo S. Hot, horny and healthy-online intervention to incentivize HIV and sexually transmitted infections (STI) testing among young Mexican MSM: a feasibility study. Mhealth 2020 Jul;6:28 [FREE Full text] [doi: 10.21037/mhealth.2020.03.01] [Medline: 32632366]
14. Plyuto A. Skeuomorph Mobile Banking. Dribbble. URL: https://dribbble.com/shots/7994421-Skeuomorph-Mobile-Banking [accessed 2022-06-10]
15. Sharma K. Neumorphism (Soft UI) in User interface design - Tutorial. UX Planet. 2019. URL: https://uxplanet.org/neumorphism-in-user-interface-tutorial-c353698ac5c0
16. Osinusi K. The New Trend in UI Design: An Overview of Neumorphism. Toptal. URL: https://www.toptal.com/designers/ui/neumorphic-ui-design [accessed 2022-06-10]
17. Kwan TH, Wong NS, Lui GCY, Chan KCW, Tsang OTY, Leung WS, et al. Incorporation of information diffusion model for enhancing analyses in HIV molecular surveillance. Emerg Microbes Infect 2020 Jan 30;9(1):256-262 [FREE Full text] [doi: 10.1080/22221751.2020.1718554] [Medline: 31997717]
18. Kwan TH, Lee SS. Predictors of HIV testing and their influence on PrEP acceptance in men who have sex with men: A cross-sectional study. AIDS Behav 2018 Apr 10;22(4):1150-1157. [doi: 10.1007/s10461-017-1978-0] [Medline: 29127535]
19. Kwan TH, Lee SS. Bridging awareness and acceptance of pre-exposure prophylaxis among men who have sex with men and the need for targeting chemsex and HIV testing: Cross-sectional survey. JMIR Public Health Surveill 2019 Jul 03;5(3):e13083 [FREE Full text] [doi: 10.2196/13083] [Medline: 31271148]
20. 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, Florida: CRC Press; 1996.
21. Sauro J, Lewis J. Standardized usability questionnaires. In: Quantifying the User Experience. Burlington, Massachusetts: Morgan Kaufmann; 2016:185-248.
22. Lewis J, Sauro J. Item benchmarks for the system usability scale. Journal of Usability Studies 2018;13(3):158-167. [doi: 10.5555/3294033.3294037]
23. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
24. Bangor A, Kortun P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. Journal of Usability Studies 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]
25. 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]
26. Liao M, Wang M, Shen X, Huang P, Yang X, Hao L, et al. Bisexual Behaviors, HIV Knowledge, and Stigmatizing/Discriminatory Attitudes among Men Who Have Sex with Men. PLoS One 2015 Jun 29;10(6):e0130866 [FREE Full text] [doi: 10.1371/journal.pone.0130866] [Medline: 26121250]
27. Kamal N, Fels S, Ho K. Online social networks for personal informatics to promote positive health behavior. In: WSM '10: Proceedings of the Second ACM SIGMM Workshop on Social Media. 2010 Presented at: Second ACM SIGMM Workshop on Social Media; October 25, 2010; Firenze, Italy p. 47-52. [doi: 10.1145/1878151.1878167]

28. Liu Y, Sun X, Qian H, Yin L, Yan Z, Wang L, et al. Qualitative assessment of barriers and facilitators of access to HIV testing among men who have sex with men in China. AIDS Patient Care STDS 2015 Sep;29(9):481-489 [FREE Full text] [doi: 10.1089/apc.2015.0083] [Medline: 26186029]
29. Nielsen J. Enhancing the explanatory power of usability heuristics. In: CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1994 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 24-28, 1994; Boston, MA p. 152-158. [doi: 10.1145/191666.191729]
30. Aktivia R, Djatna T, Nurhadryani Y. Visual usability design for mobile application based on user personality. 2014 Presented at: International Conference on Advanced Computer Science and Information System; October 18-19, 2014; Jakarta, Indonesia. [doi: 10.1109/icacsis.2014.7065893]
31. Arambepola N, Munasinghe L. Empirical analysis of user factors that affect the user interface design in mobile applications. 2020 Presented at: 20th International Conference on Advances in ICT for Emerging Regions (ICTer); November 4-7, 2020; Colombo, Sri Lanka. [doi: 10.1109/icter51097.2020.9325452]
32. Quidley-Rodriguez N, De Santis JP. A literature review of health risks in the bear community, a gay subculture. Am J Mens Health 2017 Nov 30;11(6):1673-1679 [FREE Full text] [doi: 10.1177/1557988315624507] [Medline: 26718773]
33. McCoy SI, Buzdugan R, Grimball R, Natoli L, Mejia CM, Klausner JD, et al. Stick To It: pilot study results of an intervention using gamification to increase HIV screening among young men who have sex with men in California. Mhealth 2018;4:40 [FREE Full text] [doi: 10.21037/mhealth.2018.09.04] [Medline: 30363751]

Abbreviations

CLAI: condomless anal intercourse
MSM: men who have sex with men
OR: odds ratio
PrEP: pre-exposure prophylaxis
SEQ: Single Ease Question
SUS: System Usability Scale
UI: user interface
UX: user experience

Edited by N Zary; submitted 22.12.21; peer-reviewed by V Verma, K Chigudu; comments to author 14.04.22; revised version received 21.04.22; accepted 13.05.22; published 22.06.22.

Please cite as:
Kwan TH, Chan DPC, Lee SS
User Experience and Usability of Neumorphism and Gamification User Interface Designs in an HIV Self-Test Referral Program for Men Who Have Sex With Men: Prospective Open-Label Parallel-Group Randomized Controlled Trial
JMIR Serious Games 2022;10(2):e35869
URL: https://games.jmir.org/2022/2/e35869
doi:10.2196/35869
PMID:35731564

Original Paper

Impact of Visual Game-Like Features on Cognitive Performance in a Virtual Reality Working Memory Task: Within-Subjects Experiment

Eric Redlinger[1]; Bernhard Glas[2]; Yang Rong[3]

[1]Tokyo Institute of Technology, Institute of Innovative Research / Koike & Yoshimura Lab, Tokyo, Japan

[2]Technical University of Munich, Munich, Germany

[3]Tokyo Institute of Technology, Tokyo, Japan

Corresponding Author:
Eric Redlinger
Tokyo Institute of Technology
Institute of Innovative Research / Koike & Yoshimura Lab
2-12-1-I3-20 Ookayama Campus
Tokyo, 152-8552
Japan
Phone: 81 3 5734 362
Email: redlinger.e.aa@m.titech.ac.jp

Abstract

Background: Although the pursuit of improved cognitive function through working memory training has been the subject of decades of research, the recent growth in commercial adaptations of classic working memory tasks in the form of gamified apps warrants additional scrutiny. In particular, the emergence of virtual reality as a platform for cognitive training presents opportunities for the use of novel visual features.

Objective: This study aimed to add to the body of knowledge regarding the use of game-like visual design elements by specifically examining the application of two particular visual features common to virtual reality environments: immersive, colorful backgrounds and the use of 3D depth. In addition, electroencephalography (EEG) data were collected to identify potential neural correlates of any observed changes in performance.

Methods: A simple visual working memory task was presented to participants in several game-like adaptations, including the use of colorful, immersive backgrounds and 3D depth. The impact of each adaptation was separately assessed using both EEG and performance assessment outcomes and compared with an unmodified version of the task.

Results: Results suggest that although accuracy and reaction time may be slightly affected by the introduction of such game elements, the effects were small and not statistically significant. Changes in EEG power, particularly in the beta and theta rhythms, were significant but failed to correlate with any corresponding changes in performance. Therefore, they may only reflect cognitive changes at the perceptual level.

Conclusions: Overall, the data suggest that the addition of these specific visual features to simple cognitive tasks does not appear to significantly affect performance or task-dependent cognitive load.

(JMIR Serious Games 2022;10(2):e35295) doi:10.2196/35295

KEYWORDS

HMD; working memory; gamification; cognitive training; serious game; game; cognitive activity; user performance; visual memory; cognitive; user performance; mobile phone

Introduction

The Emergence of a Cognitive Training Industry

The recent widespread availability of game-like cognitive training products in the form of *apps* on smartphones and tablets, along with a growing public awareness of cognitive training in general, have all contributed to the creation of a multibillion-dollar industry [1]. However, long before the first commercial *brain training* apps appeared on smartphone app stores, a series of widely publicized studies helped set the stage for broader public acceptance of cognitive training. In one such study from 2003, Bavelier and Green [2] documented an increased attentional capacity for players of action video games. Although such differences are easily dismissed as a result of innate abilities or self-selection (eg, individuals with these capacities tend to gravitate toward gaming), the authors notably demonstrated that similar capacities could also be acquired by previously *nongamer* participants through a simple training regimen derived from the same games [2]. Another early, influential training study that received mainstream exposure was that of Jaeggi et al [3]: "Improving fluid intelligence with training on working memory." The authors documented significantly increased *fluid intelligence* (the ability to reason and solve new problems independently of previously acquired knowledge) after cognitive training using a working memory task [3]. The study subsequently received widespread media coverage in outlets such as Wired magazine. Finally, an ambitious, multisite, longitudinal study made additional news headlines in 2017 when it concluded that a kind of adaptive, speed-of-processing task known as *Useful Field of View (UFoV) training* resulted in a significantly decreased risk of dementia up to 10 years after the training intervention [4]. These studies, among others, were instrumental in increasing public awareness of the possibility that explicit training might yield cognitive dividends and contributed to the recent industry boom.

Ongoing Controversy

However, outside the world of public opinion, the overall efficacy of cognitive training remains controversial. Proponents have demonstrated benefits ranging from better scores on standard cognitive assessment tests [5,6] and improved performance in driving aptitude tests [7] to general gains in memory, attention, and visual-spatial ability [8,9]. Nevertheless, recent studies that report little or no benefit from cognitive training, including screen-based training, also exist in substantial numbers [10-13]. The discrepancies in study results are variously attributed to a lack of agreement on experimental methodology, outcome assessment, and the design and implementation of the cognitive training tasks themselves [14,15]. Even simple deviations from the convention can potentially have a major impact on the results. For example, a recent study by Linares et al [16] found no evidence of a *near-transfer* effect (ie, performance improvements in related tasks following training), even between very similar working memory tasks. However, an inspection of their protocol revealed that the training task used in the study was nonadaptive (ie, task difficulty was not adjusted to match participants' natural abilities or prior training gains). This detail may have negatively affected the study's findings, as recent studies argue that adaptive training may be an *essential* component of the success of cognitive training [15,17]. In addition, environmental factors may have contributed to the lack of an observed effect in the Linares et al [16] study as the training sessions were unsupervised, but the assessment sessions were conducted by study staff, which is a source of social stress and a potentially confounding variable [18,19].

Clinical Studies Versus Commercial Cognitive Training

Makers of commercial cognitive training apps regularly advertise their products as *proven effective* and *based on real science* [20]. Some have even licensed or repurposed the very tasks that were used in well-regarded clinical studies [21]. However, just as simple deviations from task design convention may risk negating training effects in clinical studies, the repackaging of promising cognitive training tasks for use in commercial applications must be carefully considered to minimize any risk of introducing new cognitive demands.

For example, one primary assessment outcome used in the 2017 study cited previously was the UFoV test. This test, which contains several subtests to assess short-term recall and spatial memory, requires participants to identify a previously displayed stimulus from among various similarly shaped distractors. Depending on the subtest, the stimuli may appear in either the central visual area, the peripheral area, or both. Although the original task, first developed in 1986, was designed as a clinical assessment tool for use with a standard monochrome computer monitor, commercial adaptations of the UFoV task generally add a number of additional visual and narrative elements to make the task more appealing to customers. These elements include the use of cartoon-like icons; colorful, task-irrelevant background imagery; thematic storylines; scoreboards; and others (Figure 1). Although the use of these elements has become commonplace in the commercial cognitive training industry, consensus on whether such modifications significantly affect the effectiveness of their core tasks is still elusive.

Figure 1. (A) Useful Field of View (UFoV) assessment test compared with (B) commercial cognitive training task dual decision (Posit Science Corporation) designed to train UFoV capacity. The task depicted on the right uses a similar, circular task design but includes colorful icons, scoreboards, and a game-like setting.

A Closer Look at Gamified Tasks

Gamification is generally defined as the process of adding game elements to nonentertainment settings to increase motivation and engagement [22]. Game-like elements may broadly be considered to include visual elements such as colorful icons or patterns, evocative imagery, and playful animations, along with elements designed primarily to stimulate motivation, such as scoreboards and real-time performance feedback. When coupled with cognitive tasks specifically designed to maintain or improve one’s cognitive abilities, this result may be referred to as *gamified cognitive training*. The embrace of gamification as a method of increasing user engagement and enjoyment of otherwise dull, repetitive tasks is indeed supported by a significant number of studies [22-25]. However, the full picture of the potential impact of gamification on cognitive *performance* is less conclusive. Two recent comprehensive reviews [22,26] that examined the use of gamification strategies in brain training and general cognitive assessment studies overwhelmingly found that although gamified training appears to boost participant motivation, study heterogeneity impeded the drawing of clear conclusions with respect to performance or ecological validity (ie, the degree to which experimental results are generalizable to real-life situations). For example, the authors of the first study [26] identified no fewer than 28 game-like elements used in the 33 studies surveyed. These included positive and negative task feedback, time pressure, storylines or narrative elements, performance status displays, and many others. The second survey [22], from 2020, found that of the 49 papers examined, no study reported on the effect of a single element alone and that the game elements were investigated only in combination, making it impossible to establish whether individual elements had measurable effects.

For example, the 2017 study by Mohammed et al [24] compared two adaptations of an n-back task: a stripped-down task and one that contained a visually rich display combined with multiple audio soundtracks. Although the authors found increased task enjoyment for the game condition, there were no significant differences in the long-term outcomes between the gamified and nongamified tasks. However, given the complex set of features included in the gamified version, they acknowledged that more granularity was perhaps needed to fully understand which features might prove to be more successful than others [24].

Another study with a sizable participant pool (n=107) found negative correlations between certain game elements and task performance [27]. The authors speculated that unneeded stress and new cognitive demands might have been induced by distracting game elements such as persistent score displays, leading to reduced performance. However, rather than individual game elements added to a bare-bones task, the study design removed specific game elements from a larger group of game features. This approach seems to leave the possibility open for the remaining elements to compensate for the removal of a single element, making it difficult to know for sure which element or elements might have specifically accounted for the *new* cognitive demands [27].

In summary, as gamification encompasses a great number of individual elements, a lack of precision and homogeneity between studies has hampered the ability to draw consensus conclusions regarding which game elements, if any, may affect task performance. In addition, although motivational features such as scoreboards and real-time performance feedback have been widely studied [25,27-31], the specific impact of certain purely visual features, such as 3D depth and colorful, immersive backgrounds, is less well-documented, despite being increasingly encountered in consumer products such as game systems and dedicated virtual reality (VR) headsets.

Therefore, this study aimed to add to the body of knowledge regarding the use of game-like visual design elements by specifically examining the application of two particular visual features: immersive, colorful backgrounds and the use of 3D depth. These features were specifically chosen because of their underrepresentation in previous studies and their increased use in VR and augmented reality technology, a rapidly growing consumer market segment that also contains cognitive training products. We hypothesized that task performance may be adversely affected by additional visual processing demands but that the motivational effects documented by previous researchers may, in turn, compensate or reverse these effects. Finally, using electroencephalography (EEG) as an additional quantitative outcome, we hoped to gain insight into the possible neural correlates for any observed performance impact.

Methods

Study Design and Sample Size Considerations

Two primary outcomes were used to examine the impact of visual gamified design elements on cognitive task performance. Cognitive activity will be broadly measured along the midline using EEG (see the *EEG Data* section for details regarding EEG). Raw task performance was assessed by analyzing task accuracy and participant response time. The experimental task was a simple visual working memory task that required the participant to pick out the previously displayed stimulus from several distractors. To better control the testing environment, the task was coded for display in a head-mounted display (HMD) environment rather than a traditional monitor screen (see the *Experimental Task* section).

The use of an HMD serves two purposes: (1) to precisely control the display brightness and task visual angle (VA) across participants and experimental conditions and (2) to minimize potentially distracting external stimuli. For these and other reasons, several recent papers have recommended the use of HMDs, describing them as among the "most fitting platforms for applying nonpharmacological computerized neurocognitive assessments" [14] and a "frontier for neurorehabilitation" [32].

The current experimental task was previously used in a related study exploring changes in the size and position of visual stimuli and showed a robust effect size (>0.5) between conditions [33].

For this study, we undertook several additional modifications to further boost statistical power. First, to reduce between-subject variability, an adaptive task design was used in which task difficulty was automatically modulated to ensure maximum participant engagement. The precise method is described in more detail in the *Adaptive Task* section.

Second, an intrasubject protocol design exposed each participant to all experimental conditions. This enabled the use of repeated-measures ANOVA and Wilcoxon signed-rank sum tests, which are known to be particularly robust in establishing significance in small-n situations [34,35]. With this study design, we used the G-Power algorithm [36] to determine that a sample size of n=20 should be sufficient to enable us to achieve adequate statistical power at the 5% confidence level.

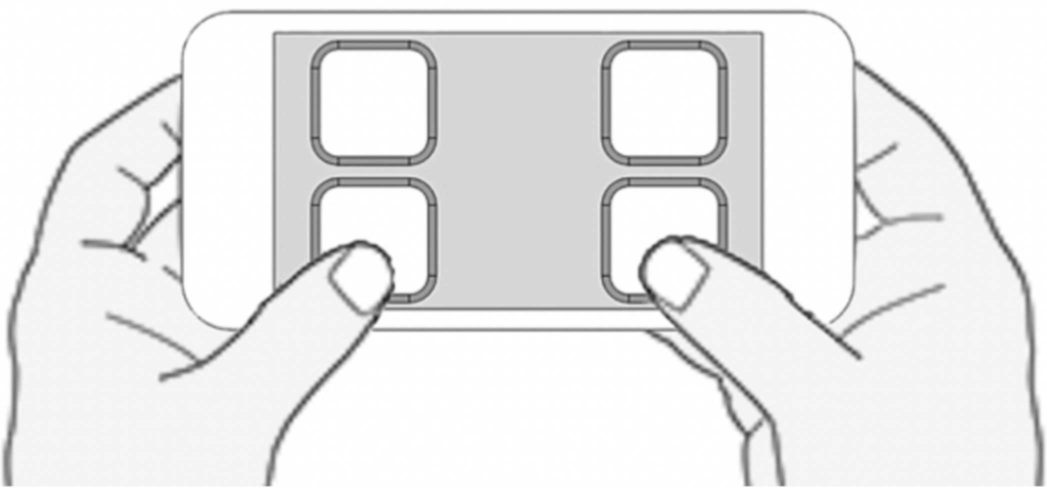
Test Environment

A standalone HMD (HTC Vive Focus, HTC Corp) in its default configuration was chosen for the test environment. The cognitive training task was created in *Unity 3D*, a programming environment commonly used for creating 3D visual content for VR headsets (Unity 3D; Unity Software Inc).

HMD systems typically rely on handheld pointers for user input. However, such input devices are not appropriate for EEG studies, as they could introduce muscle-related artifacts. To address this, a touch screen smartphone was programmed to wirelessly send network commands to the HMD. A soft foam overlay with holes corresponding to the locations of the on-screen virtual buttons was added to the screen. With this combination, the participants could identify the smartphone controls in a tactile manner using only their hands without any need to view the screen. This is crucial as the participant cannot see the smartphone screen while wearing the headset.

During the experiment, participants were seated and instructed to hold the smartphone controller in their laps, cradled by both hands (Figure 2). The experimental task was performed by tapping the virtual buttons on the screen with both thumbs while minimizing other body movements.

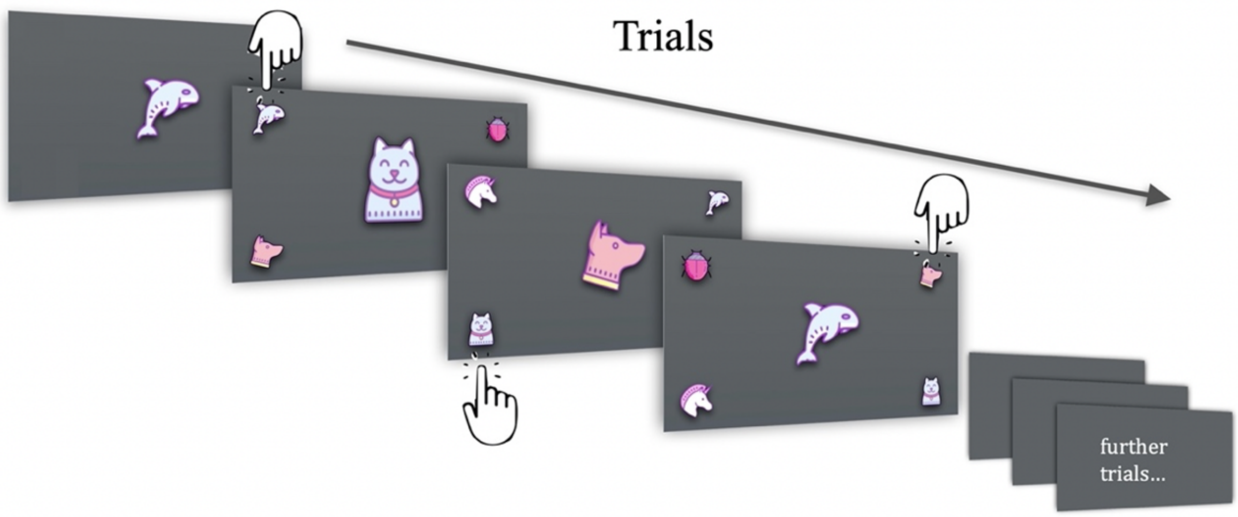
Figure 2. Smartphone interface with a foam overlay.



Experimental Task

To emulate a typical commercial cognitive training task, we designed our core task to incorporate a number of cognitive processes drawn from both gaming research [2,37] and the cognitive training literature [3,4,38]. These included visual memory recognition, divided attention, perceived time pressure, and distractor avoidance. The experimental task required participants to focus on a sequence of stimuli located in the center of the HMD screen. At the start of each new trial, the previously displayed center stimulus was moved to 1 of the 4 corners of the display, and a new stimulus took its place in the center. A total of 3 randomly chosen images were placed in the remaining 3 corners so that the screen always contained 1 center image and 4 images in the outer corners. To proceed to the next trial, the participant was asked to identify the stimulus that was *previously* in the center of the display. Participants performed this task by tapping the virtual button on the smartphone screen corresponding to the location of the object they wished to select. Once a choice was made by the participant, the answer choices disappeared, and the stimulus currently at the center of the display was reassigned to 1 of the 4 corners. A new stimulus then took its place in the center (Figure 3). A trial was also considered ended if the allotted time elapsed before a selection was made. Please see the *Adaptive Task* and *Experimental Protocol* sections for specific details related to trial times and durations.

Figure 3. Sample trials showing the center stimulus and peripheral answer choices: the current answer choices (small images) and the next stimulus (large image) are displayed simultaneously. Preceding the first trial, only the initial stimulus is displayed. Participants select correct answers in the subsequent trials, as demonstrated with cartoon hands. In each case, the correct answer corresponds to the center stimulus from the previous slide.



In each trial, the center stimulus and incorrect answer choices were selected at random by the software such that no duplicate images appeared together. The trials lasted approximately 1.2 seconds (SD 116 milliseconds) on average and were designed to elicit continuous cognitive load as both the current answer choices and the following stimulus were displayed simultaneously. This was to minimize the usual *peaks and valleys* in the cognitive activity that often accompany tasks that alternate between stimulus presentation and participant response.

The goal of choosing this experimental task was to create a minimally complex task that could nevertheless reliably elicit sufficient cognitive load with little prior task training. Although the basic mechanism is inspired by the classic n-back task, we restricted our task to 1-back to minimize individual differences in performance ability commonly associated with higher degrees of *n* [39].

The figures themselves are from a set of 20 cartoon animal images, all drawn in a similar style but differing in shape and color. The image collection was licensed for noncommercial use from a popular internet vendor. It was chosen for its design similarity to prevailing commercial cognitive product designs, which frequently use a similar cartoon design aesthetic.

Adaptive Task

An adaptive model was chosen for the experimental task to ensure similar engagement levels for all the participants. As the experiment progressed, the task difficulty increased incrementally until the participant failed to respond within the allotted time window or made ≥2 sequential mistakes. The task difficulty level was reflected in the amount of time available for the participant to choose an answer. As the difficulty level rose, this amount of time decreased in 50-millisecond intervals. Conversely, if the difficulty level decreased, more time (50 milliseconds) was made available to complete each trial. The prevailing task difficulty level affected the experiment in the following two ways:

1. A visible countdown timer just below the task area displayed the amount of time allocated to make a selection. As the trial time progressed, the bar's contents filled incrementally from left to right, reminding the participant to answer as quickly as possible. The bar was purposefully designed to be as unobtrusive as possible so as not to distract from the primary task (Figure 4).
2. Failure to make a selection within the allotted time resulted in the trial being marked incorrect, and the next stimulus was presented. Making any selection (correct or incorrect)

resulted in the timer pausing briefly (200 milliseconds) before being reset for the next trial.

At the end of each trial, the response (or failure to respond), the reaction speed, and accuracy were recorded. Only trials in which the participant actively made a selection were included in the reaction time assessment.

Figure 4. Adaptive task countdown timer.

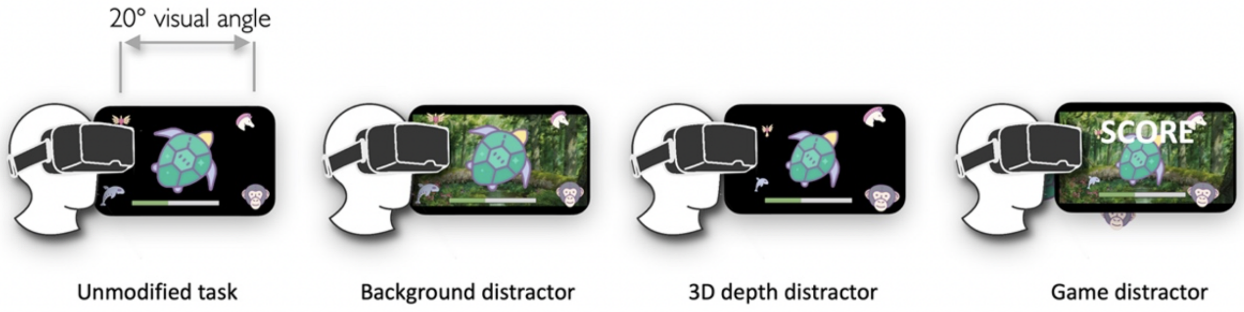
Experimental Protocol

A total of 20 participants, aged 21 to 48 (mean 28.6, SD 7.7) years, were recruited from among students and staff at the Tokyo Institute of Technology and agreed to participate in the experiment after signing an informed consent form. The 20 participants included 6 (30%) women and 14 (70%) men, all right-handed, with no history of color vision disorders. In addition, all participants reported having had a previous experience using an HMD.

The protocol was executed in the following order: task training, EEG baseline activity measurement, and experimental conditions. The EEG baseline measurement phase (60 seconds) involved viewing a black background with open eyes to record nominal cognitive activity with no visual stimuli.

The experimental conditions comprised 4 distinct visual representations of the same core task: unmodified (the stimuli were simply placed on a flat plane against a black background), background distractor (stimuli + irrelevant background image), 3D depth distractor (stimuli presented at different virtual distances from the participant), and game distractor (dynamic motivational features in addition to the 2 previous distractors; Figure 5).

Figure 5. The four experimental conditions: the unmodified task on a black background; the task performed atop an irrelevant, colorful background; the task performed in 3D space; and the task with both background and depth distractors plus an interactive scoreboard and user feedback. Horizontal dimensions of core task limited to a 20º visual angle.



The image used in the *background distractor* condition was a cartoon forest scene obtained from the same provider as the stimulus images. The colors, detail level, and visual style were similar to those of the stimuli; however, there was no other obvious contextual connection. The game condition's dynamic features comprised a scoreboard and real-time performance feedback. The performance feedback was implemented as follows: an incorrect user response caused the selected answer choice to briefly shake back and forth to indicate *no*, whereas a correct choice caused the item to gently pulse outward toward the user. These animations lasted exactly 200 milliseconds. In addition, a scoreboard at the top of the display indicated the current accuracy rate and total score for the current trial set.

All experimental conditions were repeated twice in a randomized order for a total of 8 sets per participant. Each set contained 50 trials and lasted approximately 60 seconds. A 30-second break (black screen; no visual stimulus) was imposed between the training and baseline phases. This was done to prevent contamination of the baseline EEG data by lingering arousal from training. Between each set of trials, there were additional 10-second rest breaks.

The task VA for all conditions was set at 20º, corresponding to the outer edges of the answer choices, measured horizontally. The VA was calculated using the following standard formula:

$$VA = (S \times 57.29) / D \quad \textbf{(1)}$$

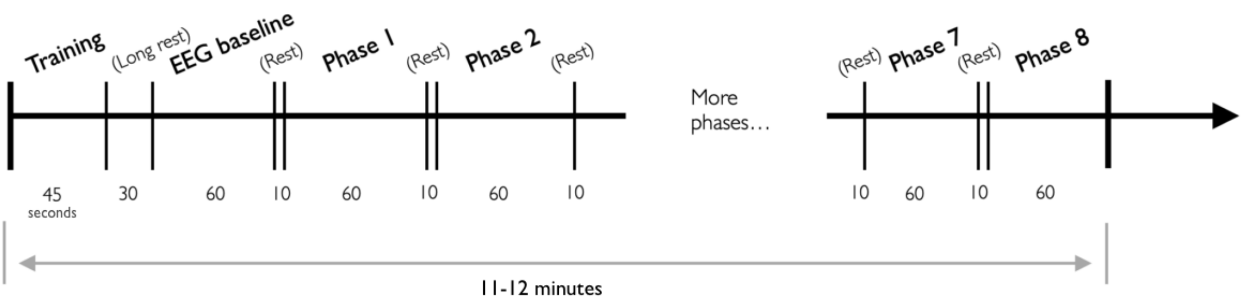
Here, *S* is the size of the object, and *D* is the distance from the observer.

This VA was shown in a previous experiment to be optimal for maximizing the task training performance [33]. With the exception of the *3D depth distractor* and *game distractor* conditions, all visual task elements were precisely placed at a virtual distance of 2 m from the user, as viewed within the HMD. In the conditions that made use of 3D depth, the answer choices (and colorful background) remained at the same virtual distance of 2 m; however, the primary central stimulus moved forward to appear at a distance of 1 m from the user. In the Unity 3D programming environment, 1 unit of space is equivalent to 1 perceived meter of distance. To set the VA for each experimental condition, we specified the desired VA and solved the abovementioned equation for *S*. The value of *S* was applied to the visual task automatically by the software for each new experimental condition before the presentation of the first task trial.

Body movements, particularly eye movements, have a high possibility of introducing movement artifacts into the EEG data. Therefore, participants were instructed to blink and adjust their posture as needed during rest breaks but to refrain from doing so during the trial sets themselves.

Visual text messages on display announced the beginning and end of these break periods. The latter message flashed off 2 seconds before the start of the following set. The total time required to complete each set of trials varied according to participant ability (as dictated by the rules of the adaptive task) but lasted approximately 60 (SD 7.49) seconds on average. This resulted in an overall experimental protocol duration of 11 to 12 minutes (Figure 6).

Figure 6. Protocol flow: following training and electroencephalograph (EEG) baseline recording, 8 experimental phases, each containing 50 trials, were conducted. A 10-second rest separated each experimental phase. The content of the experimental phases was randomly selected from the 4 condition types (unmodified task, background distractor, 3D depth, and game distractor) and balanced so that each participant experienced each condition twice. Unless otherwise noted, all times are in seconds; completion times for training and experimental phases are approximate. Total time to complete the protocol varied from 11 to 12 minutes per participant.



Training

Before the start of the protocol, the task rules were explained, and each participant was granted time to practice the task until they were able to achieve a 75% average accuracy rate for at least 10 trials. Some participants mastered the task more quickly than others, such that the training period lasted between 30 and 90 seconds, with an average of 44 (SD 17) seconds. As the adaptive mechanism was also engaged during the training period, the training process also served to establish the starting difficulty for the participant for the following experimental trial sets.

EEG Data

EEG signals (microvolts) were acquired from the frontal, central, occipital, and parietal regions using a wireless 8-channel EEG amplifier (OpenBCI 32-bit Board Kit, OpenBCI, Inc) with a sampling rate of 250 Hz. The electrode locations were Fz, Cz, Oz, and Pz, placed according to the international 10 to 20 system, and were specifically selected to capture a broad range of activity along the midline. In particular, we were interested in electrode positions Fz and Cz because of their frequently cited relationship with concentration and cognitive load, whereas Oz and Pz were chosen because of their proximity to the visual cortex and prior association with both attention and complex visual decoding [40-43]. Gold cup electrodes were attached to the scalp and ear lobes using an electroconductive gel, and an initial impedance of <5 kΩ across all electrode positions was ensured. Additional electrodes were affixed above and below the participants' eyes to record electrooculogram signals caused by blinking or other facial movements for later use in noise reduction and signal optimization [44].

EEG data were recorded throughout the experiment, although only the final 30 seconds of activity were analyzed for each phase. This was to ensure that the task adaptation algorithm had been given sufficient time to adjust the difficulty levels for each participant before reaching the analysis time window. Time markers for determining the analysis epochs were embedded in the EEG data stream directly using real-time network packets generated by the experimental task. Through the use of this mechanism, we hoped to precisely measure similar levels of cognitive engagement for each participant.

Task Performance

Overall reaction time and task accuracy were calculated for each phase and averaged across all trials for a given experimental condition.

Analysis Method

The software used for EEG data preprocessing and analysis was MATLAB R2019b (MathWorks, Inc). The raw EEG data were notch filtered (50 Hz) and high-pass filtered at 4 Hz using built-in Butterworth and bandpass filters in MATLAB. As noted earlier, the electrooculogram data were recorded in tandem with the EEG for each participant. This enabled us to create customized artifact recognition routines that were individually applied during the data preprocessing phase for each participant. Additional muscle artifacts identified from a visual inspection of the EEG data plots were also removed in full from the time series before analysis.

Fast Fourier transforms were calculated for the following spectral ranges: theta (4-8 Hz), alpha (8-13 Hz), low beta (13-20 Hz), and high beta (20-28 Hz), with 30-second windows for each phase of the experiment. The total sum of the power values from each range was divided by the total number of EEG data samples. The resulting score was normalized by subtracting the overall population mean (combined EEG data of all participants divided by the number of participants) and dividing by the SD to obtain the power index. Fast Fourier transforms and statistical analyses were performed using built-in MATLAB functions.

Shapiro-Wilk tests showed that we could not necessarily operate under an assumption of normally distributed data. Therefore, statistical significance was determined with a repeated-measures ANOVA, followed by a nonparametric Wilcoxon signed-rank test to determine the significance of any changes in power between the experimental phases. The Wilcoxon test was chosen because of the large individual differences in performance observed among participants, nonnormally distributed data, and the within-subjects nature of the study.

Task performance data were averaged to obtain an overall accuracy and reaction time value for each participant per task condition. Individual results were averaged, and similar Wilcoxon signed-rank tests were conducted.

When looking at the preliminary data, it became quickly apparent that the performance levels varied significantly from participant to participant. Some individuals were able to complete the task quickly and accurately, whereas others struggled to respond and made frequent mistakes. This contributed to a large SD in the overall results, which could potentially complicate the drawing of meaningful conclusions. To address this, participants were additionally subclassified into *high*- and *low*-performance groups for further analysis. The selection criteria were based on the average overall task difficulty level achieved by each participant.

Ethics Approval

The experimental protocol was approved by the ethics board of the Tokyo Institute of Technology (2019059).

Results

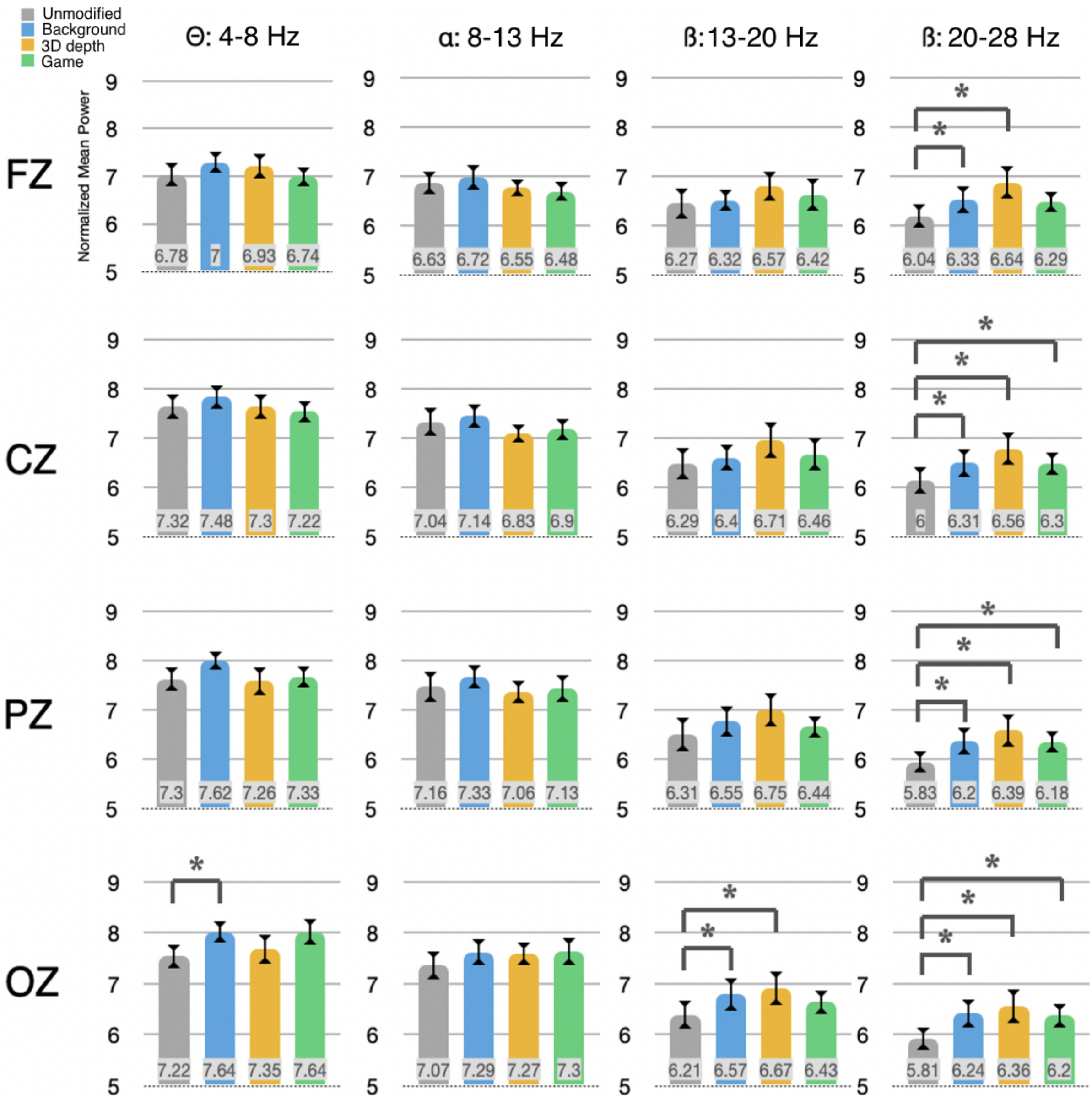
EEG Data

The presence of gamified visual features led to observable changes in the spectral power at all EEG locations. In particular, the occipital and parietal areas showed noticeable increases in beta EEG power for the 3D depth distractor condition and in the theta rhythm during the background distractor condition. Overall, 1-way repeated-measures ANOVA showed significant differences in the high-beta range for all electrodes tested (Fz: $F_{3,76}$=3.75, P=.02; Cz: $F_{3,76}$=4.09, P=.01; Pz: $F_{3,76}$=2.82, P=.046; Oz: $F_{3,76}$=2.97, P=.04). Post hoc Wilcoxon signed-rank tests revealed that with the exception of the *game* condition at Fz, all individual increases in the high-beta rhythm between the unmodified and experimental conditions were significant at the 5% confidence level. However, the differences between individual experimental conditions were not significant.

In contrast, for the theta range, only the results at Oz displayed significant variation ($F_{3,76}$=3.20; P=.03), and only one individual experimental condition, the background distractor, proved to be significant (n=20; Z=−2.81; P=.00495) in the post hoc analysis. Changes in the alpha rhythm did not prove to be significant at any electrode position (Figure 7).

It is noteworthy that the game condition, which also included the 3D depth distractor, did not reach the same levels of cognitive activity as the depth-only condition for the beta range. This may indicate that the presence of additional distractions in the game condition inhibits the overall impact of the 3D depth effect. However, in the theta range, the presence of background distraction in both the background and game conditions led to similar cognitive responses.

Figure 7. Spectral power by condition and frequency at electroencephalograph locations Fz, Cz, Pz, and Oz; n=20, SE; statistical significance calculated with Wilcoxon signed-rank test (*P<.05).



Performance Data

The 1-way repeated-measures ANOVA comparing the 4 conditions showed no significance for either task speed ($F_{3,72}$=1.21; P=.31) or accuracy ($F_{3,72}$=0.143; P=.93). In general, the presence of colorful, task-irrelevant backgrounds led to slight reductions in accuracy but had little impact on performance speed. Conversely, the presence of 3D depth cues seems to have slightly affected reaction time but not accuracy (Figure 8).

As noted previously, we took the additional step of separating participants into high- and low-performance groups according to ability (average maximum task difficulty achieved during all trial sets) as a supplemental analysis. This was because of a large SE observed in the performance data, which we felt had the potential to mask underlying trends. Although the resulting subgroups were too small to deliver meaningful statistical power, the results revealed several nuances and presented a potentially interesting direction for a follow-up investigation.

For task accuracy, the additional visual distractions present in the multiple-distraction game condition appear to have had a cumulative negative impact on high performers. However, a seemingly opposite effect was observed in the low-performance group, which cumulatively achieved the highest accuracy in this condition.

In terms of task completion speed, our results did not show any significant differences between conditions, even when observing only the more internally homogenous high-performance subgroup (Figure 9).

Figure 8. Mean task performance (reaction time and accuracy) by condition and group; SE; statistical significance calculated with Wilcoxon signed-rank test (*P*<.05).

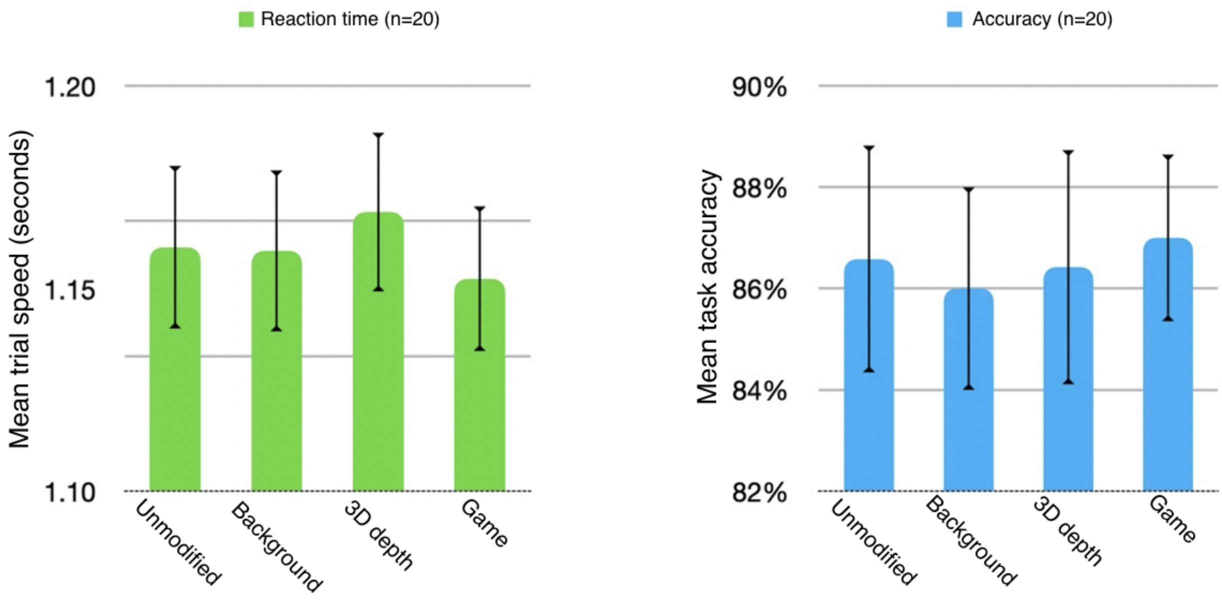
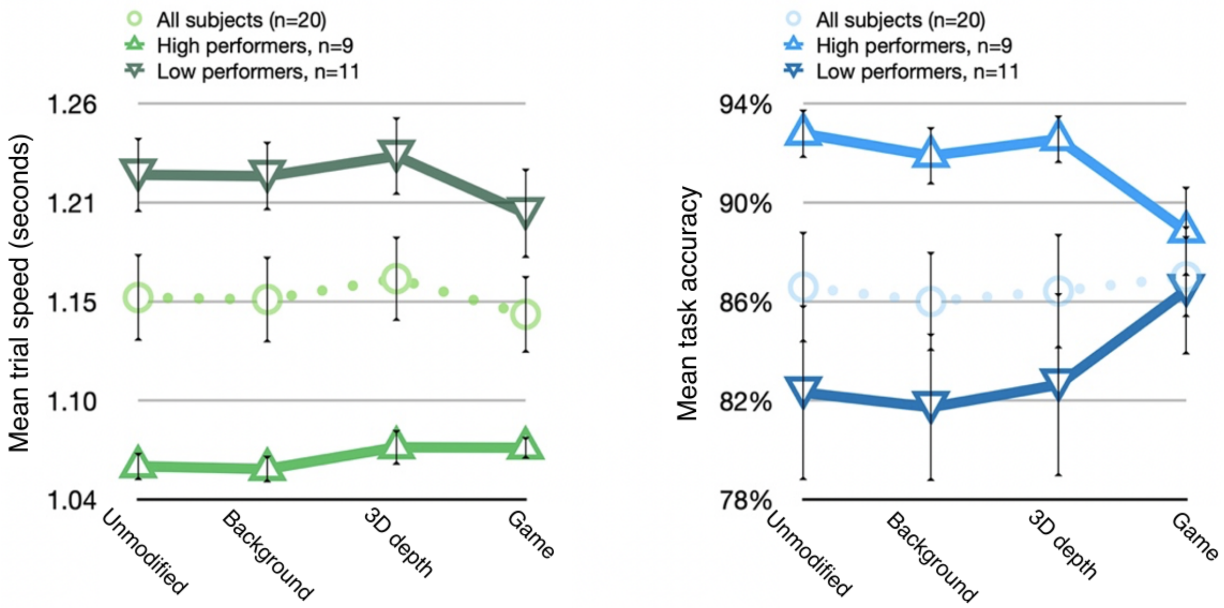


Figure 9. Mean task performance (reaction time and accuracy) by condition and high- and low-performance subgroup; SE.



Performance and EEG Compared

Perhaps because of a lack of significant differences in performance between experimental conditions, regressing EEG spectral power onto performance results produced no meaningful correlations for either the overall group or either of the subgroups. Large individual differences in participant performance likely also contributed to the lack of significant results.

Discussion

Principal Findings

The objective of this study was to examine the impact of visual, game-like elements on task performance and cognitive activity in a visual working memory task. No significant differences in performance could be determined for both reaction time and task accuracy. Nevertheless, certain performance trends can be observed that seem to leave open the possibility that specific types of visual distractions may affect some aspects of cognitive performance while leaving others unaffected. For example, our data show that visually distracting backgrounds had no observable impact on reaction speed but had a slight impact on accuracy. Conversely, 3D depth decoding appears to have slightly affected the speed of processing but not the task accuracy.

Similarly, the EEG power analysis revealed no significant differences in the crucial frontal theta rhythm at Fz, which often serves as a proxy for participant concentration and task engagement [41,43]. In contrast, significant differences between conditions were observed in the beta band and theta band at the occipital electrode. Although these results return to insignificance if one corrects for multiple comparisons using Bonferroni or a similar method, the question nevertheless arises as to what might have caused these observed effects in the beta and theta rhythms, particularly given the lack of correlation with performance. For instance, the higher theta power observed at Oz was actually accompanied by slightly *reduced* accuracy in the background distractor condition. The proximity of Oz to the occipital area and the visual cortex suggests that rather than being directly linked with cognitive effort related to the task, perhaps the theta rhythm is simply more sensitive to certain underlying ocular processes required by the visually rich background used in this condition.

For example, although the current experimental task is designed to prohibit voluntary eye movements by requiring the participant to continually focus on a center stimulus, the presence and frequency of *involuntary* eye movements such as saccades were unfortunately not recorded as part of the current experimental design. Indeed, evidence suggests that saccades may be highly correlated with theta power during periods of memory encoding [45]. Other studies have similarly observed links between increased cognitive stress related to memory tasks and elevated saccadic frequency and duration [46,47]. Thus, the possibility that the background condition may have elicited a disproportionate amount of ocular activity and, along with it, increased theta power represents one hypothesis for the observed results.

At the same time, increased high-beta (20-28 Hz) spectral power in the 3D depth condition was accompanied by generally slower reaction times. Although previous research has implicated beta rhythm in a variety of assistive roles with regard to visual perception [48], studies that specifically examine 3D decoding are less conclusive. For example, although some researchers found that 3D environments elicited greater cognitive activity than their 2D counterparts, particularly in the beta range [49], Dan et al [50] found a *reduction* in EEG power during the 3D condition versus the 2D condition in their experiment involving a learning task [50]. However, the latter study involved complex *reality-like* visuals, focused on the Fz theta/Pz alpha ratio for EEG feature classification rather than a broad-spectrum analysis, and did not specifically target the beta range. Therefore, the possibility remains that, as with the theta band, underlying cognitive demands related to visual processes may have obscured task-related cognitive activity. As noted earlier, the cognitive task used incorporates several cognitive processes, including visual working memory and divided attention. This multimodality presents a further challenge when trying to determine the exact reason for unexpected EEG results, as it is difficult to ascertain the cognitive process responsible for the observed effects.

The supplementary analysis of performance by participant ability, although not statistically meaningful, nevertheless revealed an unexpected trend with regard to task accuracy. The performance results from the high group appeared to be cumulatively reduced by successive layers of distractions, with the game condition eliciting the lowest average accuracy levels. The poorer performers paradoxically appeared to perform best during this condition. However, it must be noted that the average degree of accuracy obtained in the *low* group was still well below that of the average overall performance from the *high* group.

We offer two hypotheses: throughout the experiment, the low-performance group may have experienced a form of performance anxiety that led to generally slower decision-making and lower overall accuracy. However, the presence of multiple additional visual elements in the game condition may have provided a certain degree of reassurance and encouragement, an effect of gamified design documented by previous researchers [14,25]. Similarly, the inclusion of a scoreboard and positive and negative response feedback after every trial in the game condition may have helped to refocus participant attention and encourage less experienced or more easily distracted participants to improve their performance.

Finally, it is worth noting the limitations of the current results. First, as the context of this study was potential users of commercial cognitive training products, we used broadly inclusive criteria for participant selection, which resulted in a wide range of ages and an uneven gender balance. This may have affected the results in unexpected ways. Second, although all experimental conditions differed significantly from the unmodified task in the high-beta range (except for the game condition at Fz), they did not differ significantly from each other. This lack of precision reinforces the possibility that any visual novelty, whether it is the presence of 3D depth or a colorful background, triggers an increased cognitive response in the high-beta range. Greater EEG channel density and separating the multimodal task into its component cognitive processes could potentially help isolate and differentiate the observed responses.

Conclusions

In isolation, a small performance impact was incurred by the inclusion of a colorful, task-irrelevant background and the use of 3D depth elements. However, that impact was mitigated or reversed for some participants when combined with *motivating* features such as real-time feedback and scoreboards. Overall, the primary finding of this study is that performance in simple memory tasks of the kind that are frequently found in commercial cognitive training apps is not significantly affected by the use of visually distracting backgrounds or 3D depth or by common motivational game elements such as scoreboards and real-time performance feedback. Particularly in light of the user engagement and motivational advantages of gamification documented by previous researchers, the observed impacts may not be substantial enough to warrant specific design patterns or the redesigning of existing gamified cognitive tasks unless the specific goal is to maximize the speed and accuracy, in which case, the current findings may provide some useful guidance.

Acknowledgments

This publication would not have been possible without the support of Dr Yasuharu Koike at the Tokyo Institute of Technology. The authors would like to express their gratitude for his advice, material support, and assistance with our research. The authors received additional guidance and support from Dr Tohru Yagi and Dr Takako Yoshida, directors of their respective laboratories at the Tokyo Institute of Technology.

Authors' Contributions

ER and BG helped in conceptualization and methodology; ER and YR helped with the software; ER, BG, and YR helped in validation; and ER performed the formal analysis and helped in the investigation. ER, BG, and YR helped with resources; ER and YR performed the data curation; ER prepared the original draft; ER and BG performed review and editing; and ER was involved in the supervision, project administration, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Cognitive Assessment and Training Market by Component, Organization Size, Application (Clinical Trials, Learning, and Research), Vertical (Healthcare and Life Sciences, Education, and Corporate), and Region - Global Forecast to 2025. Markets and Markets. 2020. URL: https://www.marketsandmarkets.com/Market-Reports/cognitive-assessment-market-1039.html [accessed 2020-07-09]
2. Green CS, Bavelier D. Action video game modifies visual selective attention. Nature 2003 May 29;423(6939):534-537. [doi: 10.1038/nature01647] [Medline: 12774121]
3. Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A 2008 May 13;105(19):6829-6833 [FREE Full text] [doi: 10.1073/pnas.0801268105] [Medline: 18443283]
4. Edwards JD, Xu H, Clark DO, Guey LT, Ross LA, Unverzagt FW. Speed of processing training results in lower risk of dementia. Alzheimers Dement (N Y) 2017 Nov;3(4):603-611 [FREE Full text] [doi: 10.1016/j.trci.2017.09.002] [Medline: 29201994]
5. Wolinsky FD, Vander Weg MW, Howren MB, Jones MP, Dotson MM. A randomized controlled trial of cognitive training using a visual speed of processing intervention in middle aged and older adults. PLoS One 2013 May 1;8(5):e61624 [FREE Full text] [doi: 10.1371/journal.pone.0061624] [Medline: 23650501]
6. Turunen M, Hokkanen L, Bäckman L, Stigsdotter-Neely A, Hänninen T, Paajanen T, et al. Computer-based cognitive training for older adults: determinants of adherence. PLoS One 2019 Jul 10;14(7):e0219541 [FREE Full text] [doi: 10.1371/journal.pone.0219541] [Medline: 31291337]
7. Eramudugolla R, Kiely KM, Chopra S, Anstey KJ. Effect of speed of processing training on older driver screening measures. Front Aging Neurosci 2017 Oct 17;9:338 [FREE Full text] [doi: 10.3389/fnagi.2017.00338] [Medline: 29089888]
8. Chi H, Agama E, Prodanoff ZG. Developing serious games to promote cognitive abilities for the elderly. In: Proceedings of the IEEE 5th International Conference on Serious Games and Applications for Health. 2017 Presented at: SeGAH '17; April 2-4, 2017; Perth, Australia p. 1-8. [doi: 10.1109/segah.2017.7939279]
9. Conklin HM, Ashford JM, Clark KN, Martin-Elbahesh K, Hardy KK, Merchant TE, et al. Long-term efficacy of computerized cognitive training among survivors of childhood cancer: a single-blind randomized controlled trial. J Pediatr Psychol 2017 Mar 01;42(2):220-231 [FREE Full text] [doi: 10.1093/jpepsy/jsw057] [Medline: 27342301]
10. Kable JW, Caulfield MK, Falcone M, McConnell M, Bernardo L, Parthasarathi T, et al. No effect of commercial cognitive training on brain activity, choice behavior, or cognitive performance. J Neurosci 2017 Aug 02;37(31):7390-7402 [FREE Full text] [doi: 10.1523/JNEUROSCI.2832-16.2017] [Medline: 28694338]
11. Sala G, Gobet F. Cognitive training does not enhance general cognition. Trends Cogn Sci 2019 Jan;23(1):9-20. [doi: 10.1016/j.tics.2018.10.004] [Medline: 30471868]
12. Stojanoski B, Lyons KM, Pearce AA, Owen AM. Targeted training: converging evidence against the transferable benefits of online brain training on cognitive function. Neuropsychologia 2018 Aug;117:541-550. [doi: 10.1016/j.neuropsychologia.2018.07.013] [Medline: 30009838]
13. Souders DJ, Boot WR, Blocker K, Vitale T, Roque NA, Charness N. Evidence for narrow transfer after short-term cognitive training in older adults. Front Aging Neurosci 2017 Feb 28;9:41 [FREE Full text] [doi: 10.3389/fnagi.2017.00041] [Medline: 28293188]
14. García-Betances RI, Cabrera-Umpiérrez MF, Arredondo MT. Computerized neurocognitive interventions in the context of the brain training controversy. Rev Neurosci 2018 Jan 26;29(1):55-69. [doi: 10.1515/revneuro-2017-0031] [Medline: 28822228]
15. Edwards JD, Fausto BA, Tetlow AM, Corona RT, Valdés EG. Systematic review and meta-analyses of useful field of view cognitive training. Neurosci Biobehav Rev 2018 Jan;84:72-91. [doi: 10.1016/j.neubiorev.2017.11.004] [Medline: 29175362]

16. Linares R, Borella E, Lechuga MT, Carretti B, Pelegrina S. Nearest transfer effects of working memory training: a comparison of two programs focused on working memory updating. PLoS One 2019 Feb 13;14(2):e0211321 [FREE Full text] [doi: 10.1371/journal.pone.0211321] [Medline: 30759135]
17. Ross LA, Webb CE, Whitaker C, Hicks JM, Schmidt EL, Samimy S, et al. The effects of useful field of view training on brain activity and connectivity. J Gerontol B Psychol Sci Soc Sci 2019 Sep 15;74(7):1152-1162 [FREE Full text] [doi: 10.1093/geronb/gby041] [Medline: 29757433]
18. Grissmann S, Faller J, Scharinger C, Spüler M, Gerjets P. Electroencephalography based analysis of working memory load and affective valence in an N-back task with emotional stimuli. Front Hum Neurosci 2017 Dec 19;11:616 [FREE Full text] [doi: 10.3389/fnhum.2017.00616] [Medline: 29311875]
19. Mühl C, Jeunet C, Lotte F. EEG-based workload estimation across affective contexts. Front Neurosci 2014 Jun 12;8:114 [FREE Full text] [doi: 10.3389/fnins.2014.00114] [Medline: 24971046]
20. Lumosity for mobile. Lumos Labs. 2021. URL: https://www.lumosity.com/en/ [accessed 2021-08-06]
21. Brain Training that Works. BrainHQ. 2020. URL: https://www.brainhq.com/ [accessed 2020-07-09]
22. 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]
23. Boendermaker WJ, Veltkamp RC, Peeters M. Training behavioral control in adolescents using a serious game. Games Health J 2017 Dec;6(6):351-357. [doi: 10.1089/g4h.2017.0071] [Medline: 28926286]
24. Mohammed S, Flores L, Deveau J, Hoffing RC, Phung C, Parlett CM, et al. The benefits and challenges of implementing motivational features to boost cognitive training outcome. J Cogn Enhanc 2017 Dec;1(4):491-507 [FREE Full text] [doi: 10.1007/s41465-017-0047-y] [Medline: 30221244]
25. 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 1st International Conference on Gameful Design, Research, and Applications. 2013 Oct Presented at: Gamification '13; October 2-4, 2013; Toronto, Canada p. 66-73. [doi: 10.1145/2583008.2583017]
26. 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]
27. Katz B, Jaeggi S, Buschkuehl M, Stegman A, Shah P. Differential effect of motivational features on training improvements in school-based cognitive training. Front Hum Neurosci 2014 Apr 24;8:242 [FREE Full text] [doi: 10.3389/fnhum.2014.00242] [Medline: 24795603]
28. Landers RN, Landers AK. An empirical test of the theory of gamified learning: the effect of Leaderboards on time-on-task and academic performance. Simul Gaming 2015 Jan 26;45(6):769-785. [doi: 10.1177/1046878114563662]
29. 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]
30. Ninaus M, Pereira G, Stefitz R, Prada R, Paiva A, Neuper C, et al. Game elements improve performance in a working memory training task. Int J Serious Games 2015 Feb 10;2(1):3-16. [doi: 10.17083/ijsg.v2i1.60]
31. 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]
32. Raggi A, Tasca D, Ferri R. A brief essay on non-pharmacological treatment of Alzheimer's disease. Rev Neurosci 2017 Jul 26;28(6):587-597. [doi: 10.1515/revneuro-2017-0002] [Medline: 28422708]
33. Redlinger E, Glas B, Rong Y. Impact of screen size on cognitive training task performance: an HMD study. Int J Psychophysiol 2021 Aug;166:166-173. [doi: 10.1016/j.ijpsycho.2021.06.003] [Medline: 34119616]
34. Smith PL, Little DR. Small is beautiful: in defense of the small-N design. Psychon Bull Rev 2018 Dec;25(6):2083-2101 [FREE Full text] [doi: 10.3758/s13423-018-1451-8] [Medline: 29557067]
35. Normand MP. Less is more: psychologists can learn more by studying fewer people. Front Psychol 2016 Jun 17;7:934 [FREE Full text] [doi: 10.3389/fpsyg.2016.00934] [Medline: 27379004]
36. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
37. 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]
38. Ball K, Edwards JD, Ross LA. The impact of speed of processing training on cognitive and everyday functions. J Gerontol B Psychol Sci Soc Sci 2007 Jun;62 Spec No 1:19-31. [doi: 10.1093/geronb/62.special_issue_1.19] [Medline: 17565162]
39. Kirchner WK. Age differences in short-term retention of rapidly changing information. J Exp Psychol 1958 Apr;55(4):352-358. [doi: 10.1037/h0043688] [Medline: 13539317]

40. Gevins A, Smith ME, McEvoy L, Yu D. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex 1997 Jun;7(4):374-385. [doi: 10.1093/cercor/7.4.374] [Medline: 9177767]
41. Yamada F. Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. Ergonomics 1998 May;41(5):678-688. [doi: 10.1080/001401398186847] [Medline: 9613228]
42. Sauseng P, Klimesch W, Stadler W, Schabus M, Doppelmayr M, Hanslmayr S, et al. A shift of visual spatial attention is selectively associated with human EEG alpha activity. Eur J Neurosci 2005 Dec;22(11):2917-2926. [doi: 10.1111/j.1460-9568.2005.04482.x] [Medline: 16324126]
43. Ewing KC, Fairclough SH, Gilleade K. Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a Biocybernetic loop. Front Hum Neurosci 2016 May 18;10:223. [doi: 10.3389/fnhum.2016.00223] [Medline: 27242486]
44. Zahan S. Removing EOG artifacts from EEG signal using noise-assisted multivariate empirical mode decomposition. In: Proceedings of the 2nd International Conference on Electrical, Computer & Telecommunication Engineering. 2017 Presented at: ICECTE '16; December 8-10, 2016; Rajshahi, Bangladesh p. 1-5. [doi: 10.1109/icecte.2016.7879634]
45. Sato N, Yamaguchi Y. Computational model-based prediction of human episodic memory performance based on eye movements. IEICE Trans Commun 2008 Jul 01;E91-B(7):2142-2143. [doi: 10.1093/ietcom/e91-b.7.2142]
46. Zargari Marandi R, Madeleine P, Omland Ø, Vuillerme N, Samani A. Reliability of oculometrics during a mentally demanding task in young and old adults. IEEE Access 2018 Mar 26;6:17500-17517. [doi: 10.1109/access.2018.2819211]
47. Di Stasi LL, McCamy MB, Catena A, Macknik SL, Cañas JJ, Martinez-Conde S. Microsaccade and drift dynamics reflect mental fatigue. Eur J Neurosci 2013 Aug;38(3):2389-2398. [doi: 10.1111/ejn.12248] [Medline: 23675850]
48. Kloosterman NA, Meindertsma T, Hillebrand A, van Dijk BW, Lamme VA, Donner TH. Top-down modulation in human visual cortex predicts the stability of a perceptual illusion. J Neurophysiol 2015 Feb 15;113(4):1063-1076 [FREE Full text] [doi: 10.1152/jn.00338.2014] [Medline: 25411458]
49. Bilgin P, Agres K, Robinson N, Wai AA, Guan C. A comparative study of mental states in 2D and 3D virtual environments using EEG. In: Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics. 2019 Presented at: SMC '19; October 6-9, 2019; Bari, Italy p. 2833-2838. [doi: 10.1109/smc.2019.8914326]
50. Dan A, Reiner M. EEG-based cognitive load of processing events in 3D virtual worlds is lower than processing events in 2D displays. Int J Psychophysiol 2017 Dec;122:75-84. [doi: 10.1016/j.ijpsycho.2016.08.013] [Medline: 27592084]

Abbreviations

ACTIVE: Advanced Cognitive Training for Independent and Vital Elderly
EEG: electroencephalograph
HMD: head-mounted display
UFoV: Useful Field of View
VA: visual angle
VR: virtual reality

Edited by N Zary; submitted 01.12.21; peer-reviewed by V Verma, P Wang, AM Bezabih, L Martinez; comments to author 05.03.22; revised version received 08.03.22; accepted 25.03.22; published 28.04.22.

Please cite as:
Redlinger E, Glas B, Rong Y
Impact of Visual Game-Like Features on Cognitive Performance in a Virtual Reality Working Memory Task: Within-Subjects Experiment
JMIR Serious Games 2022;10(2):e35295
URL: https://games.jmir.org/2022/2/e35295
doi:10.2196/35295
PMID:35482373

Original Paper

Interactive Digital Game for Improving Visual–Perceptual Defects in Children With a Developmental Disability: Randomized Controlled Trial

Wen-Lan Wu[1,2], PhD; Yu-Ling Huang[1], MS; Jing-Min Liang[1], MS; Chia-Hsin Chen[3,4,5], MD, PhD; Chih-Chung Wang[3], MS; Wen-Hsien Ho[2,6,7], PhD

[1]Department of Sports Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
[2]Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
[3]Department of Physical Medicine and Rehabilitation, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
[4]Department of Physical Medicine and Rehabilitation, School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
[5]Regenerative Medicine and Cell Therapy Research Center, Kaohsiung Medical University, Kaohsiung, Taiwan
[6]Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung, Taiwan
[7]Department of Mechanical Engineering, National Pingtung University of Science and Technology, Pingtung, Taiwan

Corresponding Author:
Wen-Hsien Ho, PhD
Department of Healthcare Administration and Medical Informatics
Kaohsiung Medical University
No.100, Shih-Chuan 1st Road
Kaohsiung, 807
Taiwan
Phone: 886 7 3121101 ext 2648
Email: whho@kmu.edu.tw

Abstract

Background: Visual–perceptual defects in children can negatively affect their ability to perform activities of daily living. Conventional rehabilitation training for correcting visual–perceptual defects has limited training patterns and limited interactivity, which makes motivation difficult to sustain.

Objective: We aimed to develop and evaluate an interactive digital game system for correcting visual–perceptual defects and evaluate its effectiveness.

Methods: Participants were children aged 5 to 10 years with a diagnosis of visual–perceptual defect associated with a developmental disability. The children were randomized into a digital game group who received the traditional course of rehabilitation combined with an interactive digital game intervention (n=12) and a standard rehabilitation group (n=11) who only received the traditional course of rehabilitation. Each group underwent rehabilitation once a week for 4 weeks. Overall improvement in Test of Visual Perceptual Skills 3rd edition (TVPS-3) score and overall improvement in performance in the interactive digital game were evaluated. Parents and therapists were asked to complete a satisfaction questionnaire.

Results: After 4 weeks, the TVPS-3 score had significantly increased (P=.002) in the digital game group (pre: mean 41.67, SD 13.88; post: 61.50, SD 21.64). In the standard rehabilitation group, the TVPS-3 score also increased, but the increase was not statistically significant (P=.58). Additionally, TVPS-3 score increases were significantly larger for the digital game group compared with those for the standard rehabilitation group (P=.005). Moreover, both parents and therapists were highly satisfied with the system. All 5 themes of satisfaction had mean scores higher than 4 in a 5-point scale questionnaire (mean 4.30, SD 0.56).

Conclusions: The system has potential applications for improving visual–perceptual function in children undergoing medical rehabilitation for developmental disability.

Trial Registration: ClinicalTrials.gov NCT05016492; http://clinicaltrials.gov/ct2/show/NCT05016492

(JMIR Serious Games 2022;10(2):e34756) doi:10.2196/34756

KEYWORDS

interactive digital game; visual–perceptual defect; developmental disability; Children; Test of Visual Perceptual Skills; rehabilitation

Introduction

Children with developmental disabilities often have difficulty performing activities that require well-developed visual–perceptual skills or often perform them inefficiently or with poor form. Such activities include activities of daily living that require gross motor skills or fine motor skills [1]. Children with developmental disabilities are usually referred for professional evaluation when their skill level is 20% lower than that considered to be normal for their age [2]. Parents can usually estimate the baseline ability and learning speed of children with developmental disabilities by referring to well-established childhood development standards [3,4]. Beginning therapy when the child is young—at an age when neuroplasticity is still high—can reduce future medical, educational, familial, economic, and social burdens of a developmental disability. Reducing these burdens can then enhance the dignity, adaptability, personal functional performance, and quality of life of the child and help to avoid development of emotional and psychological issues. Therefore, beginning therapy early after diagnosis of a developmental abnormality is vitally important [5].

Developmental disabilities in children are usually complicated by visual–perceptual problems. Although these problems can be improved by stimulation and training, conventional methods of visual–perceptual training have limited effectiveness because their training patterns lack extensibility, which limits the motivation, goal orientation, and sense of achievement in the learner.

The numerous scientific and technological advances in computer and digital learning technologies that have successively emerged in recent decades are currently driving broad applications of computer technology in education and therapy as well as changes in the patterns of education delivery, especially for students with special needs. Thus, researchers have begun to investigate the extent to which interactive digital games can replace traditional paper-based teaching modes.

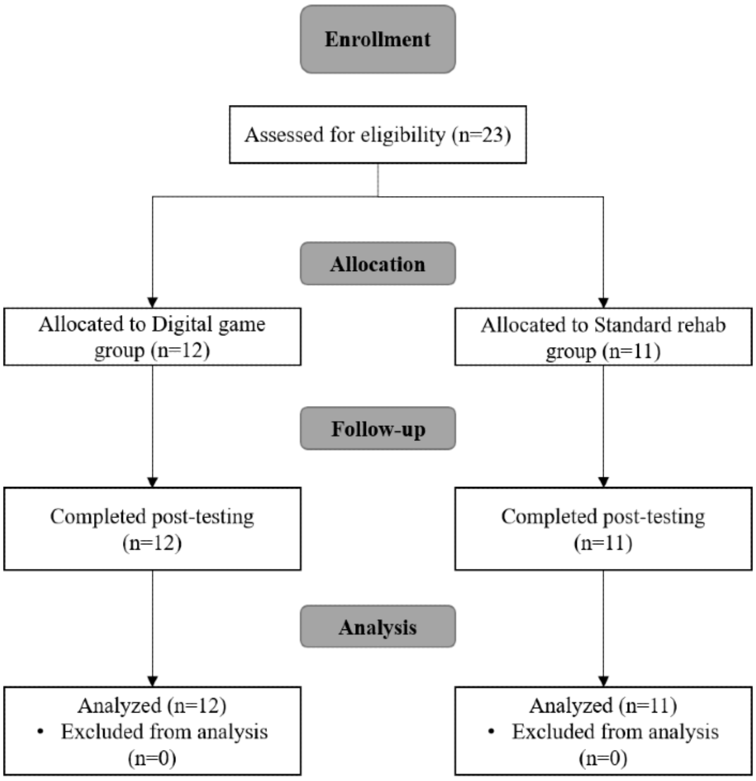
Some research suggests that digital game designs must provide an enjoyable experience to attract users and maintain their engagement. Clinical use of multimedia for early rehabilitation and therapy for children is now very common. For example, a multimedia game used for rehabilitation may depict a skill training item as an animation of a fictitious character that provides interactive feedback, for example, dynamic acoustic-optic sounds or static images, to introduce learning concepts, provide sensory stimulation, and guide learning. The emergence of interactive digital games indeed provides clinicians with an attractive option for rehabilitation that can effectively arouse a sense of participation in the learner and sustain the interest of the learner [6-10]. Thus, to motivate learners and to stimulate their interest in the content, positive feedback and increasingly difficult skill levels should be essential elements of a multimedia digital game used as treatment to improve developmental disability. Researchers have identified therapeutic effects for multimedia digital games designed for children—improving learning in mathematics [9], find motor skill [8], or visual perception abilities [10-13] in children with autism [6], poor academic performance [7], developmental delay [8], intelligence disorder [9], attention deficit hyperactivity disorder [10], and developmental disabilities [11-13]. One of these studies [11] used the Test of Visual Perceptual Skills (TVPS) to evaluate the effectiveness of multimedia digital games for improving developmental disability but only investigated changes in a limited selection of TVPS items. Another study [12] recently reported the use of kinesthetic games for improving visual–perceptual skills in children with delayed development of these skills [13]; although the novel design of the kinesthetic games included the integration of gesture-based interaction in a digital game environment, the kinesthetic games were mainly designed to provide training in integrating visual and motor skills. Thus, Wuang et al [13] expected that the functional improvements achieved by training on the kinesthetic games would be jointly contributed by improvements in both visual attention and motor control rather than by improvements in cognitive skills alone. In order to improve cognitive skills alone, we aimed to design a digital game that provides cognitive training in the full range of TVPS items and to evaluate its feasibility for testing visual–perceptual skills.

Methods

Participants

This was a prospective randomized, parallel, single-center clinical study. We recruited 23 schoolchildren with a diagnosis of visual–perceptual defect associated with a developmental disability, aged 4 to 10 years, from the pediatric rehabilitation division of a hospital in southern Taiwan (Figure 1). The enrollment criteria for this study were (1) record of developmental disability diagnosis, (2) ability to understand instructions, (3) TVPS-3 score lower than 25% of the norm reference and diagnosis of visual–perceptual defect, and (4) Test of Nonverbal Intelligence third edition score higher than 70. Participants were excluded if they (1) did not follow or did not understand the instructions for participating in the study, or if they (2) had severe defects in vision or hearing. The standard rehabilitation group (n=11) received a standard 4-week course of rehabilitation delivered in one 30-minute session per week. The digital game group (n=12) received the standard 4-week course of rehabilitation but with an additional 30-minute interactive digital game training session per week. Signed assent was obtained from all participants and signed consent was obtained from at least one parent prior to the start of the study.

Figure 1. Study flowchart.



Ethics Approval

This study was approved by the institutional review board of Kaohsiung Medical University Chung-Ho Memorial Hospital (KMUHIRB-SVII-20150090).

Visual–Perceptual Interactive Game System

The system was developed with Visual Studio (version 2015; Microsoft Inc). The system allows the training skill level difficulty to be adjusted to sustain the interest of the user.

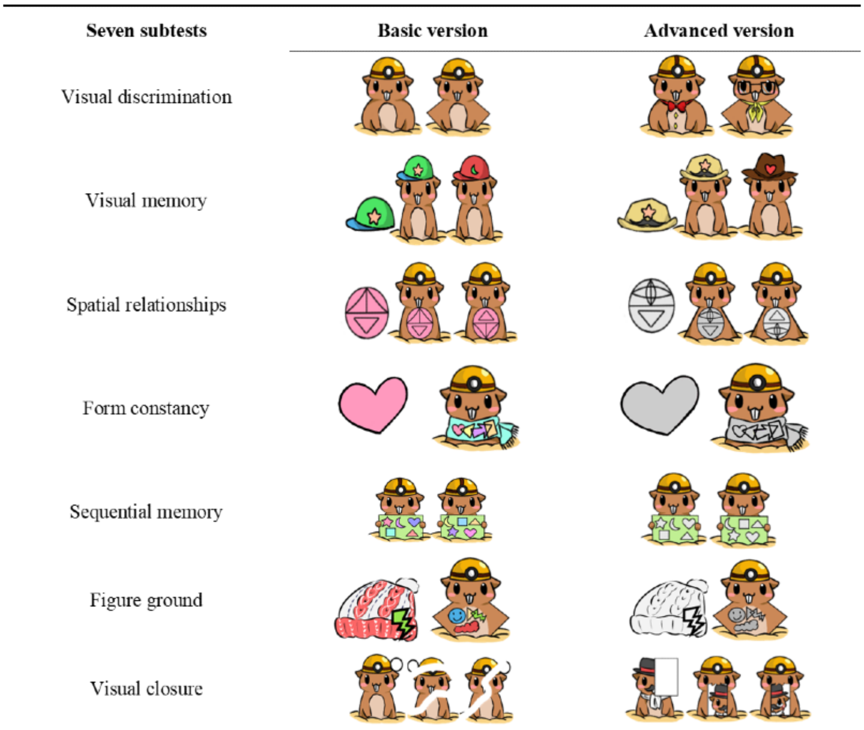
When the game begins (Figure 2), the child is asked to hold the Bluetooth ball with the dominant hand and touch the starting position on the screen. A response signal (animation of mole) appears in the lower part of the screen, and the task is to find a mole with similar characteristics at the top of the screen. The objective of the game is to complete each task as quickly as possible, and the computer displays the average response time required to complete each task. When the game ends, the results are presented to enhance learning motivation and satisfaction in the user.

After training, the responses and response times were stored in the database. Each child was required to perform 7 visual perceptual skills learning tasks, each of which was named to correspond to a TVPS-3 subscale. The game had 2 levels for each skill: basic and advanced (Figure 3). The graphics for the basic version were in color, while those for the advanced version were in black and white. The number of distractors in the advanced version was also greater than that in the basic version. Each skill level was further subdivided into primary, intermediate, and advanced. In the primary version, there was no time limit for answering each matching question or memory question. In the intermediate version, the questions were answered within a certain time limit—3 minutes for each of the questions. In the advanced version, error rate limitation was used, that is, the game ended when the number of errors exceeded 3.

The system also had a testing mode. The test difficulty was the same as the practice in the basic level of the training mode. If the user selected the testing mode, the system automatically started a sequence of 7 subtests. Each subtest had 10 questions. In accordance with the 7-item rules of the TVPS-3, 1 point was assigned to each question, for a score of 10 points for each subtest, and a total score of 70. After the test began, the user continuously answered the questions from the first test, and the test stopped if a wrong answer was given to any of the questions more than 3 times and jumped to the next test. The test scores were recorded for further analysis.

Figure 2. Visual–perceptual interactive digital game system.

Figure 3. Graphics used in 7 subtests: basic version and advanced version of interactive digital game system.



Traditional Paper-Based TVPS-3

The TVPS is a widely used tool for clinical evaluation of visual perception [14]. The tool, which is applicable for children aged 4 to 12 years, has 7 subtests for visual discrimination, visual memory, spatial relationships, form constancy, sequential memory, figure-ground, and visual closure scoring. Each of the 7 subtests has 16 items of varying difficulty. Therefore, the test has 112 total questions. A correct response is scored as 1, and an incorrect response is scored as 0. A ceiling is reached after 3 consecutive incorrect responses on a subtest, the score recorded for the subtest is the number of questions answered correctly on the subtest so far. Raw scores are tallied for each subtest and recorded on a score sheet. Raw scores are reported as scaled scores and percentile ranks for each of the 7 subtests. The total score is reported as a percentile rank and standard score. Age-equivalent scores are provided for subtest and total scores [15].

User Satisfaction Survey

After the experimental group had completed 4 weeks of training, a satisfaction questionnaire was administered to the parents and therapists of the children with developmental disabilities. The 20-item questionnaire content was divided into 5 themes: perceived ease of use (6 questions), perceived usefulness (4 questions), perceived joyfulness (4 questions), satisfaction (3 questions), and continued use (3 questions). We referred to the literature when designing questionnaire items for perceived ease of use [16], perceived usefulness [17], and perceived joyfulness [18]. Questionnaire items were answered on a Likert scale from

1 (completely disagree) to 5 (completely agree). A high total score indicated high satisfaction with the system. The maximum score was 100 points.

Procedure

The TVPS-3 and interactive digital game test were carried out for the 2 groups before the intervention in the first week and the after the intervention in the fourth week of the training.

Statistical Data Analysis

Basic characteristics and preintervention scores were compared between the 2 groups using 2-tailed independent *t* tests for interval data (age, intelligence score, and test scores) and the chi-square tests for categorical data (gender). Within-group changes in TVPS-3 and the digital game test scores were analyzed using 2-tailed paired *t* tests, and between-group differences of changes in TVPS-3 and the digital game test scores were analyzed using independent *t* tests (2-tailed). A *P* value<.05 indicated a significant difference. The Pearson correlation coefficient was calculated to verify the relevance between the digital game test score and the traditional paper TVPS-3 score. In a correlation coefficient analysis, a low correlation is defined as |*r*|<0.39; a medium correlation is defined as a |*r*|=0.40-0.69; a high correlation is defined as a |*r*|=0.70-0.99; a full correlation is defined as a |*r*|=1 [19]. The reliability of the questionnaire measurement results was assessed with Cronbach α; Cronbach α>0.70 indicated high reliability, and Cronbach α<0.35 indicated low reliability [20].

Results

Basic Characteristic and Preintervention Data

Children enrolled in this study had clinical diagnoses of developmental disability complicated with attention deficit hyperactivity disorder of the inattentive type (17 children), autistic spectrum disorder (1 child), speech articulation disorder (3 children), emotional disorder (1 child), and behavioral disorder (1 child). The digital game group included 10 male children and 2 female children (age: mean 7.33, SD 1.61 years; intelligence score: mean 91.50, SD 6.71 points). The standard rehabilitation group included 11 male children (age: mean 7.18, SD 1.33 years; intelligence score: mean 88.36, SD 5.78 points). There were no significant differences in gender (*P*=.16), age (*P*=.81), and intelligence score (*P*=.24) between groups. The groups did not significantly differ in preintervention scores for the TVPS-3 test (*P*=.09-.92) or for the digital game test (*P*=.15-.95).

Comparison of Pre- and Postintervention Scores for TVPS-3 Between Groups

After training, visual discrimination (*P*=.007), visual memory (*P*=.008), form constancy (*P*=.01), sequential memory (*P*=.06), and full-test scores (*P*=.002) for the digital game group significantly increased. Scores for the remaining 3 components of the TVPS-3 were higher after the intervention, but the differences did not reach statistical significance—spatial relationships (*P*=.53), figure-ground (*P*=.08), visual closure (*P*=.06). In contrast, there were no significant changes in TVPS-3 scores for the standard rehabilitation group (*P*=.12-.73). Form constancy (*P*=.03), sequential memory (*P*=.03), visual closure (*P*=.02), and full score (*P*=.005) increases for the digital game group were greater than those for the standard rehabilitation group (Table 1).

Table 1. Pre- and postintervention TVPS-3 scores.

	Digital game group (n=12)			Standard rehabilitation group (n=11)			Difference (between pre- and postintervention)		
	Preintervention	Postintervention	*P* value	Preintervention	Postintervention	*P* value	Digital game group	Standard rehabilitation group	*P* value
Visual discrimination	6.50 (4.44)	9.42 (3.78)	.007	5.09 (2.17)	5.73 (2.15)	.31	2.92 (3.09)	0.64 (1.96)	.05
Visual memory	8.08 (3.99)	11.67 (3.00)	.008	5.55 (2.47)	6.27 (1.90)	.38	3.58 (3.82)	0.73 (2.61)	.05
Spatial relationships	7.33 (4.64)	8.25 (4.18)	.53	5.27 (3.38)	4.91 (3.81)	.73	0.92 (4.87)	–0.36 (3.41)	.48
Form constancy	5.75 (2.42)	8.83 (4.43)	.01	6.09 (4.39)	5.55 (4.18)	.65	3.08 (3.45)	–0.55 (3.86)	.03
Sequential memory	6.83 (3.79)	10.67 (4.36)	.006	5.09 (2.63)	5.91 (3.24)	.12	3.83 (3.90)	0.82 (1.60)	.03
Figure-ground	4.00 (1.71)	6.58 (4.38)	.08	3.91 (2.30)	4.73 (3.41)	.39	2.58 (4.68)	0.82 (2.99)	.30
Visual closure	3.17 (2.13)	6.08 (4.17)	.06	3.55 (2.54)	2.64 (1.43)	.18	2.92 (4.89)	–0.91 (2.07)	.02
Full score	41.67 (13.88)	61.50 (21.64)	.002	34.55 (13.00)	36.36 (13.40)	.58	19.83 (16.44)	1.82 (10.53)	.005

Digital Game Test Results Before and After Intervention

For the digital game group, there were significant increases in scores for the visual discrimination component (*P*=.03) of the digital game test and the full digital game test (*P*=.01). While scores for the remaining 6 components (visual memory: *P*=.05; spatial relationships: *P*=.05; form constancy: *P*=.54; sequential memory: *P*=.76; figure-ground: *P*=.06; visual closure: *P*=.09) did not significantly differ in the digital game group, postintervention scores tended to be slightly higher than preintervention scores. In contrast, the standard rehabilitation group revealed no significant differences in digital game test scores for visual discrimination (*P*=.53), visual memory (*P*=.13), spatial relationships (*P*=.89), form constancy (*P*=.93), sequential memory (*P*=.43), figure-ground (*P*=.49), visual closure (*P*=.40), and full score (*P*=.30) (Table 2).

Table 2. Pre- and postintervention scores on digital game test.

	Digital game group (n=12)			Standard rehabilitation group (n=11)			Difference (between pre- and postintervention)		
	Preintervention	Postintervention	*P* value	Preintervention	Postintervention	*P* value	Digital game group	Standard rehabilitation group	*P* value
Visual discrimination	8.33 (1.72)	9.75 (0.45)	.03	8.00 (3.97)	8.73 (1.10)	.53	1.42 (2.02)	0.73 (3.74)	.58
Visual memory	7.25 (3.25)	9.08 (0.90)	.05	7.45 (3.05)	9.00 (0.77)	.13	1.83 (2.82)	1.55 (3.08)	.82
Spatial relationships	6.83 (2.79)	8.67 (2.84)	.05	5.91 (3.11)	6.09 (3.62)	.89	1.83 (2.92)	0.18 (4.14)	.28
Form constancy	8.42 (1.73)	8.67 (1.56)	.54	6.82 (3.06)	6.91 (2.07)	.93	0.25 (1.36)	0.09 (3.48)	.89
Sequential memory	6.00 (2.95)	6.25 (3.22)	.76	5.91 (3.81)	6.73 (2.49)	.43	0.25 (2.77)	0.82 (3.31)	.66
Figure-ground	5.17 (3.13)	7.50 (3.53)	.06	5.45 (3.78)	6.09 (3.02)	.49	2.33 (3.92)	0.64 (2.98)	.26
Visual closure	7.67 (2.64)	9.08 (1.38)	.09	6.45 (3.56)	7.09 (2.98)	.40	1.42 (2.68)	0.64 (2.42)	.47
Full score	49.92 (11.69)	58.92 (9.59)	.01	45.82 (19.15)	50.27 (10.13)	.30	9.00 (10.20)	4.45 (13.43)	.37

Correlation Between TVPS-3 Test and the Digital Game Test

Spatial relationship items (*r*=0.66, *P*=.001), figure-ground items (*r*=0.65, *P*=.001), and items for the full score (*r*=0.47, *P*=.03) in the digital game test and in the TVPS-3 were significantly moderately correlated (Table 3).

Table 3. Analysis of Correlation between TVPS-3 Test and Digital Game Test.

	r	*P* value
Visual discrimination	0.06	.77
Visual memory	0.06	.78
Spatial relationships	0.66	.001
Form constancy	–0.01	.97
Sequential memory	0.25	.26
Figure-ground	0.65	.001
Visual closure	0.10	.64
Full score	0.47	.03

Satisfaction Questionnaire

Out of 30 satisfaction questionnaires distributed to parents and therapists, 26 valid and complete questionnaires were retrieved (24 from parents and 2 from therapists). All 5 themes had mean scores higher than 4, which indicated that parents and therapists were highly satisfied with the interactive digital game system, and Cronbach α=0.89-0.96, which indicated high reliability (Table 4).

Table 4. Internal consistency of the satisfaction questionnaire.

	Mean (SD)	Cronbach α
Perceived ease of use	4.01 (0.71)	0.89
Perceived usefulness	4.35 (0.58)	0.91
Perceived joyfulness	4.44 (0.61)	0.95
Satisfaction	4.41 (0.58)	0.96
Continued use	4.31 (0.69)	0.96

Discussion

Pre- and postintervention scores in the digital game group differed for all items except spatial relationships, figure-ground, and visual closure (Table 1). Previous literature has indicated that children with growth delays have low scores on visual discrimination and spatial relationships tests because insufficient visual attention and cortical visual dysfunction limits discrimination and classification capabilities [21]. Thus, these capabilities cannot be enhanced by a short period of training. In visuospatial ability training studies, verbal and visual prompts are the key factors for training success [22]. An effective digital multimedia system for learners should integrate multisensory stimulation and widely varying teaching strategies (eg, use of visual prompts such as photographs, auditory prompts, repetition) [23,24]. An example of a visual prompt is stimulus fading (ie, simultaneously presenting subjective and pictorial prompts and then fading the prompted picture in the sequence), which allows children to visually slowly concentrate on the subject or guides them to learn how to judge the subject with graphic animation paired with line changes. Moreover, so that children could more easily discriminate and judge the images, voice prompts were used to facilitate repeated training to help them learn through visualized clues [23,24].

Improvements after the digital game intervention for most TVPS-3 items were satisfactory, which is consistent with the findings of previous studies [10-13]. Multimedia provide acoustic-optic effects and interactivity, which attract the attention of learners and guides their learning. In comparison with traditional teaching materials (eg, static pictures) and traditional teaching methods (eg, paper-based activities), multimedia teaching materials can provide greater audiovisual stimuli (eg, animations and sound) to increase learning efficiency and motivation [24-27]. Additionally, a digital multimedia interface can motivate and guide the user to explore issues and accept challenges by providing highlights or keywords [28]. According to the literature [29], 60% of the learners achieve better learning outcomes when learning is supported by digital assistance multimedia. Thus, the abovementioned research results indicate that a digital multimedia program is more effective than traditional paper-based training. Integrating a digital multimedia interface in teaching can enhance learning motivation and efficiency.

There were moderate to high correlations between the TVPS-3 and digital game test (spatial relationships: r=0.66; figure-ground: r=0.65; full score: r=0.47) (Table 3); however, only the visual discrimination component (P=.03) and the full score (P=.01) in the digital game group exhibited significant differences between the pre- and postintervention scores. We investigated why the digital game test had lower sensitivity in detecting pre versus postintervention change compared to the TVPS-3. The digital game test was based on the 7 TVPS-3 themes (10 questions per theme and 70 total questions). The standard TVPS-3 had 16 questions per theme and 112 total questions. Therefore, the lower number of questions on the digital game test resulted in a higher rate of repetition, which decreased its sensitivity in detecting change. In addition to the high rate of repetition, the difficulty level was not high enough, since the questions in the digital game test were chosen from the primary version. Thus, digital game test questions may not have been sufficiently challenging for the identification and assessment of children with relatively higher function or older age. In future studies, the instrument should be modified so that the number of test items is identical to that of the TVPS-3, which would reduce repeatability of the image gallery and possibly enhance the credibility and validity of the system.

On the satisfaction questionnaire, 5 items had scores higher than 4 points (Table 4), which indicated that parents and therapists were highly satisfied with the interactive digital game system. The 2 clinical therapists who had participated in the course reported that the digital game could be a useful intervention for improving visual–perceptual skills in children since the operating interface was easy to use and visually attractive to children. Additionally, the improvement was comparable to that obtained from conventional paper-based game training. The 24 parents reported that their children found the digital game easy to play and found the Bluetooth controller easy to operate. As a result, their scores on the visual perception test increased, which indicates that development of visual–perceptual abilities can be enhanced when rehabilitation includes an interactive digital game assistance system [30]. Relevant literature also mentions that a digital game system that attracts the interest of users tends to have high user satisfaction and high continued use [16,17].

Additionally, traditional teaching materials are mainly paper-based. They are not easily modified and have limited modes of interactivity (mainly paper-pen mode). They cannot be shared conveniently; paper-based media can only be shared by making additional paper copies whereas digital media can be easily shared and copied via the internet. While digital multimedia has good interactivity (online synchronous learning application or operating mode combined with other virtual reality equipment), is strong in media integrity, can be flexibly adjusted at any time, can be conveniently downloaded from the cloud, and can be represented and edited with 3C software in real time [31].

In summary, the interactive digital game assistance system developed in this study has practical applications in clinical rehabilitation treatment. The system is applicable and suitable for visual–perceptual training in children with developmental disabilities.. Thus, application of this learning mode in specialized rehabilitation for children with visual–perceptual defects can substantially improve their visual–perceptual skills and can potentially increase active participation in other relevant courses of treatment.

The interactive digital game designed and developed in this study positively benefitted visual perception training for children with developmental disabilities. When used in combination with conventional training, the interactive digital game system not only confers larger improvements in vision and perception in children with developmental disabilities, it also increases their willingness to participate in the training. The interactive digital game training system developed in this study can be implemented by medical institutions, family care givers, and school educators to increase the effectiveness of conventional programs for developing visual–perceptual skills in children with developmental disabilities. To improve the applicability of the proposed system as a testing tool, our future work will expand the question bank to enable it to have a broader range of skill levels.

Acknowledgments

This research was funded by the Ministry of Science and Technology in Taiwan (MOST 109-2221-E-037-004 and MOST 109-2221-E-037-005), and in part, by the Intelligent Manufacturing Research Center from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 15418 KB - games_v10i2e34756_app1.pdf]

References

1. Majnemer A, Shevell MI. Diagnostic yield of the neurologic assessment of the developmentally delayed child. J Pediatr 1995 Aug;127(2):193-199. [doi: 10.1016/s0022-3476(95)70294-6] [Medline: 7543566]
2. Chang CL. Using case-based reasoning to diagnostic screening of children with developmental delay. Expert Syst Appl 2005 Feb;28(2):237-247. [doi: 10.1016/j.eswa.2004.09.002]
3. Shevell MI. The evaluation of the child with a global developmental delay. Semin Pediatr Neurol 1998 Mar;5(1):21-26. [doi: 10.1016/s1071-9091(98)80014-4] [Medline: 9548637]
4. Trower D, Nicol P. Life-span intellectual development of people with mental retardation. Dev Med Child Neurol 1996 Jul;38(7):645-650. [doi: 10.1111/j.1469-8749.1996.tb12129.x] [Medline: 8674915]
5. Haber JS. Early diagnosis and referral of children with developmental disabilities. Am Fam Physician 1991 Jan;43(1):132-140. [Medline: 1702577]
6. Huang MC, Liu MN, Kao YH, Huang SC. The physical interactive game design for autistic children. J Erg Stud 2009;10(2):1-10.
7. Chen FL. The Effects of App Game-based Remedial Instruction on Learning Area for School Children Master. Taipei, Taiwan: National Taiwan Normal University; 2013.
8. Yu YH. The Research of Digital Game Design on Improving Fine Motor Skills of Children with Developmental Delay. Taipei, Taiwan: National Taiwan University of Science and Technology; 2014.
9. Wu CP, Liu YC. Designa digital game to facilitate the mathematical learning process of students with mental retardation. 2015 Presented at: Open Source and Educational Technology; May 1; Taipei, Taiwan.
10. Buckle F. Effectiveness of Additional Computer-Based Intervention for the Remediation of Visual Perceptual Skills in 7 to 9 Year Old Learners Diagnosed with Attention Deficit Hyperactivity Disorder Doctor. Johannesburg, Republic of South Africa: University of the Witwatersrand; 2019.
11. Wuang YP, Chiu YH, Chen YJ, Chen CP, Wang CC, Huang CL, et al. Game-based auxiliary training system for improving visual perceptual dysfunction in children with developmental disabilities: a proposed design and evaluation. Comput Educ 2018 Sep;124:27-36. [doi: 10.1016/j.compedu.2018.05.008]
12. Lin HC, Chiu YH, Chen YJ, Wuang YP, Chen CP, Wang CC, et al. Continued use of an interactive computer game-based visual perception learning system in children with developmental delay. Int J Med Inform 2017 Nov;107:76-87. [doi: 10.1016/j.ijmedinf.2017.09.003] [Medline: 29029695]
13. Wuang YP, Chen YJ, Chiu YH, Wang CC, Chen CP, Huang CL, et al. Effectiveness of kinesthetic game-based training system in children with visual-perceptual dysfunction. IEEE Access 2021;9:153838-153849. [doi: 10.1109/access.2021.3128109]
14. Martin N. Test of Visual Perceptual Skills 3rd ed. Novato, California: American Therapy Publications; 2006.
15. Haynes C, Kornbluth J, Okonsky L. Test of visual perceptual skills -3rd edition. EBS Healthcare. 2016. URL: https://ebsuniversity.com/wp-content/uploads/2018/07/TVPS-3-overview-.pdf [accessed 2018-07-01]
16. Lee DY, Lehto MR. User acceptance of YouTube for procedural learning: an extension of the technology acceptance model. Comput Educ 2013 Feb;61:193-208. [doi: 10.1016/j.compedu.2012.10.001]
17. Cho V, Cheng TE, Lai WJ. The role of perceived user-interface design in continued usage intention of self-paced e-learning tools. Comput Educ 2009 Sep;53(2):216-227. [doi: 10.1016/j.compedu.2009.01.014]
18. Terzis V, Moridis CN, Economides AA. The effect of emotional feedback on behavioral intention to use computer based assessment. Comput Educ 2012 Sep;59(2):710-721. [doi: 10.1016/j.compedu.2012.03.003]

19. Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med 2018 Sep;18(3):91-93 [FREE Full text] [doi: 10.1016/j.tjem.2018.08.001] [Medline: 30191186]
20. Lance C, Butts M, Michels L. The sources of four commonly reported cutoff criteria. Organ Res Methods 2016 Jun 29;9(2):202-220. [doi: 10.1177/1094428105284919]
21. Wuang YP, Tsai HY. Sensorimotor and visual perceptual functioning in school-aged children with Williams syndrome. J Intellect Disabil Res 2017 Apr 30;61(4):348-362. [doi: 10.1111/jir.12346] [Medline: 27900816]
22. Chabani E, Hommel B. Effectiveness of visual and verbal prompts in training visuospatial processing skills in school age children. Instructional Science 2014 Apr 5;42(6):995-1012. [doi: 10.1007/s11251-014-9316-7]
23. Braswell R, Flynt S, Mosley V. Teaching sight-words to special needs students using technology. 1996 Presented at: Society for Information Technology & Teacher Education International Conference; March; Waynesville, North Carolina URL: https://www.learntechlib.org/p/46985
24. Mechling LC, Gast DL, Langone J. Computer-based video instruction to teach persons with moderate intellectual disabilities to read grocery aisle signs and locate items. J Spec Educ 2016 Aug 18;35(4):224-240. [doi: 10.1177/002246690203500404]
25. Mayer RE, Sims VK. For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. J Educ Psychol 1994 Sep;86(3):389-401. [doi: 10.1037/0022-0663.86.3.389]
26. Park JH, Park JH. A randomized controlled trial of the computer-based cognitive rehabilitation program for children (CoTras-C) to examine cognitive function and visual perception in children with developmental disabilities. J Phys Ther Sci 2015 Dec;27(12):3623-3626 [FREE Full text] [doi: 10.1589/jpts.27.3623] [Medline: 26834318]
27. Park O, Gittelman SS. Selective use of animation and feedback in computer-based instruction. Educ Technol Res Dev 1992 Dec;40(4):27-38. [doi: 10.1007/bf02296897]
28. Visser J, Keller JM. The clinical use of motivational messages: an inquiry into the validity of the ARCS model of motivational design. Instructional Science 1990;19(6):467-500. [doi: 10.1007/bf00119391]
29. Fletcher-Flinn CM, Gravatt B. The efficacy of computer assisted instruction (CAI): a meta-analysis. J Educ Comput Res 2016 Nov 28;12(3):219-241. [doi: 10.2190/51d4-f6l3-jqhu-9m31]
30. Hackbarth G, Grover V, Yi MY. Computer playfulness and anxiety: positive and negative mediators of the system experience effect on perceived ease of use. Inf Management 2003 Jan;40(3):221-232. [doi: 10.1016/s0378-7206(02)00006-x]
31. Lee CY, Wu HJ, Kuo HC. The educational technology and elearning content developing trends. J Curriculum Stud 2011 Mar;6(1):1-26.

Abbreviations

TVPS: Test of Visual Perceptual Skills

Edited by N Zary; submitted 08.11.21; peer-reviewed by CS Ouyang, JT Tsai, YJ Chen; comments to author 11.12.21; accepted 17.02.22; published 15.04.22.

Please cite as:
Wu WL, Huang YL, Liang JM, Chen CH, Wang CC, Ho WH
Interactive Digital Game for Improving Visual–Perceptual Defects in Children With a Developmental Disability: Randomized Controlled Trial
JMIR Serious Games 2022;10(2):e34756
URL: https://games.jmir.org/2022/2/e34756
doi:10.2196/34756
PMID:35436215

Original Paper

Effects of Virtual Reality–Based Multimodal Audio-Tactile Cueing in Patients With Spatial Attention Deficits: Pilot Usability Study

Samuel Elia Johannes Knobel[1], BSc, MSc, MD; Brigitte Charlotte Kaufmann[2], BSc, MSc, PhD; Nora Geiser[3], BSc, MSc; Stephan Moreno Gerber[1], BSc, MSc, PhD; René M Müri[1,4,5], MD, Prof Dr; Tobias Nef[1,5,6], Prof Dr; Thomas Nyffeler[1,3,4,5], Prof Dr; Dario Cazzoli[1,3,7], PhD, PD

[1]Gerontechnology & Rehabilitation Group, University of Bern, Bern, Switzerland

[2]Sorbonne Université, Institut du Cerveau - Paris Brain Institute (ICM), Inserm, Centre national de la recherche scientifique, Hôpital de la Pitié-Salpêtrière, Paris, France

[3]Neurocenter, Luzerner Kantonsspital, Luzern, Switzerland

[4]Perception and Eye Movement Laboratory, Departments of Neurology and BioMedical Research, Inselspital, Bern University Hospital, Bern, Switzerland

[5]Department of Neurology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[6]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

[7]Institute of Psychology, University of Bern, Bern, Switzerland

Corresponding Author:
Dario Cazzoli, PhD, PD
Neurocenter
Luzerner Kantonsspital
Spitalstrasse
Luzern, 6004
Switzerland
Phone: 41 205 56 86
Email: dario.cazzoli@luks.ch

Abstract

Background: Virtual reality (VR) devices are increasingly being used in medicine and other areas for a broad spectrum of applications. One of the possible applications of VR involves the creation of an environment manipulated in a way that helps patients with disturbances in the spatial allocation of visual attention (so-called hemispatial neglect). One approach to ameliorate neglect is to apply cross-modal cues (ie, cues in sensory modalities other than the visual one, eg, auditory and tactile) to guide visual attention toward the neglected space. So far, no study has investigated the effects of audio-tactile cues in VR on the spatial deployment of visual attention in neglect patients.

Objective: This pilot study aimed to investigate the feasibility and usability of multimodal (audio-tactile) cueing, as implemented in a 3D VR setting, in patients with neglect, and obtain preliminary results concerning the effects of different types of cues on visual attention allocation compared with noncued conditions.

Methods: Patients were placed in a virtual environment using a head-mounted display (HMD). The inlay of the HMD was equipped to deliver tactile feedback to the forehead. The task was to find and flag appearing birds. The birds could appear at 4 different presentation angles (lateral and paracentral on the left and right sides), and with (auditory, tactile, or audio-tactile cue) or without (no cue) a spatially meaningful cue. The task usability and feasibility, and 2 simple in-task measures (performance and early orientation) were assessed in 12 right-hemispheric stroke patients with neglect (5 with and 7 without additional somatosensory impairment).

Results: The new VR setup showed high usability (mean score 10.2, SD 1.85; maximum score 12) and no relevant side effects (mean score 0.833, SD 0.834; maximum score 21). A repeated measures ANOVA on task performance data, with presentation angle, cue type, and group as factors, revealed a significant main effect of cue type ($F_{30,3}$=9.863; P<.001) and a significant 3-way interaction ($F_{90,9}$=2.057; P=.04). Post-hoc analyses revealed that among patients without somatosensory impairment, any cue led to better performance compared with no cue, for targets on the left side, and audio-tactile cues did not seem to have additive effects. Among patients with somatosensory impairment, performance was better with both auditory and audio-tactile cueing than with no cue, at every presentation angle; conversely, tactile cueing alone had no significant effect at any presentation angle.

Analysis of early orientation data showed that any type of cue triggered better orientation in both groups for lateral presentation angles, possibly reflecting an early alerting effect.

Conclusions: Overall, audio-tactile cueing seems to be a promising method to guide patient attention. For instance, in the future, it could be used as an add-on method that supports attentional orientation during established therapeutic approaches.

(JMIR Serious Games 2022;10(2):e34884) doi:10.2196/34884

KEYWORDS

virtual reality; search task; stroke; neglect, multimodal cueing; bird search task

Introduction

Virtual reality (VR) devices are being increasingly used, and can be found in industries [1,2] and areas of entertainment, military [3,4], and medicine [5-11]. Skills learned in VR have been shown to be transferable to the real world [7,12,13], leading to a broad spectrum of possibilities, from training to rehabilitation. Virtual environments have the advantage of being fully customizable, and therefore, they are potentially adaptable to different situations and even different user abilities [14,15]. Particularly in the field of rehabilitation of neurological disorders, VR has been shown to be a promising approach because of this adaptability and the inclusion of different sensory information [16]. This is believed to not only promote cortical reorganization but also facilitate the activation of neuronal plasticity [17]. One of the possible applications of VR in this area is the creation of an environment that can be manipulated in a way that would help patients with disturbances in the spatial allocation of visual attention. These disturbances are often subsumed under the label of visual neglect, a frequent condition occurring after right hemispheric stroke (up to 70%) [18,19]. Visual neglect is characterized by the inability to respond or react to targets coming from the contralesional side of space. It is a negative prognostic factor for the overall outcome after stroke and is difficult to treat [20].

One approach to ameliorate visual neglect is to apply cross-modal cues (ie, cues in sensory modalities other than the visual one) in order to guide visual attention toward the neglected side. Most commonly, auditory cues have been successfully applied in neglect patients in order to guide visual attention toward the contralesional side [21-24]. Although less often investigated, tactile cues (alone or in combination with auditory cues) seem also to be able to ameliorate the contralesional allocation of visual attention in this patient population [21,25].

Most setups of the above-mentioned studies were based on 2D screens, complex speaker arrays, or loudspeakers on moving robot arms. The use of a 3D VR environment can reduce the complexity of the system and increase its ecological validity, owing to the higher immersion, larger visual field, and possibility to freely move the head and body [26].

Indeed, the effects of auditory cues, presented in a VR setting, on visuospatial attention deployment have been successfully explored in neglect patients and show promising results [27,28]. In contrast, to the best of our knowledge, no study has so far investigated the effects of tactile cues in VR on the spatial deployment of visual attention in neglect patients, although results obtained in healthy controls seem encouraging [29,30]. Finally, previous results in 2D settings suggest that multimodal cueing (ie, combining auditory and tactile cueing at the same time) may result in superior effects than single cueing [29,31]. However, the effects of this combination for patients with visual attention deficits in a 3D VR setting have not been studied.

Thus, this pilot study aimed to investigate the feasibility and usability of multimodal (auditory and tactile) cueing, as implemented in a 3D VR setting (our bird search task), in patients with visual neglect, as well as obtain results concerning the effects of different types of cues on visual attention allocation compared with noncued conditions in these patients. We hypothesized that (1) the implementation of a new system, including tactile, auditory, and combined audio-tactile cueing, is feasible and usable for patients with impaired spatial attention; (2) the different cue types have a positive effect on the visuospatial attention allocation ability of patients; and (3) the use of multimodal cues (combined auditory and tactile) has larger effects than unimodal cues (auditory or tactile alone).

Methods

Demographics

Between June 2020 and November 2020, 12 patients with left-sided visual neglect after subacute right hemispheric stroke were recruited and included in the study.

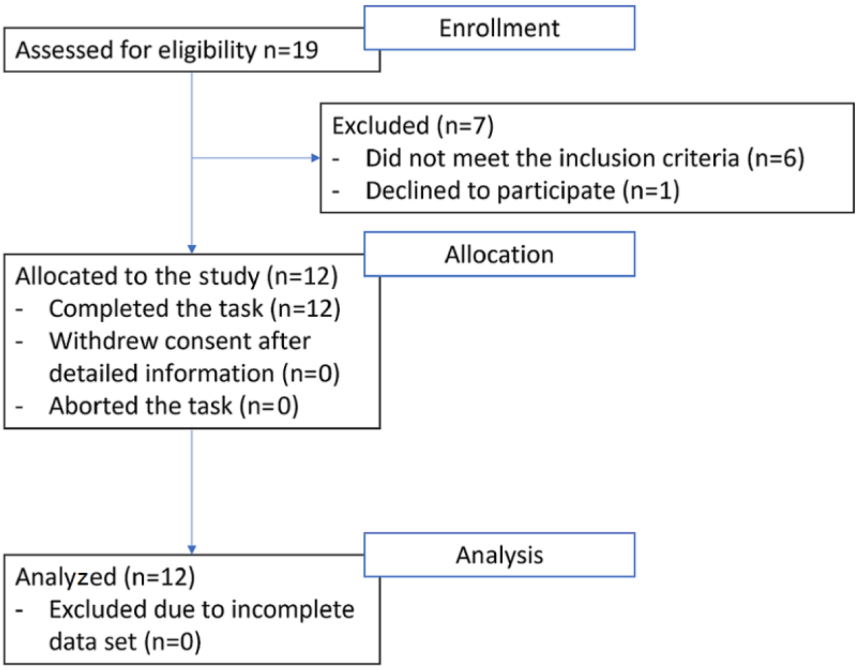
They were inpatients at the Neurorehabilitation Clinics of the Inselspital, Bern University Hospital, or the Kantonsspital Luzern, Switzerland. All patients had normal or corrected-to-normal vision. The mean age of the patients was 58.2 years (SD 9.70 years), and 4 were female. See Table 1 for detailed information and Figure 1 for the recruitment flowchart.

Table 1. Detailed demographics and results of the neuropsychological tests of the participants.

Patient ID	Gender	Age (years)	CBS[a] score	CoC[b]	LBT[c]	Somatosensory impairment
P01	Male	54	8	0.084	12.68%	Yes
P02	Male	42	4	0.043	4.98%	No
P03	Female	63	4	−0.107	5.28%	No
P04	Female	63	10	0.490	65.57%	No
P05	Male	46	8	0.024	11.11%	No
P06	Male	61	15	0.699	5.75%	No
P07	Female	66	4	0.003	2.70%	Yes
P08	Male	69	13	−0.005	9.34%	No
P09	Female	44	9	−0.004	1.24%	Yes
P10	Male	58	12	0.863	37.13%	Yes
P11	Male	61	14	0.739	3.32%	Yes
P12	Male	71	19	0.740	3.81%	No

[a]CBS: Catherine Bergego Scale [32].
[b]CoC: center of cancellation [33].
[c]LBT: Line Bisection Text (relative deviation in % [34]).

Figure 1. Patient flowchart from enrollment to analysis.



The inclusion criterion was a pathological score in 1 of the following 3 assessments: Catherine Bergego Scale (CBS) [32], Sensitive Neglect Test (SNT) [35], and Line Bisection Test [34].

The CBS is an established questionnaire allowing to detect the presence and severity of neglect based on the observation of everyday life activities (cutoff for neglect: CBS score >1). The SNT is a paper-pencil cancellation task. The spatial distribution and number of missed targets are evaluated using the center of cancellation (CoC; cutoff: CoC >0.081) [33]. The CoC reflects the mean position from the center to the missed targets and is normalized to values from −1 to 1. Zero indicates no spatial bias, negative values indicate a shift towards the left, and positive values indicate a shift towards the right. Furthermore, the Line Bisection Test [34] is another frequently used neuropsychological task to assess neglect severity, with a cutoff value of ≥11% mean relative rightward deviation [34].

Somatosensory impairment, and hearing and auditory extinction were assessed clinically. Somatosensory impairment was assessed by comparing the sensitivity for touch on the forehead and the temporal head region between the left and right. To assess hearing and auditory extinction, rustling was presented to each ear individually or to both ears simultaneously, and patients' reports were compared. Based on this assessment, patients were assigned to the following 2 subgroups: patients with and without sensory impairment. As no patient had auditory extinction, no auditory extinction group was formed.

Ethics Approval

The study was approved by the Ethics Committee of the Cantons of Bern and Lucerne (ID 2017-02195), and was conducted in accordance with the latest version of the Declaration of Helsinki. All participants signed written informed consent forms before participation.

Questionnaires

A selection of different questionnaires was administered to assess the usability and side effects of the VR task, as previously described by Gerber et al [36,37].

For assessments of acceptance, usability, and participant perception of the visual search task and the VR system, the System Usability Scale (SUS) [38] was used. To assess side effects and cybersickness [39], the Simulator Sickness Questionnaire (SSQ) [40] was used.

Technical Setup

The VR hardware consisted of a stand-alone head-mounted display (HMD) and a hand-held controller. During the task, the positions of the controller and the HMD were continuously recorded (20 Hz sample rate).

The resolution of the HMD (Oculus Quest, Facebook Technologies) was 1440×1600 pixels, with a horizontal field of view of 110 degrees and a frame rate of 72 Hz. For the development of our bird search task, the platform Unity [41] was used.

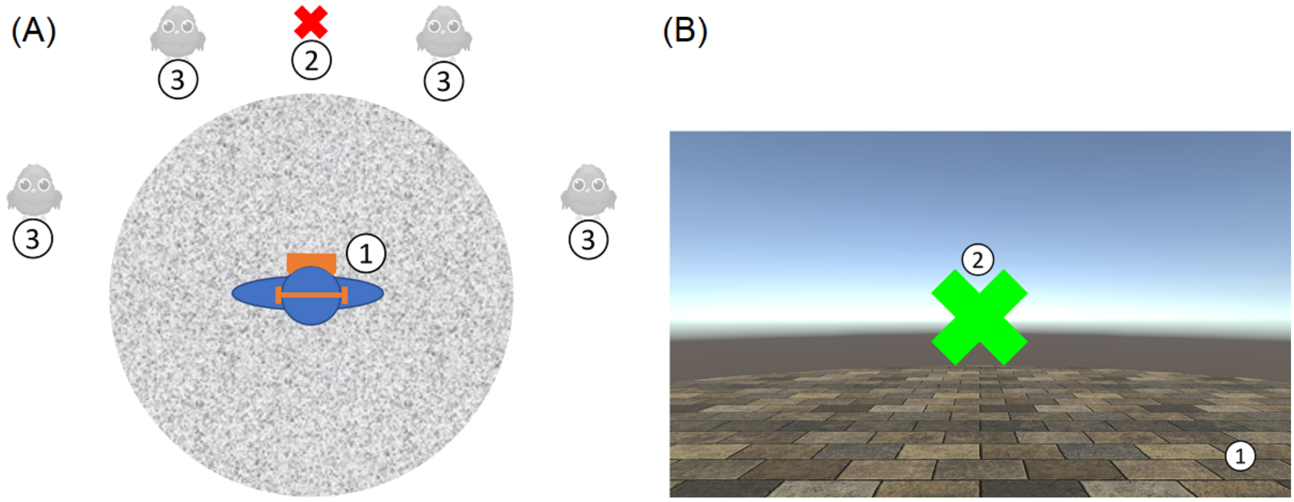
To monitor the patient's behavior during the task, the patient's view was projected to a laptop computer using a wireless streaming tool via Sidequest [42]. The experimenter was thus able to see what the patient saw in real-time and could intervene if needed.

Additionally, for the application of the tactile cue, a special cushion as an inlay piece for the HMD was used, as described by Knobel et al [29]. The cushion contains 6 coin vibrators symmetrically distributed over the forehead. Each coin vibrator could be started individually. The sound was presented via over-ear headphones.

Bird Search Task and Study Procedure

The main goal of the bird search task was to detect the appearing birds as quickly as possible. The task took place in a virtual environment and contained 3 main features (Figure 2).

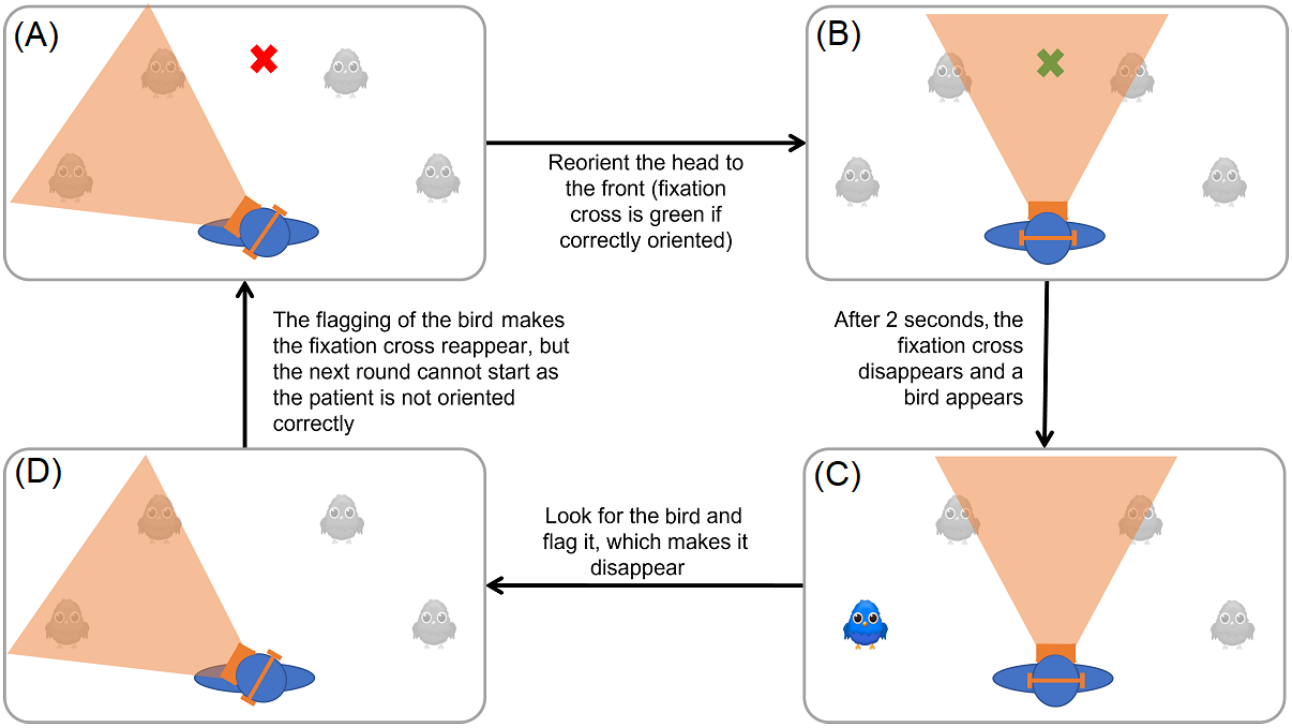
Figure 2. (A) Scenario setup in the virtual environment, with the following 3 main features: (1) An empty floor, on which the player was positioned at the center of the scene; (2) A central fixation cross; and (3) Four presentation angles (−70°, −30°, 30°, and 70°) where the birds could appear in the bird search task. (B) Patients view when starting the task with the floor (1) and the central fixation cross (2).



First, there was a plane surface, looking like a floor, which prevented the patient from feeling like floating or even falling. Second, directly in front of the patient, a colored fixation cross was presented. The patient's initial position was aligned, so the fixation cross was straight ahead. Third, there were 4 presentation angles at which the stimuli (blue birds), one at a time, could appear. These presentation angles were on a horizontal line at −70°, −30°, 30°, and 70°, as defined with respect to the patient's trunk (see the presentation angles of 4 birds in Figure 2). In this study, we will consider the 30° and −30° angles as paracentral angles and the 70° and −70° angles as lateral angles.

At the beginning of each trial (a full round in Figure 3), the patients were asked to orient the HMD toward a central fixation cross (red cross in Figure 2; Figure 3A), thereby aligning their head to the trunk straight ahead. The fixation cross had the property to turn green if aligned with the HMD, thereby giving feedback for being correctly fixated (Figure 3B). After fixating for 2 seconds on the central fixation cross, it disappeared and, at the same time, a target bird appeared (Figure 3C). Consequently, the disappearance of the central fixation cross was a signal for the patient to start searching for the target bird. The patient was asked to confirm the detection of the target bird by flagging it with a hand-held controller. A flagged bird disappeared from the scene by falling. This started a new trial (ie, central fixation cross for 2 seconds, disappearance of the latter, and concomitant appearance of the next target bird; Figure 3D). If the target bird was not found within 11 seconds, it simply disappeared, and the next trial started.

Figure 3. Schematic representation of an exemplary trial of the bird search task. The grey birds represent the potential presentation angles, where the birds could appear. (A) The patient's head (in blue) is not oriented toward the fixation cross, and therefore the task cannot start. (B) After reorienting the head toward the fixation cross, the cross turns green. (C) After 2 seconds of being correctly oriented, the fixation cross disappears and the bird appears (in this case, the bird appears at the presentation angle −70°). (D) The patient orients toward the bird and flags it, so it disappears. This results in the reappearance of the fixation cross that stays red as long as the patient has not reoriented toward it.



Each target bird could appear either with or without additional spatial cues. There were 2 different possible spatial cues (auditory or tactile), resulting in the following 4 different conditions: None (no cue), Audio (only auditory cue), Tactile (only tactile cue), and Combo (auditory and tactile cues; for more details, see the "Cues" section below).

In the None condition, the patient had to start looking for the target as soon as the fixation cross disappeared, but no additional hint for the direction was given. In the 3 other conditions, the patient received an additional hint from the presentation angles of the newly appeared target bird at the same time the target appeared. The cues were repeated after 3 seconds until the maximum presentation time of 11 seconds.

The bird search task was organized into 4 sessions, each containing a single spatial cue type (ie, None, Audio, Tactile, and Combo). In each session, 80 target birds were presented (ie, each of the 4 presentation angles were tested 20 times). The order of the presentation angles at which the target birds appeared was random, with the constraint that 2 target birds could not appear at the same presentation angle in 2 consecutive trials. The order of the 4 sessions was randomized. The 4 sessions took place over 2 consecutive days (ie, 2 sessions per day). On the first day, all the conditions were explained to the patient, and 3 practice trials were performed. During the practice trials, each cue type and each presentation angle was presented. First, the patient was verbally guided to the target, and if the position was not detected correctly, the corresponding practice trial was repeated up to 3 times. If the position was detected correctly, the next angle/cue was presented. After completing the practice, the first 2 sessions were performed. On the second day, the corresponding practice trials of the 2 remaining cueing conditions were repeated, and then, the 2 remaining sessions were performed.

Cues

The task included auditory and tactile spatial cues. The auditory cue was a 1-second tone of 500 Hz with 0.1 seconds fading in and out to reduce the sound's sharpness. The spatial information of the auditory cue was generated by the audio spatializer software development tool kit that is part of the Unity game development platform [41]. The head-related transfer function of the spatializer is based on the KEMAR data set (set of pre-ear impulse recordings of a dummy head; Bill Garnder at MIT Media Lab [43]). In the task, the sound was realistically perceived as coming from the left (due to different sound volume levels between the 2 ears) if the sound source was on the left side of the head and as coming from the front if the head was oriented toward the sound source.

The tactile cue (1-second vibration) was applied using punctual vibration that came from a special inlay of the HMD that contained 6 symmetrically positioned individually controllable coin vibrators [29]. The coin vibrators were controlled in a way that mimicked the spatial behavior of sound. This means that if the head is oriented to the front and a cue is given on the left side, the most left coin vibrator is activated, signaling that the cue is on the left side. If the head is then turned and faces the cue, the middle coin vibrators are activated, signaling that the cue is now directly in the front of the patient. If the head is turned even more, the right coin vibrators are activated, signaling that the cue is now on the right side.

In the Combo condition, auditory and tactile cues were given simultaneously for 1 second each.

In-Task Parameters

Performance

The main parameters assessed were the mean time until targets were flagged and the percentage of found and flagged birds for each presentation angle and condition. From the percentage found and the mean time of found targets, we calculated a performance measure [44]. The performance measure was calculated for every presentation angle and condition by dividing the percentage of found targets by the mean time needed to flag the targets. A higher value as a result of this calculation thus represents better performance.

This performance measure allows combining both behavioral aspects (ie, accuracy and speed of reactions) within a single parameter. With respect to the cognitive impairments of neglected patients, this performance measure offers several advantages. First, it allows quantifying performance even if a neglect patient does not find any target for a given angle/condition. Indeed, in this case, it would not be possible to calculate the mean time to flag targets. However, as the value is multiplied by the percentage of found targets (in this case, 0%) in the formula, the result would correctly be 0 (indicating very low performance for that particular presentation angle and condition). Second, this performance measure is robust against extreme values.

If, for instance, a patient finds only 1 target in a very fast way in a certain presentation angle/condition combination, this low reaction time would be overrepresented if only the mean time to flag targets is analyzed. Instead, in the performance measure formula, the inclusion of the percentage of found targets reduces the weight of this extreme value [44].

Early Orientation

During the task, the angular position of the HMD with respect to the trunk was continuously recorded, thus reflecting the head rotation over time. Based on these data, the early orientation [45-47], that is, the direction (to the left or right) of the first head rotation movement after the presentation of a stimulus, was extracted. We then used the ratio of correct early orientation instances (ie, the instances in which the head was initially turned to the left even though the target appeared on the right, and vice versa; Figure 4) per condition and presentation angle as an in-game performance parameter.

Figure 4. Schematic representation of 7 trials and how the ratio of correct early orientation was calculated. The x-axis shows the head angle from the point where the target appears until it is found. The angle is in relation to the starting position of the trunk. Green represents traces of correct orientation and red represents traces of incorrect orientation.

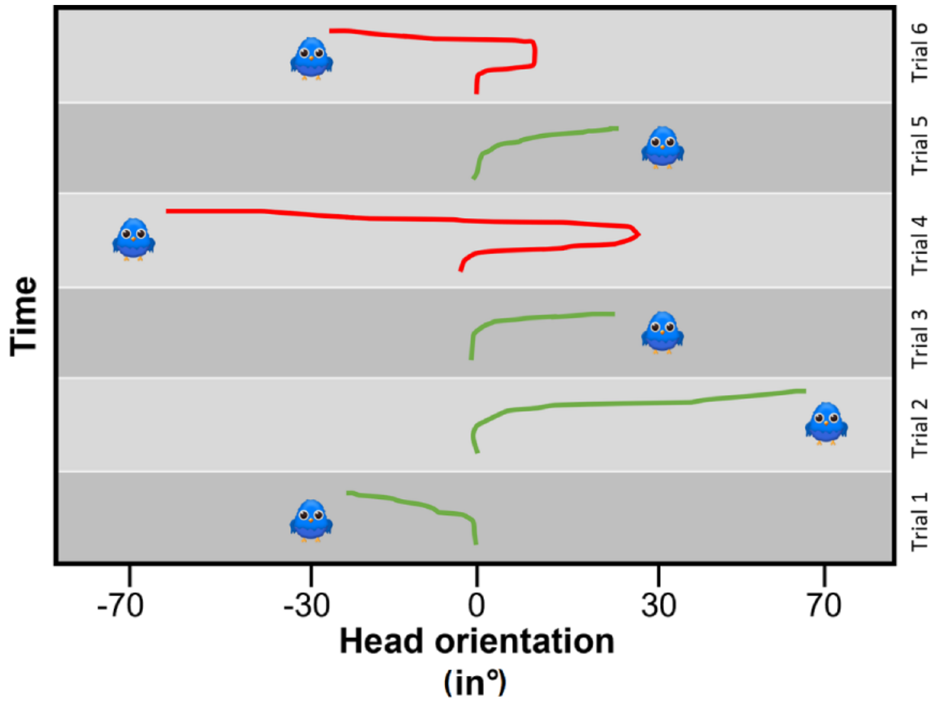


Figure 4 shows a schematic representation of the early orientation parameter. The traces (green and red lines) represent the head orientation during the presentation of 6 targets (blue birds). In this example, when the fourth and sixth targets were presented (red traces), the patient first turned the head toward the right, even though the target appeared on the left. When the first, second, third, and fifth trials (green traces) were performed, the orientation was correct from the beginning. Thus, the ratio of correct early orientation instances for the right side is 1 (ie, all early orientation instances were correct when the target appeared on the right side) and for the left side is 1/3.

The early orientation was determined by comparing the head's angle when the target appeared (central) with the head's angle after 1.5 seconds. If the difference (start angle minus angle after 1.5 seconds) was negative, the patient had turned the head to the right; otherwise, the patient had turned the head to the left.

Data Collection and Statistical Analyses

Statistical analyses were performed once for all patients together, and comparisons were performed between the with and without somatosensory impairment groups (both groups had no auditory extinction).

First, we analyzed the effects of the different cueing conditions and of the different spatial positions on in-game parameters by means of a 2-way repeated measures ANOVA, with cue type (levels: None, Audio, Tactile, and Combo) and angle (levels: −70°, −30°, 30°, and 70°) as within-subject factors for all participants.

Second, since it is assumed that auditory and tactile cueing require intact somatosensory processing to be effective, we aimed to assess the effect of impairment in the respective modality. For this purpose, we grouped the patients according to somatosensory impairment (yes/no) and auditory extinction (yes/no), and reran the ANOVA with group as a between-subject factor (levels: with and without somatosensory impairment), and cue type (levels: None, Audio, Tactile, and Combo) and angle (levels: −70°, −30°, 30°, and 70°) as within-subject factors.

Post-hoc analyses with least significant difference were applied for the identification of significant differences for cues within presentation angles. The study data were managed using Research Electronic Data Capture [48,49], a web-based tool to support data handling for research studies. Data analyses were performed with R (R Foundation for Statistical Computing) and MATLAB (MathWorks).

Results

Subgroups

With regard to the neuropsychological measures, the mean results were as follows: 10.0 (SD 4.79) points for the CBS score [32], 0.30 (SD 0.37) for the CoC [33], and 13.6% (SD 19.0%) relative deviation toward the right for the Line Bisection Test [34].

On separation for somatosensory impairment, there were 7 patients without and 5 with somatosensory impairment of the front head. On comparing age and neuropsychological measures between the groups, there were no significant differences for age ($t_{9.94}$=0.480; *P*=.64) and the neuropsychological measures (CBS: $t_{10.0}$=0.376, *P*=.72; CoC: $t_{7.78}$=−0.289, *P*=.78; Line Bisection Test: $t_{9.99}$=0.343, *P*=.74).

Questionnaires

Usability (based on the SUS ratings) was rated with the sum of 3 questions (responses ranging from 0 [unusable] to 4 [highly usable]) and reached a value of 10.2 (SD 1.85) out of a maximum of 12. Both groups provided similar ratings, with no significant difference between the groups ($t_{8.95}$=−1.54; *P*=.16).

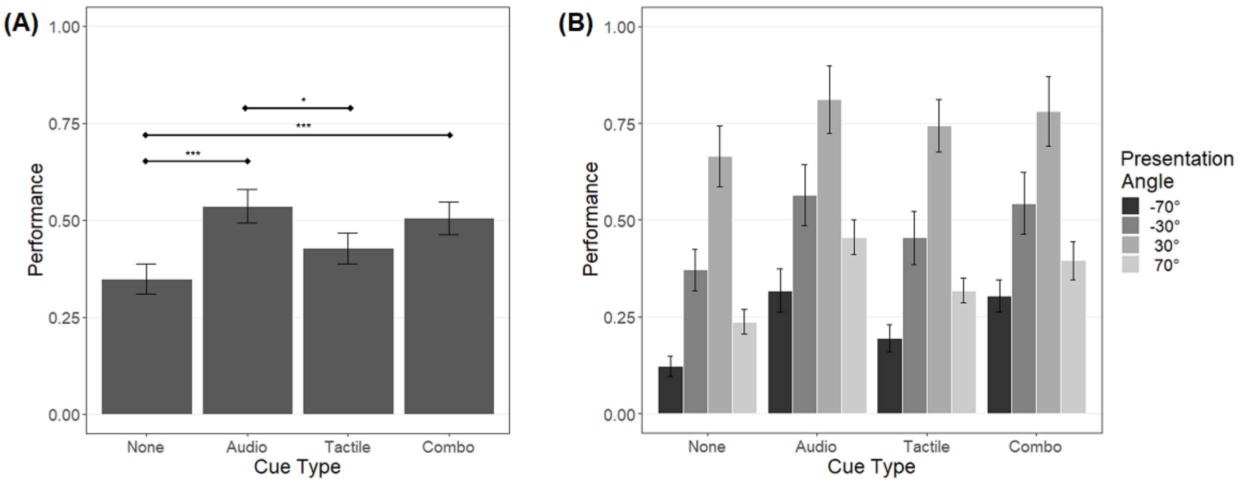
The occurrence of side effects was assessed with the sum of 7 frequent items from the SSQ (responses ranging from 0 [no side effects] to 3 [severe side effects]), and the score was 0.833 (SD 0.834) out of a theoretical maximum of 21, indicating a very low rate of side effects. The results were not significantly different between the groups (SSQ: $t_{9.86}$=−1.77; *P*=.11).

In-Task Parameters

Performance: All Patients

Results concerning all patients are shown in Figure 5. Two-way repeated measures ANOVA showed significant main effects of cue type ($F_{33,3}$=7.60; *P*<.001) and presentation angle ($F_{19.35,1.76}$=50.9; *P*<.001), but not of their interaction (cue type × presentation angle) ($F_{99,9}$=0.469; *P*=.89).

Figure 5. (A) Performance per cue type for all patients (n=12). The whiskers represent the standard error of means. The asterisks represent the level of significance of post-hoc tests (**P*<.05, ***P*<.01, ****P*<.001). (B) Visualization of the performance per cue type and presentation angle.

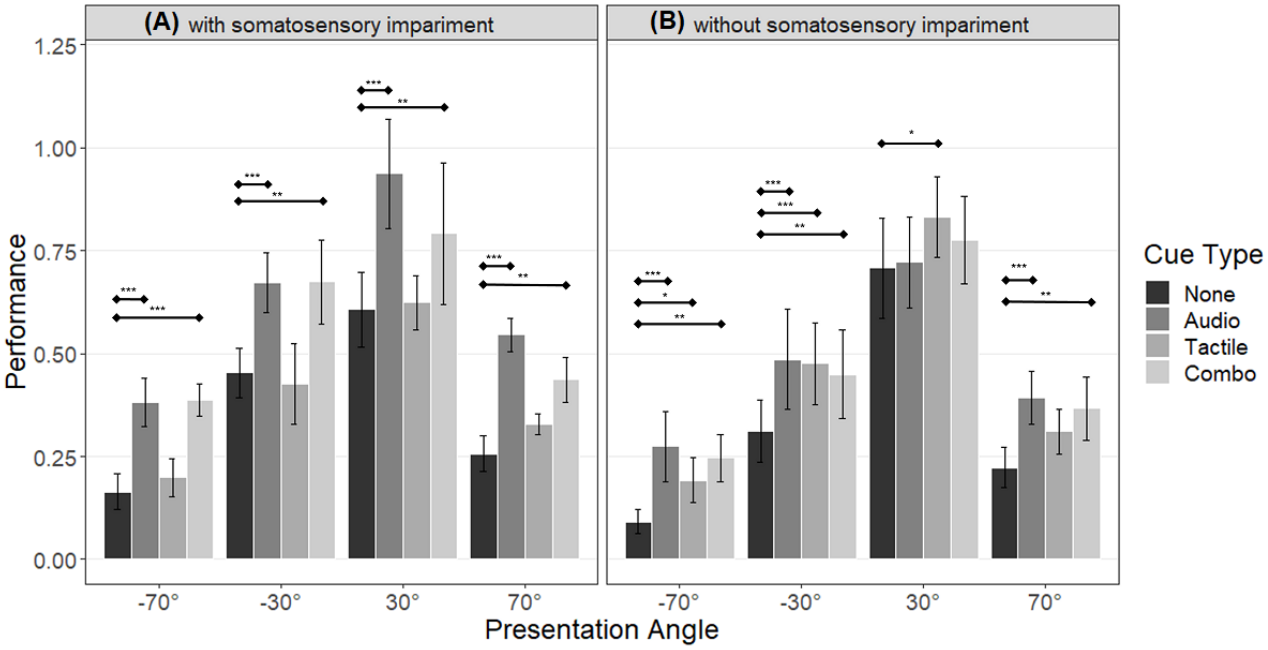


Post-hoc analyses concerning the main effects of cue type (ie, irrespective of presentation angle) revealed a significantly better performance for auditory or combined cues but not for tactile cues compared with no cue. Furthermore, auditory cues were better than tactile cues.

Performance: Subgroups

Detailed examination of the results (Figure 6) included a 3-way mixed model ANOVA that showed significant main effects of cue type ($F_{30,3}$=9.863; *P*<.001) and presentation angle ($F_{17.24,1.72}$=47.609; *P*<.001), and a significant 3-way interaction between cue type, presentation angle, and group ($F_{90,9}$=2.057; *P*=.04). There were no other significant interactions.

Figure 6. Visualization of the performance, split for the group with somatosensory impairment (A; n=5) and the group without somatosensory impairment (B; n=7). Results of the post-hoc tests are shown as significance bars (level of significance: **P*<.05, ***P*<.01, ****P*<.001). The whiskers represent the standard error of means.

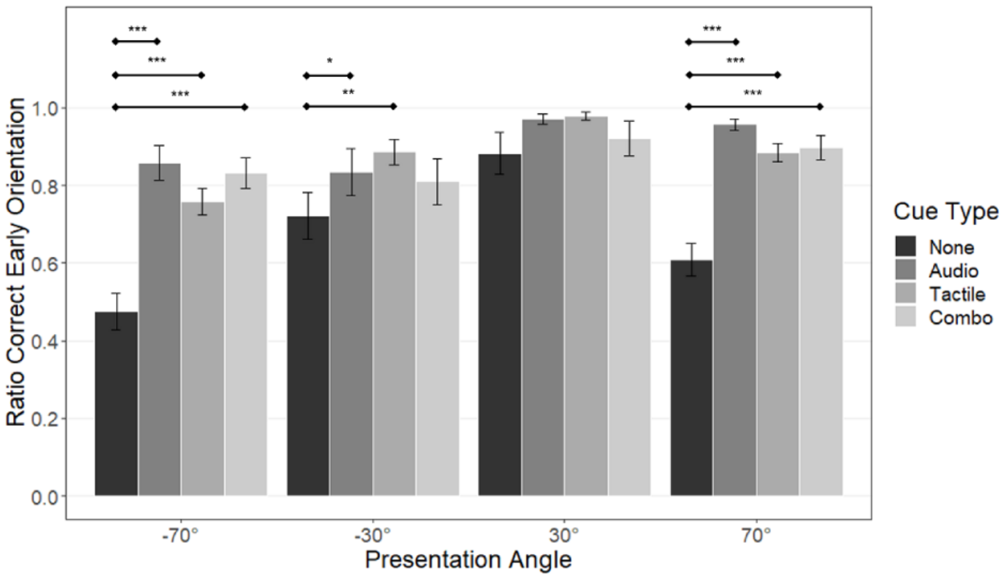


The post-hoc analysis revealed that for patients without somatosensory impairment, any cue led to better performance than no cue for targets on the left side. In patients with somatosensory impairment, performance was better with both auditory and audio-tactile cueing than with no cue, at every presentation angle; conversely, tactile cueing alone had no significant effect at any presentation angle.

Early Orientation: All Patients

Results of the ratio of correct early orientation for all patients are shown in Figure 7. Repeated measures ANOVA showed significant main effects of cue type ($F_{36,3}$=26.934; *P*<.001), presentation angle ($F_{36,3}$=10.207; *P*<.001), and their interaction ($F_{99,9}$=4.798; *P*=.001). Having cues at lateral angles led to an improvement in the ratio of correct early orientation. Comparison of the same cues at different presentation angles showed that in the None and Audio conditions, there was a significant left-right difference for the paracentral (None: *P*=.002; Audio: *P*=.008) and lateral (None: *P*=.008; Audio: *P*=.047) angles. This pattern was not observed in the Tactile and Combo conditions, where only the lateral (Tactile: *P*=.01) and paracentral (Combo: *P*=.03) angles differed significantly. The significance bars for direct cue comparisons are not shown in Figure 7.

Figure 7. Visualization of the ratio of correct early orientation instances. The whiskers represent the standard error of means. Results of the post-hoc tests are shown as significance bars. The asterisks above the bars represent the level of significance (**P*<.05, ***P*<.01, ****P*<.001).

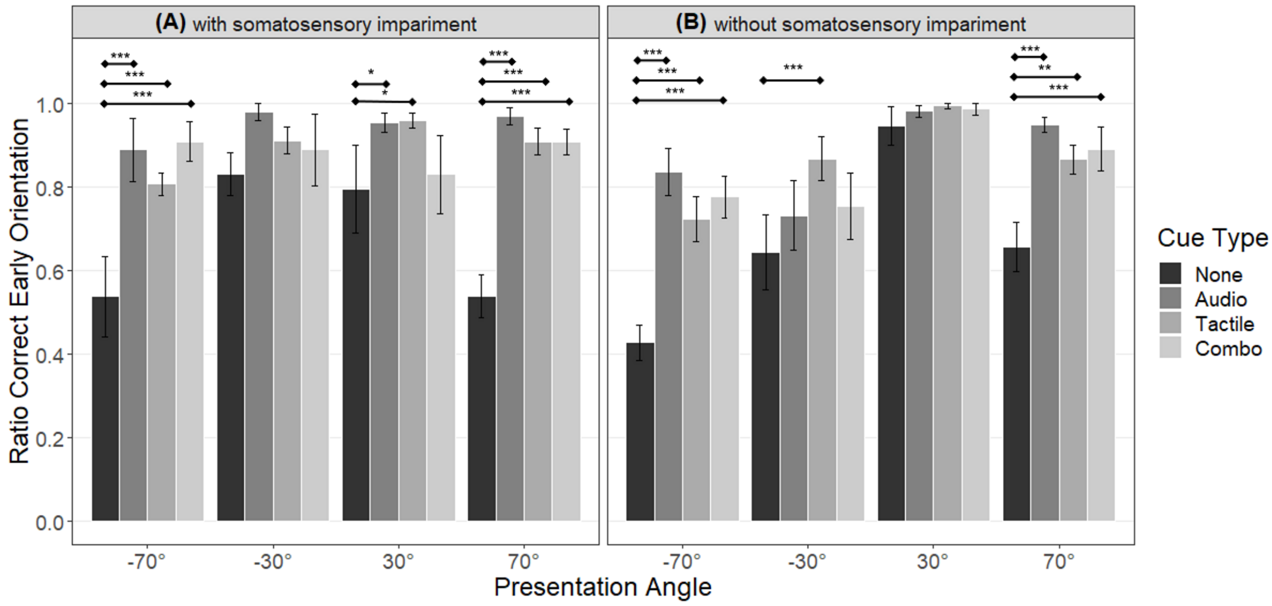


Early Orientation: Subgroups

A mixed model ANOVA for cue type, presentation angle, and group showed significant main effects of cue type ($F_{30,3}$=32.671; P<.001) and presentation angle ($F_{30,3}$=13.97; P<.001), as well as the following 2 significant 2-way interactions: group × presentation angle ($F_{30,3}$=6.738; P=.001) and cue type × presentation angle ($F_{90,9}$=4.412; P<.001). No other interaction reached significance.

The lack of a significant 3-way interaction also showed that any type of cue (auditory, tactile, or combo) triggered better orientation in both groups for lateral angles (Figure 8).

Figure 8. Visualization of the interaction result of presentation angle and cue type, split for the group with somatosensory impairment (A) and the group without somatosensory impairment (B). The whiskers represent the standard error of means. Results of the post-hoc tests are shown as significance bars. The asterisks above the bars represent the level of significance (*P<.05, **P<.01, ***P<.001).



The significant interaction between group and presentation angle showed a significant difference between the groups for left paracentral angles and a left-right difference for lateral and paracentral angles in the group without somatosensory impairment.

Discussion

Overview

This study showed that our new multimodal VR setup, including visual, auditory, and tactile cues, is highly usable and that it is suitable to provide multimodal spatial cues to patients who have spatial attention deficits. Furthermore, the measurements provided some important information concerning the peculiarities of the application of such a setup in this population. Indeed, all patients seemed to benefit from auditory cues; however, the positive effect of tactile cues on in-task performance seemed to depend on the presence of somatosensory impairment. Patients without somatosensory impairment did benefit from tactile cues. Nevertheless, there seemed to be a consistent alerting effect for all cue conditions (auditory, tactile, and combined) among all patients, which could help them to overcome the initial tendency to orient toward the right, a common symptom observed in neglect. Finally, we did not find an additive effect of combining auditory and tactile cues in any measure or group.

Usability

In line with our first hypothesis, patients rated the device's usability as very high and did not report any relevant side effects. The high usability of our new VR setting, including tactile stimulation, is similar to that in one of the few previous studies in stroke patients with a similar visual and auditory setting, which however did not entail any tactile stimulation [14,37]. Moreover, similar levels of usability have been reported by healthy participants using new VR tools including visual and auditory stimulation, but without tactile stimulation [36,50,51]. Hence, in our study, the additional presentation of tactile stimulation did not change usability or cause discomfort. Interestingly, in our last study in healthy participants [29], tactile stimulation, similar to that presented here, caused some side effects. We hypothesize that one reason for the minor discomfort caused by tactile stimulation is the duration of application of the tactile stimulation. Indeed, in our previous study, tactile stimulation was continuously presented over several minutes, whereas in this study, it was only presented during short intervals (ie, up to a maximum of 4 seconds per target). This study and our previous study [29] thus represent in some way maximum and minimum usage of tactile stimulation. Future studies should therefore target an optimal tactile stimulation time for maximum effectiveness, but without side effects.

In-Task Parameters

Our second hypothesis concerned in-task parameters and assumed that patients with spatial attention disorder would benefit from spatial cues, as assessed by means of a compound performance measure, considering both search time and percentage of found targets. Usually, those cues are presented on a 2D screen with a limited visual angle. The advantage of our VR approach is that we can present the patients with far more lateral cues in a virtual environment, supporting them in orienting toward the left by means of not only eye movements, but also head movements. Even though some tools have been developed and tested to examine the effects of visual, auditory, or even tactile cues in patients with impaired spatial attention [21,52], to the best of our knowledge, so far, no study has examined the effects of tactile cues in neglect patients using a VR system.

The auditory cueing of our VR setup was effective in all patients, increasing performance and confirming our hypothesis. The new approach with tactile cueing in VR was also effective, as long as patients had preserved somatosensory processing. This is in line with the principles of unimodal or multimodal cueing, a common approach to guide the attention of patients with spatial attention disorders toward the neglected side [24,53,54].

The fact that tactile cueing can induce positive effects in patients without somatosensory impairment is particularly remarkable. Indeed, tactile cues were presented at different locations in space than visual stimuli; the tactile cues were presented on the skin of the forehead (ie, within the personal space), whereas visual targets were presented in the peripersonal space. This configuration is different from the one concerning auditory cues, were visual targets and auditory cues are both presented at the same location in the peripersonal space [55,56]. In this group of patients, this suggests the activation of spatial attention networks through not only spared supramodal mechanisms [52,57] but also "supraspatial" (ie, entailing different spatial reference frames) mechanisms.

Our results seem to not provide support for the third hypothesis, that is, an advantage of the combined application of tactile and auditory cues in comparison with unimodal cues, since this did not lead to a significant increase in performance or a higher rate of correct early orientation. Some evidence in the literature shows that the use of multimodal cues can have an additive effect [29,31], whereas other studies have shown no benefit of using multimodal cues when compared with unimodal cues [54]. A possible explanation for these disparities could be associated with the level of spatiotemporal matching [58,59] and cross-modal intensity matching [60] needed for cross-modal cues to have an additive effect. In our study, auditory cues had a high spatiotemporal match with the visual target, but tactile cues did not [54]. This might be even more relevant for neglect patients, as they typically show problems with spatial transformations [61-63].

Besides performance parameters (eg, search time and percentage of found targets), we also assessed early orientation behavior. Interestingly, in this case, all types of cues (auditory, tactile, and combination) had positive effects on early orientation, even in patients who had a somatosensory impairment and could not correctly localize a tactile cue. This might be explained through a general alerting effect, as it has been shown that a temporary alertness enhancement can ameliorate attentional orienting in neglect patients [64-66], but this effect is short lived [67,68]. Therefore, it seems reasonable to hypothesize that all types of cues can ameliorate early orientation through their alerting effect and that this effect is not long enough to further support performance to find targets.

Limitations

In this pilot study, we aimed to show the feasibility and usability of our VR setup in a diverse and complex patient group. On the one hand, the variance increases the difficulty of the interpretability of the task results, and on the other hand, the variety increases the representativeness for this patient group. Neglect severity shows spontaneous fluctuations and intraday variations [69]. We tried to minimize this effect by having the 2 sessions on consecutive days and at the same time; nevertheless, factors like medication changes and exhausting therapies beforehand could not be controlled. Future studies with a larger sample size should try to control or at least assess such confounders

Conclusion and Outlook

This study showed the usability and feasibility of a new approach entailing auditory, tactile, and combined audio-tactile cueing in VR among patients with different combinations of attentional and somatosensory impairments. Our data suggest that auditory and tactile cues may be equally efficient in ameliorating attentional performance in neglect patients, at least in those with spared somatosensory processing. Moreover, combined audio-tactile stimulation did not show an additive effect in our setup. Future studies are needed to assess these preliminary findings in a larger group of patients. One of the possible directions to reach intermodal additive effects would be to assess the effects of tactile cues presented "nearer" to the source of the visual and auditory stimuli (eg, tactile stimulation on the hand, with variation in presentation side and intensity according to how near the patient moves the hand to the target).

Overall, audio-tactile cueing seems to be a promising method to guide patient attention. For instance, in the future, it could be used as an add-on approach that supports attentional orientation during established therapeutic approaches (eg, optokinetic stimulation).

Acknowledgments

This work was supported by Swiss National Science Foundation (SNSF) Grants (P2BEP3_195283 to BCK, 32003B_196915 and 320030_169789 to TNyffeler, and Z00P3_154714/1 to DC). We are grateful to the patients who took part in our study and to the clinical teams of the Inselspital and the Luzerner Kantonsspital for their support.

Authors' Contributions

SEJK, DC, BCK, RMM, TNef, and TNyffeler designed the study. SEJK, DC, and TNef developed the tool and the setup. SEJK and BCK recruited the participants. SEJK, DC, BCK, NG, TNyffeler, and RMM analyzed the data. SEJK, DC, BCK, NG, and TNef wrote the manuscript. All authors approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Berg LP, Vance JM. Industry use of virtual reality in product design and manufacturing: a survey. Virtual Reality 2016 Sep 12;21(1):1-17. [doi: 10.1007/s10055-016-0293-9]
2. Wang T, Huang H, Huang H. Design and implementation of cancellation tasks for visual search strategies and visual attention in school children. Computers & Education 2006 Aug;47(1):1-16. [doi: 10.1016/j.compedu.2004.08.012]
3. Rizzo A, Parsons TD, Lange B, Kenny P, Buckwalter JG, Rothbaum B, et al. Virtual reality goes to war: a brief review of the future of military behavioral healthcare. J Clin Psychol Med Settings 2011 Jun 7;18(2):176-187. [doi: 10.1007/s10880-011-9247-2] [Medline: 21553133]
4. Liu X, Zhang J, Hou G, Wang Z. Virtual Reality and Its Application in Military. IOP Conf. Ser.: Earth Environ. Sci 2018 Jul 17;170:032155. [doi: 10.1088/1755-1315/170/3/032155]
5. Izard SG, Juanes JA, García Peñalvo FJ, Estella JMG, Ledesma MJS, Ruisoto P. Virtual Reality as an Educational and Training Tool for Medicine. J Med Syst 2018 Mar 01;42(3):50. [doi: 10.1007/s10916-018-0900-2] [Medline: 29392522]
6. Mazurek J, Kiper P, Cieślik B, Rutkowski S, Mehlich K, Turolla A, et al. Virtual reality in medicine: a brief overview and future research directions. hm 2019;20(3):16-22. [doi: 10.5114/hm.2019.83529]
7. Zaid MB, Dilallo M, Shau D, Ward DT, Barry JJ. Virtual Reality as a Learning Tool for Trainees in Unicompartmental Knee Arthroplasty: A Randomized Controlled Trial. J Am Acad Orthop Surg 2022 Jan 15;30(2):84-90. [doi: 10.5435/JAAOS-D-20-01357] [Medline: 34520419]
8. Bartlett J, Kazzazi F, To K, Lawrence J, Khanduja V. Virtual Reality Simulator Use Stimulates Medical Students' Interest in Orthopaedic Surgery. Arthrosc Sports Med Rehabil 2021 Oct;3(5):e1343-e1348 [FREE Full text] [doi: 10.1016/j.asmr.2021.05.014] [Medline: 34712972]
9. Maples-Keller JL, Bunnell BE, Kim S, Rothbaum BO. The Use of Virtual Reality Technology in the Treatment of Anxiety and Other Psychiatric Disorders. Harv Rev Psychiatry 2017;25(3):103-113. [doi: 10.1097/hrp.0000000000000138]
10. Rothbaum BO, Hodges LF. The use of virtual reality exposure in the treatment of anxiety disorders. Behav Modif 1999 Oct 26;23(4):507-525. [doi: 10.1177/0145445599234001] [Medline: 10533438]
11. Coelho CM, Waters AM, Hine TJ, Wallis G. The use of virtual reality in acrophobia research and treatment. J Anxiety Disord 2009 Jun;23(5):563-574. [doi: 10.1016/j.janxdis.2009.01.014] [Medline: 19282142]
12. Torkington J, Smith SG, Rees BI, Darzi A. Skill transfer from virtual reality to a real laparoscopic task. Surg Endosc 2001 Oct;15(10):1076-1079 [FREE Full text] [doi: 10.1007/s004640000233] [Medline: 11727073]
13. Bossard C, Kermarrec G, Buche C, Tisseau J. Transfer of learning in virtual environments: a new challenge? Virtual Reality 2008 Mar 26;12(3):151-161. [doi: 10.1007/s10055-008-0093-y]
14. Knobel SEJ, Kaufmann BC, Gerber SM, Urwyler P, Cazzoli D, Müri RM, et al. Development of a Search Task Using Immersive Virtual Reality: Proof-of-Concept Study. JMIR Serious Games 2021 Jul 02;9(3):e29182 [FREE Full text] [doi: 10.2196/29182] [Medline: 34255653]
15. Paraskevopoulos I, Tsekleves E, Warland A, Kilbride C. Virtual Reality-Based Holistic Framework: A Tool for Participatory Development of Customised Playful Therapy Sessions for Motor Rehabilitation. 2016 Presented at: 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES); September 7-9, 2016; Barcelona, Spain p. 1-8. [doi: 10.1109/vs-games.2016.7590378]
16. Maggio MG, Latella D, Maresca G, Sciarrone F, Manuli A, Naro A, et al. Virtual Reality and Cognitive Rehabilitation in People With Stroke: An Overview. J Neurosci Nurs 2019 Apr;51(2):101-105. [doi: 10.1097/JNN.0000000000000423] [Medline: 30649091]
17. Bray S, Shimojo S, O'Doherty JP. Direct Instrumental Conditioning of Neural Activity Using Functional Magnetic Resonance Imaging-Derived Reward Feedback. Journal of Neuroscience 2007 Jul 11;27(28):7498-7507. [doi: 10.1523/jneurosci.2118-07.2007]
18. Stone SP, Halligan PW, Greenwood RJ. The Incidence of Neglect Phenomena and Related Disorders in Patients with an Acute Right or Left Hemisphere Stroke. Age Ageing 1993;22(1):46-52. [doi: 10.1093/ageing/22.1.46]
19. Bowen A, McKenna K, Tallis RC. Reasons for variability in the reported rate of occurrence of unilateral spatial neglect after stroke. Stroke 1999 Jun;30(6):1196-1202 [FREE Full text] [doi: 10.1161/01.str.30.6.1196] [Medline: 10356099]
20. Jehkonen M, Laihosalo M, Kettunen JE. Impact of neglect on functional outcome after stroke: a review of methodological issues and recent research findings. Restor Neurol Neurosci 2006;24(4-6):209-215. [Medline: 17119299]

21. Hommel M, Peres B, Pollak P, Memin B, Besson G, Gaio J, et al. Effects of passive tactile and auditory stimuli on left visual neglect. Arch Neurol 1990 May 01;47(5):573-576. [doi: 10.1001/archneur.1990.00530050097018] [Medline: 2334306]
22. Schenke N, Franke R, Puschmann S, Turgut N, Kastrup A, Thiel CM, et al. Can auditory cues improve visuo-spatial neglect? Results of two pilot studies. Neuropsychol Rehabil 2021 Jun 27;31(5):710-730. [doi: 10.1080/09602011.2020.1727931] [Medline: 32102605]
23. Frassinetti F, Pavani F, Làdavas E. Acoustical vision of neglected stimuli: interaction among spatially converging audiovisual inputs in neglect patients. J Cogn Neurosci 2002 Jan 01;14(1):62-69. [doi: 10.1162/089892902317205320] [Medline: 11798387]
24. Kaufmann B, Cazzoli D, Bartolomeo P, Frey J, Pflugshaupt T, Knobel S, et al. Auditory spatial cueing reduces neglect after right-hemispheric stroke: A proof of concept study. Cortex 2022 Mar;148:152-167. [doi: 10.1016/j.cortex.2021.12.009] [Medline: 35176552]
25. Bailey MJ, Riddoch MJ, Crome P. Treatment of Visual Neglect in Elderly Patients With Stroke: A Single-Subject Series Using Either a Scanning and Cueing Strategy or a Left-Limb Activation Strategy. Physical Therapy 2002;82(8):782-797. [doi: 10.1093/ptj/82.8.782]
26. 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]
27. Baheux K, Yoshizawa M, Tanaka A, Seki K, Handa Y. Diagnosis and rehabilitation of hemispatial neglect patients with virtual reality technology. THC 2005 Jul 25;13(4):245-260. [doi: 10.3233/thc-2005-13404]
28. Huygelier H, Schraepen B, Lafosse C, Vaes N, Schillebeeckx F, Michiels K, et al. An immersive virtual reality game to train spatial attention orientation after stroke: A feasibility study. Appl Neuropsychol Adult 2020 Sep 18:1-21. [doi: 10.1080/23279095.2020.1821030] [Medline: 32945702]
29. Knobel SEJ, Gyger N, Nyffeler T, Cazzoli D, Muri RM, Nef T. Development and evaluation of a new virtual reality-based audio-tactile cueing-system to guide visuo-spatial attention. Annu Int Conf IEEE Eng Med Biol Soc 2020 Jul;2020:3192-3195. [doi: 10.1109/EMBC44109.2020.9176516] [Medline: 33018683]
30. Lindeman RW, Sibert J, Mendez-Mendez E, Patil S, Phifer D. Effectiveness of directional vibrotactile cuing on a building-clearing task. In: CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2005 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 2-7, 2005; Portland, Oregon, USA p. 271-280. [doi: 10.1145/1054972.1055010]
31. Hancock PA, Mercado JE, Merlo J, Van Erp JB. Improving target detection in visual search through the augmenting multi-sensory cues. Ergonomics 2013 May;56(5):729-738. [doi: 10.1080/00140139.2013.771219] [Medline: 23510197]
32. Bergego C, Azouvi P, Samuel C, Marchal F, Louis-Dreyfus A, Jokic C, et al. Validation d'une échelle d'évaluation fonctionnelle de l'héminégligence dans la vie quotidienne: l'échelle CB. Annales de Réadaptation et de Médecine Physique 1995 Jan;38(4):183-189. [doi: 10.1016/0168-6054(96)89317-2]
33. Rorden C, Karnath H. A simple measure of neglect severity. Neuropsychologia 2010 Jul;48(9):2758-2763 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2010.04.018] [Medline: 20433859]
34. Wilson B, Cockburn J, Halligan P. Development of a behavioral test of visuospatial neglect. Arch Phys Med Rehabil 1987 Mar;68(2):98-102. [Medline: 3813864]
35. Reinhart S, Leonhard E, Kerkhoff G. Sensitive Neglect Test (SNT) single and dual task. Saarland University. 2016. URL: https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/kerkhoff/Materialien_f%C3%BCr_Diagnostik_Therapie/Instructions_SNT_single_and_dual_task.pdf [accessed 2022-04-29]
36. Gerber SM, Jeitziner M, Wyss P, Chesham A, Urwyler P, Müri RM, et al. Visuo-acoustic stimulation that helps you to relax: A virtual reality setup for patients in the intensive care unit. Sci Rep 2017 Oct 16;7(1):13228 [FREE Full text] [doi: 10.1038/s41598-017-13153-1] [Medline: 29038450]
37. Knobel SEJ, Kaufmann BC, Gerber SM, Cazzoli D, Müri RM, Nyffeler T, et al. Immersive 3D Virtual Reality Cancellation Task for Visual Neglect Assessment: A Pilot Study. Front Hum Neurosci 2020 May 25;14:180 [FREE Full text] [doi: 10.3389/fnhum.2020.00180] [Medline: 32528265]
38. 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, USA: CRC Press; 1996:189-194.
39. LaViola JJ. A discussion of cybersickness in virtual environments. SIGCHI Bull 2000 Jan 01;32(1):47-56. [doi: 10.1145/333329.333344]
40. Kennedy R, Lane N, Berbaum K, Lilienthal M. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. The International Journal of Aviation Psychology 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
41. Unity. Unity Technologies. URL: https://unity.com/ [accessed 2022-04-29]
42. SideQuestVR. URL: https://sidequestvr.com/ [accessed 2022-04-29]
43. HRTF Measurements of a KEMAR Dummy-Head Microphone. MIT Media Lab. URL: https://sound.media.mit.edu/resources/KEMAR.html [accessed 2022-02-17]

44. Vandierendonck A. Further Tests of the Utility of Integrated Speed-Accuracy Measures in Task Switching. J Cogn 2018 Jan 12;1(1):8 [FREE Full text] [doi: 10.5334/joc.6] [Medline: 31517182]
45. Pflugshaupt T, Bopp SA, Heinemann D, Mosimann UP, von Wartburg R, Nyffeler T, et al. Residual oculomotor and exploratory deficits in patients with recovered hemineglect. Neuropsychologia 2004 Jan;42(9):1203-1211. [doi: 10.1016/j.neuropsychologia.2004.02.002] [Medline: 15178172]
46. Kaufmann BC, Knobel SEJ, Nef T, Müri RM, Cazzoli D, Nyffeler T. Visual Exploration Area in Neglect: A New Analysis Method for Video-Oculography Data Based on Foveal Vision. Front Neurosci 2019 Jan 22;13:1412 [FREE Full text] [doi: 10.3389/fnins.2019.01412] [Medline: 32038129]
47. Azouvi P, Samuel C, Louis-Dreyfus A, Bernati T, Bartolomeo P, Beis J, French Collaborative Study Group on Assessment of Unilateral Neglect (GEREN/GRECO). Sensitivity of clinical and behavioural tests of spatial neglect after right hemisphere stroke. J Neurol Neurosurg Psychiatry 2002 Aug;73(2):160-166 [FREE Full text] [doi: 10.1136/jnnp.73.2.160] [Medline: 12122175]
48. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
49. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
50. Gerber S, Muri R, Mosimann U, Nef T, Urwyler P. Virtual reality for activities of daily living training in neurorehabilitation: a usability and feasibility study in healthy participants. 2018 Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 18-21, 2018; Honolulu, HI, USA p. 1-4. [doi: 10.1109/EMBC.2018.8513003]
51. Wenk N, Penalver-Andres J, Buetler KA, Nef T, Müri RM, Marchal-Crespo L. Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment. Virtual Reality 2021 Aug 16. [doi: 10.1007/s10055-021-00565-8]
52. Zigiotto L, Damora A, Albini F, Casati C, Scrocco G, Mancuso M, et al. Multisensory stimulation for the rehabilitation of unilateral spatial neglect. Neuropsychol Rehabil 2021 Oct 19;31(9):1410-1443. [doi: 10.1080/09602011.2020.1779754] [Medline: 32558611]
53. Kerkhoff G. Modulation and rehabilitation of spatial neglect by sensory stimulation. Prog Brain Res 2003;142:257-271. [doi: 10.1016/S0079-6123(03)42018-9] [Medline: 12693266]
54. Ngo MK, Spence C. Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. Attention, Perception, & Psychophysics 2010 Aug;72(6):1654-1665. [doi: 10.3758/app.72.6.1654]
55. Ho C, Reed N, Spence C. Assessing the effectiveness of "intuitive" vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. Accid Anal Prev 2006 Sep;38(5):988-996. [doi: 10.1016/j.aap.2006.04.002] [Medline: 16697344]
56. Ladavas E, Farnè A. Neuropsychological Evidence for Multimodal Representations of Space near Specific Body Parts. In: Spence C, Driver J, editors. Crossmodal Space and Crossmodal Attention. Oxford, UK: Oxford University Press; 2004:69-98.
57. Bolognini N, Vallar G. Hemianopia, spatial neglect, and their multisensory rehabilitation. In: Sathian K, Ramachandran VS, editors. Multisensory Perception. Cambridge, MA, USA: Academic Press; 2020:423-447.
58. Paladini RE, Diana L, Zito GA, Nyffeler T, Wyss P, Mosimann UP, et al. Attentional reorienting triggers spatial asymmetries in a search task with cross-modal spatial cueing. PLoS One 2018 Jan 2;13(1):e0190677 [FREE Full text] [doi: 10.1371/journal.pone.0190677] [Medline: 29293637]
59. Mclntire JP, Havig PR, Watamaniuk SNJ, Gilkey RH. Visual search performance with 3-D auditory cues: effects of motion, target location, and practice. Hum Factors 2010 Mar 04;52(1):41-53. [doi: 10.1177/0018720810368806] [Medline: 20653224]
60. Pitts B, Riggs SL, Sarter N. Crossmodal Matching: A Critical but Neglected Step in Multimodal Research. IEEE Trans. Human-Mach. Syst 2016 Jun;46(3):445-450. [doi: 10.1109/thms.2015.2501420]
61. Manfredini A, Mancini F, Posteraro L, Savazzi S. The two sides of spatial representation in neglect patients: the same spatial distortion for different patterns of performance. Neuropsychologia 2013 Aug;51(10):1867-1877. [doi: 10.1016/j.neuropsychologia.2013.06.013] [Medline: 23800681]
62. Karnath H. Disturbed coordinate transformation in the neural representation of space as the crucial mechanism leading to neglect. Neuropsychological Rehabilitation 1994 Jun;4(2):147-150. [doi: 10.1080/09602019408402273]
63. Bisiach E, Luzzatti C. Unilateral Neglect of Representational Space. Cortex 1978 Mar;14(1):129-133 [FREE Full text] [doi: 10.1016/s0010-9452(78)80016-1] [Medline: 16295118]
64. Posner MI. Measuring alertness. Ann N Y Acad Sci 2008;1129:193-199. [doi: 10.1196/annals.1417.011] [Medline: 18591480]
65. Niemi P, Näätänen R. Foreperiod and simple reaction time. Psychological Bulletin 1981;89(1):133-162. [doi: 10.1037/0033-2909.89.1.133]
66. Van der Burg E, Olivers CNL, Bronkhorst AW, Theeuwes J. Pip and pop: nonspatial auditory signals improve spatial visual search. J Exp Psychol Hum Percept Perform 2008 Oct;34(5):1053-1065. [doi: 10.1037/0096-1523.34.5.1053] [Medline: 18823194]

67. Cazzoli D, Kaufmann BC, Paladini RE, Müri RM, Nef T, Nyffeler T. Anterior insula and inferior frontal gyrus: where ventral and dorsal visual attention systems meet. Brain Commun 2021;3(1):fcaa220 [FREE Full text] [doi: 10.1093/braincomms/fcaa220] [Medline: 33501424]
68. Chandrakumar D, Keage HA, Gutteridge D, Dorrian J, Banks S, Loetscher T. Interactions between spatial attention and alertness in healthy adults: A meta-analysis. Cortex 2019 Oct;119:61-73. [doi: 10.1016/j.cortex.2019.03.016] [Medline: 31075552]
69. Li K, Malhotra PA. Spatial neglect. Pract Neurol 2015 Oct 28;15(5):333-339 [FREE Full text] [doi: 10.1136/practneurol-2015-001115] [Medline: 26023203]

Abbreviations

CBS: Catherine Bergego Scale
CoC: Center of Cancellation
HMD: head-mounted display
SNT: Sensitive Neglect Test
SSQ: Simulator Sickness Questionnaire
SUS: System Usability Scale
VR: virtual reality

Edited by N Zary; submitted 17.11.21; peer-reviewed by R Calabrò; comments to author 08.01.22; revised version received 17.02.22; accepted 07.04.22; published 25.05.22.

Please cite as:
Knobel SEJ, Kaufmann BC, Geiser N, Gerber SM, Müri RM, Nef T, Nyffeler T, Cazzoli D
Effects of Virtual Reality–Based Multimodal Audio-Tactile Cueing in Patients With Spatial Attention Deficits: Pilot Usability Study
JMIR Serious Games 2022;10(2):e34884
URL: https://games.jmir.org/2022/2/e34884
doi:10.2196/34884
PMID:35612894

Original Paper

Perceptions of Cognitive Training Games and Assessment Technologies for Dementia: Acceptability Study With Patient and Public Involvement Workshops

Kyle Harrington[1,2,3], BSc, MA, PhD; Michael P Craven[1,3,4], BSc, MSc, PhD; Max L Wilson[3,5], MEng, PhD; Aleksandra Landowska[3,5], BA, MRes, PhD

[1]NIHR MindTech MedTech Co-operative, Institute of Mental Health, Nottingham, United Kingdom

[2]School of Medicine, Faculty of Medicine and Health Sciences, University of Nottingham, Nottingham, United Kingdom

[3]NIHR Nottingham Biomedical Research Centre, University of Nottingham, Nottingham, United Kingdom

[4]Human Factors Research Group, Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom

[5]Mixed Reality Lab, School of Computer Science, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Kyle Harrington, BSc, MA, PhD
NIHR MindTech MedTech Co-operative
Institute of Mental Health
University of Nottingham Innovation Park
Nottingham, NG7 2TU
United Kingdom
Phone: 44 115 823 1294
Email: kyle.harrington@nottingham.ac.uk

Abstract

Background: Cognitive training and assessment technologies offer the promise of dementia risk reduction and a more timely diagnosis of dementia, respectively. Cognitive training games may help reduce the lifetime risk of dementia by helping to build cognitive reserve, whereas cognitive assessment technologies offer the opportunity for a more convenient approach to early detection or screening.

Objective: This study aims to elicit perspectives of potential end users on factors related to the acceptability of cognitive training games and assessment technologies, including their opinions on the meaningfulness of measurement of cognition, barriers to and facilitators of adoption, motivations to use games, and interrelationships with existing health care infrastructure.

Methods: Four linked workshops were conducted with the same group, each focusing on a specific topic: meaningful improvement, learning and motivation, trust in digital diagnosis, and barriers to technology adoption. Participants in the workshops included local involvement team members acting as facilitators and those recruited via Join Dementia Research through a purposive selection and volunteer sampling method. Group activities were recorded, and transcripts were analyzed using thematic analysis with a combination of a priori and data-driven themes. Using a mixed methods approach, we investigated the relationships between the categories of the Capability, Opportunity, and Motivation–Behavior change model along with data-driven themes by measuring the ϕ coefficient between coded excerpts and ensuring the reliability of our coding scheme by using independent reviewers and assessing interrater reliability. Finally, we explored these themes and their relationships to address our research objectives.

Results: In addition to discussions around the capability, motivation, and opportunity categories, several important themes emerged during the workshops: family and friends, cognition and mood, work and hobbies, and technology. Group participants mentioned the importance of functional and objective measures of cognitive change, the social aspect of activities as a motivating factor, and the opportunities and potential shortcomings of digital health care provision. Our quantitative results indicated at least moderate agreement on all but one of the coding schemes and good independence of our coding categories. Positive and statistically significant ϕ coefficients were observed between several coding themes between categories, including a relatively strong positive ϕ coefficient between capability and cognition (0.468; $P<.001$).

Conclusions: The implications for researchers and technology developers include assessing how cognitive training and screening pathways would integrate into existing health care systems; however, further work needs to be undertaken to address barriers to adoption and the potential real-world impact of cognitive training and screening technologies.

International Registered Report Identifier (IRRID): RR2-10.1007/978-3-030-49065-2_4

(JMIR Serious Games 2022;10(2):e32489) doi:10.2196/32489

KEYWORDS

dementia; cognitive assessment; cognitive training; serious games; patient and public involvement

Introduction

Background

The World Health Organization has estimated the number of people living with dementia worldwide to be approximately 55 million and the total societal cost of dementia to be approximately US $2.8 trillion [1]. These figures are expected to rise with the increase in life expectancy; however, over one-third of these cases are thought to be preventable by reducing modifiable risk factors [2]. As such, in recent years, there has been growing interest in the application of digital technologies designed to promote healthy lifestyles, particularly in the prevention, management, and mitigation of dementia [3-6]. There is a growing evidence base indicating that healthy lifestyle choices such as exercising, maintaining social engagement, reducing smoking, and engaging in cognitively stimulating activities may help to reduce the overall lifetime risk of dementia [7,8]. Lifestyle factors such as occupation, leisure activities, and educational attainment may help to build *cognitive reserve*, which could help to mitigate or modify the clinical expression of dementia despite underlying neuropathology [9]. Other evidence suggests that certain cognitive domains may be amenable to training [10,11]. Despite this, the extent to which these technologies may help improve cognition or reduce the overall lifetime risk of dementia is debated, and the evidence remains inconclusive [12-15].

Concurrently, cognitive screening offers the potential for early intervention and support but can often lead to frustration or stress and affect patient dignity [16]. There are still challenges around the acceptability and uptake of digital technologies, which are designed to improve or assess cognition [17], including concerns about privacy, decision-making, dignity, and liberty in the use of technology for people living with dementia [18].

The inclusion of gamified cognitive training interventions could enhance motivation, and positive mood and improve assessment [19]. Increasing engagement and adherence to serious games for health often relies on the development and implementation of motivating elements [20], whereas some argue that the use of extrinsic incentives (eg, rewards and penalties) may hinder intrinsic motivation in the long term, potentially affecting self-confidence [21,22].

Lifestyle advice aimed at reducing the modifiable risk factors of dementia and cognitive training and screening regimes falls under the broad category of public health interventions [23], which depend on behavior change to be effective. A widely used model used to understand behavior change interventions is the Capability, Opportunity, and Motivation–Behavior (COM-B) model developed by Michie et al [24]. The model identifies capability, motivation, and opportunity as the primary drivers of behavior. Using this framework, public health behavior change interventions can be systematically evaluated, and appropriate strategies can be developed to maximize the effectiveness of interventions.

Interventions designed to form part of a person's daily routine need to ensure that there is involvement from patients and the public during the testing and evaluation of cognitive technologies to ensure that the technology is being used as intended and increase uptake [25].

Patient and public involvement (PPI) is advocated in health care research for many reasons; it may help generate research opportunities, identify research priorities, lead to better research methodology, and improve the communication and application of findings [26]. PPI research is conducted *with* or *by* patients and the public rather than being conducted *on* them, *about* them, or *for* them [27,28]. There has been growing acknowledgment of the importance of involving patients and the public in research that places patients and their experiences at the center. PPI research is of particular importance to dementia research in which it is understood to contribute to the quality, relevance, and ethical conduct of research [29].

Aims and Objectives

This paper reports on PPI activity that formed a part of an industry-academic partnership, Alzheimer's Disease: Detect Prevent, led by Brain+, a Danish commercial app developer. Brain+ is developing a suite of app-based technologies designed to help reduce modifiable risk factors of dementia by cognitively stimulating gaming and lifestyle coaching. In addition, it is introducing an element of detection of cognitive deficits through a working memory test. A series of PPI workshops were held to discuss the potential benefits of cognitive training games and assessment technologies related to cognition and capture contextual opinions about the dementia journey, such as support after diagnosis, the role of family and other social relationships, and the practical and ethical factors that arise in the adoption of new health technologies.

We aimed to answer the following research questions:

1. How can cognition be measured in a way that is acceptable and meaningful to people with the experience of dementia?
2. What are the barriers to and facilitators of the adoption of digital cognitive training games and assessment technologies?
3. What factors may affect motivation to use serious games for cognitive assessment and training?
4. What are the potential benefits, drawbacks, and risks of assessment technologies for cognitive impairment, and how may such tools fit within the existing health care infrastructure?

Methods

Workshop Procedure

Workshops were conducted physically within conference rooms at the University of Nottingham. The workshops lasted between 10:30 AM and 2 PM for each of the days they were run. Lunch and regular breaks were provided to our participants, and each of our workshops was divided into 2 sessions.

In total, four workshops were conducted between October 2019 and January 2020, each with a focus on a different topic relevant to cognitive training and screening technologies:

1. What counts as meaningful improvement?
2. Learning and motivation
3. Trust in digital diagnosis
4. Barriers to digital inclusion

These topics were developed by the researchers and agreed upon by the research and commercial partners involved in the wider project consortium. The topic of *meaningful improvement* facilitated the exploration of appropriate and patient-centered outcome measures for a concurrent feasibility trial of cognitive training. The topic of *learning and motivation* was chosen for the purposes of investigating motivational strategies and the perceived potential of cognitive training, how at-risk individuals and those with lived experience conceptualize learning *cognitive reserve*, and the importance of cognitively taxing activities to delay the onset of dementia. Our third session on *trust in digital diagnosis* allowed for the elicitation of factors that affect trust in digital diagnosis and screening tools, including the social, environmental, and legislative contexts in which screening and the process of diagnosis occur, as well as attitudes and concerns surrounding the storage, handling, and sharing of personal data. In our concluding session, we sought to identify real-life, contextual, and social factors that may influence technology use with the intention of developing strategies to reach otherwise underrepresented and hard-to-engage populations. A summary of our workshop protocols can be found in Multimedia Appendix 1 and the *Materials* section.

Materials

All participants, facilitators, and researchers were supplied with name badges to facilitate an informal first-name basis tone for the discussions. Voice recorders were used to record the sessions for subsequent transcription. Two of our workshops included presentations about the topic (1 and 2), and the *trust in digital diagnosis* workshop involved an activity that included the use of UnBias Fairness Toolkit Ideation cards [30].

Recruitment

Participants were recruited via Join Dementia Research (JDR), a UK-based service that connects people living with dementia and caregivers with researchers for the purposes of conducting dementia research. JDR provides a list of names and contact details based on a set of criteria selected by the researcher. As we were interested in recruiting both people living with dementia and their carers, we decided to target carer-patient dyads. We also wished to discover issues around diagnosis; therefore, dyads without a formal diagnosis of dementia were excluded. Finally, we wanted to investigate issues surrounding the potential of cognitive assessments and cognitive training to facilitate early intervention and recruited participants living with mild dementia for whom cognitive assessment and training may serve a therapeutic or rehabilitative purpose. Using purposive selection, 18 dyads presented on the JDR website results were contacted for recruitment, from which 22% (4/18) agreed to be enrolled. Subsequently, 2 carer-patient dyads took part: one living with dementia participated alone without their partner, and another took part as an individual carer after losing his partner to dementia.

Participants were compensated in line with the INVOLVE guidelines [28], which also included transportation costs. Upon attending each workshop, participants were invited to subsequent workshops.

Participants

Sessions comprised 6 participants (ie, n=3, 50% carers and n=3, 50% people living with dementia), 2 or 3 group facilitators from the local involvement team, and members of the research team (between 2 and 4 in different sessions). Workshop participants were *experts by experience* because of their lived experience and first-hand knowledge of dementia. Within the participant group in the workshops, there was an equal gender split: the 2 carer-patient dyads were married couples, the individual with dementia was female, and the individual carer was male. Of the people living with dementia, one had working-age Alzheimer, one had vascular dementia, and the third had mixed (Alzheimer with vascular dementia), all describing the severity of their symptoms as mild. They were all aged between 57 and 76 years. Although small, the group was culturally and ethnically diverse, with a range of cultural backgrounds and experiences.

The involvement team facilitators had personal lived experiences and shared awareness of a range of mental and cognitive health conditions and were knowledgeable about digital technologies and the processes of coproduction. As such, they contributed additional knowledge and opinions captured in the focus groups along with those of the enrolled participants. During each session, they ensured that the enrolled participants were comfortable and were there to assist them should they have any difficulty. Facilitators were also involved in reviewing the content of the workshops, proofreading documents and instructions sent ahead of each session, and sense checking the analysis.

Ethics Approval

Ethical approval was granted by the Faculty of Medicine and Health Sciences Research Ethics Committee (approval number 333-1906). Participants were made aware of the aims and methods of the study and had a chance to ask questions before signing the consent form. All participants met the definition of capacity in line with the legal definition given under the Mental Capacity Act of 2005.

Analysis

Qualitative Analysis

Each of the audio-recorded sessions was transcribed via a third-party contractor before being split into meaningful excerpts, which was achieved by attempting to split the entire transcript into the smallest possible excerpts, which were intelligible without any accompanying text. As such, no two excerpts referred to the same part of the transcript. This created a total of 605 such excerpts.

A 2-stage approach to thematic analysis was used [31,32]; excerpts were initially coded with a priori (theory-driven) themes based on the COM-B behavior change model proposed by Michie et al [24]. This model targets 3 areas of importance in behavior change: capability (the individual's psychological and physical capacity to engage in the activity), motivation (processes that energize and direct behavior; both planned and habitual), and opportunity (factors that lie outside of the individual, which make behavior possible, prompt it, or present a barrier to it). Once COM-B coding was completed, an additional data-driven coding scheme was devised based on the content of the sessions and issues pertinent to the research project.

Finally, the qualitative validity of our results was assessed by asking the PPI facilitators to comment on our findings before publication, with the aim of increasing the dependability of our results.

Quantitative Analysis

Both coding schemes were quantitatively evaluated by another reviewer (coauthor), who coded the excerpts according to the coding scheme during our initial analysis using a small subsection of our excerpts (64/605, 10.6%), which were assessed for consistency using proportional agreement and Cohen κ. This helped ensure the validity and reliability of our findings.

We also looked at the intersections of excerpts coded in our themes, both to ensure that coding schemes within respective coding categories were sufficiently independent, as well as to analyze the relationships between themes of separate coding categories. Best practice dictates that where possible, coding categories should be exclusive [31], indicating that there should be little overlap between separate coding schemes. However, because the analysis used 2 separate coding schemes (one theory driven and another data driven), themes from different schemes do not need to meet this criterion. The ϕ coefficient calculations were conducted to determine the correlation between each theme. This served 2 distinct purposes. First, coefficients for themes *within* respective coding schemes indicate the extent to which each of the thematic categories could be considered independent and exclusive categories. Strong positive correlations would potentially indicate that the themes were not sufficiently independent.

Second, we also wanted to determine the relationship between the various thematic categories *between* the 2 separate frameworks to understand how our data-driven categories were related to broader theoretical determinants of behavior expressed in the COM-B model. Positive correlations indicated the extent to which excerpts matched 2 separate themes simultaneously, whereas negative correlations indicated a decreased likelihood of a single excerpt being coded for both of those categories.

Results

Overview

A description of the main research findings can be found in the following sections; more information about the themes and how the excerpts were coded can be found in Multimedia Appendices 2 and 3, along with the coding criteria used by both analysts and example excerpts. Quotations are given using *P* to denote a person with dementia, *C* to denote a carer or spouse, and *F* to denote a facilitator having lived experience of mental or cognitive health problems.

COM-B Themes

Overview

Analysis of the transcripts using the COM-B model revealed several interesting findings. Capability was often viewed with respect to self-care and the tasks of everyday living. Group attendees generally agreed that capability diminishes over time and that a good intervention was one that would lessen the rate of decline and allow people to live independent lives, where they could pursue their hobbies and passions. The theme of opportunity typically revealed several complex and interrelated issues. The theme of opportunity included the health care system and diagnostic pathway, the activities and support available for people living with dementia, cultural and social norms, and how technology may interact with all of the above. Finally, motivation was tied into several areas, including the motivation to play cognitively challenging games and develop new skills later in life and, finally, the importance of appropriate and timely feedback that did not overburden users of cognitive games.

Capability

People living with dementia were very much aware of their diminished capabilities because of their cognitive impairment.

Impact on Everyday Life

Cognitive difficulties can affect every aspect of a person's life, including socializing, work commitments, hobbies, social activities, and activities of daily living. A participant said, "I found out that I had Alzheimer's because I couldn't do my job any more as a lawyer" and that their "whole life has changed" (P3).

Fluctuating Capability

Attendees also mentioned how capability was not static but could be affected by external factors and could fluctuate, thereby affecting people's abilities to engage in meaningful tasks related to hobbies and the activities of daily living:

> *It might just be something on that day, there's an external factor that's going to have a great deal of influence on that day.* [F1]

Maintaining Active Lifestyles

However, attendees agreed upon the importance of maintaining an active lifestyle to support quality of life and independence. Participants were eager to retain and even build upon their existing capacities, although they typically viewed their ability to learn and retain information as gradually decreasing over time. It was important for our participants to have meaningful things to look forwards to in their lives:

> *...you need to try and carry on as normal because if you start taking things away from people, they start to feel bad.* [P1]

Risk and Independence

Several excerpts also discussed capability with regard to risk and independence; for example, certain activities were thought to enable people to live independently but could be dangerous when performed by someone with cognitive difficulties. Safety in the home was a concern expressed by one carer in the group, whereas there was an understanding that the capability to remain safe changes over time, and it was indicated that over the course of the progression of dementia, safety may become prioritized over independence:

> *Appliances are very dangerous, even falling down the stairs can be dangerous.* [C3]

Opportunity

In addition to the definition of opportunity in its positive sense (described previously), we also considered the barriers that lie outside the individual, such as the absence of opportunities, restrictions, coercion, and social pressure.

Services and Support

The discussion around opportunity included the availability of local social activities designed for people living with dementia, such as *Singing for the Brain*, and personal activities with friends and family, and participants mentioned several areas where technology may be able to facilitate these tasks, such as navigation tools and personal reminders. They also recounted contact with health services where opportunity was expressed negatively as a lack of support following diagnosis:

> *I think that when people are diagnosed with cognitive impairment it is quite a shock right, you do not expect it will happen to you even though you know it happens. Then there is no support to help you...* [P1]

The Impact of Cultural Norms

The same participant stressed that an activity they had managed to find and attended with their partner had an assumed shared history that was not relevant for them, having moved from outside the area:

> *The memory class we didn't find helpful at all because they were concentrating on the past in Nottingham.* [P1]

In addition, people felt that cultural norms had, in the past, limited their opportunities and that this belief could be internalized, limiting their own expectations and self-belief. For instance, a person living with dementia said the following about their childhood:

> *The girls were not really encouraged as much as the boys were.* [P1]

Other negative evaluations, even at a young age, were thought to have a negative impact on self-perception:

> *I think there were often subliminal negative reports at school, this boy will never be very good at whatever there was quite a lot of that in school reports, which might have held sway.* [C3]

A group facilitator agreed as follows:

> *If you're told often enough that you can't do something, you start to believe it.* [F1]

Opportunities Afforded by Technology

New opportunities for support, monitoring, and health care services facilitated by technology were also mentioned. The use of technologies and internet-based platforms was seen as "an opportunity to be proactive" (F3). However, participants were not entirely uncritical of additional opportunities afforded by monitoring technologies for people with dementia and showed some concern about restrictive measures that may hinder privacy and lead to concern about how their data might be used:

> *You're being compulsorily monitored, you know, like being tagged almost. It can be an invasion of human rights, I think.* [F1]

Cognitive Screening

Some also expressed concern about being told that they were at risk from a certain condition with no known cure and questioned the need to tell people about something that would cause them a great deal of anxiety. Although cognitive screening affords new opportunities for early diagnosis, the potential for early cognitive assessment was not necessarily seen as a universal good:

> *They're now finding ways to check people to say, "This person is possibly going to have it in twenty years' time." Well, do you tell them twenty years beforehand?* [C3]

Motivation

Many motivating factors were discussed, both in relation to controlling the symptoms of dementia and regarding ambitions more generally.

Acquiring New Skills

Many group attendees acquired new skills later in life, including painting, gardening, cooking, use of technology, and being involved in dementia research:

> *I think that older people that have learnt a computer later in life, I feel my brain is better from all the learning that I have done with a computer.* [F1]

Cognitive health and cognitive training were viewed in a similar, positive way as physical health, and exercise and cognitive training were considered analogous to "going out for a jog" (C2).

Challenging Oneself

Attendees were particularly interested in being challenged by cognitive games, wanted to know whether skills could be improved, thought that having short-term attainable goals was a good motivational factor, and valued resilience in the face of adversity. On numerous occasions, workshop participants mentioned that they knew their cognitive skills would deteriorate over time but felt no point dwelling on the fact and insofar as possible wanted to *carry on* and live their lives to the fullest extent possible:

> *One thing I have learnt is that to do the things on my own as much as possible is to carry on to be who I am until the day I can't do it.* [P1]

Gaming Features

Competitive games were also seen as motivational, playing against either friends and relatives or other people worldwide. Activities that provided a level of challenge were seen as both intrinsically motivating factors and opportunities for learning. The participants seemed to value positive feedback given by mobile phone apps:

> *I think positive feedback is good and I feel because I am doing this diet app at the moment and every day you get a little quiz and if I get 5 out of 5, I feel really good...* [C2]

Notification Fatigue

However, although feedback and other motivational strategies were seen as valuable, if these were seen to place too much of an expectation on people, it was thought that this may have a demotivating impact in the long term. This could be reflective of the broader trend of extrinsic motivation undermining intrinsic motivation in some instances:

> *I have got an app for my iPad and every time I open it, it tells me, "Is your memory still good?" and then I want to say, "Get off!" Every time I open it it's there, "How is your memory doing?" It tells me every day, that's what I was saying. [laughter] I don't want to get rid of the apps completely, but I want it to stop telling me everything.* [P2]

Emergent Themes

Overview

Emergent themes were devised by considering the main recurrent subject topics and the particular exploratory purposes of our research. The details of the theme descriptions and example excerpts are available in Multimedia Appendix 3. The themes were technology; friends, family, and support; work and hobbies; health care system; and cognition. Performing the second round of analysis and coding the transcripts independently based on specific domain issues allowed for more issues to be identified, which closely aligned with our aims. For instance, the benefits and drawbacks of technology could be identified more easily. Participants saw the value of technology in helping to diagnose and support people living with dementia but were worried that technology may replace human-centered care and that this may have a negative impact on their overall experience because of a lack of support and a lack of holistic understanding. Family, friends, and support networks were seen as integral to a person's life, either from a spouse helping to detect changes in cognition and enabling them to live independently or from hobby interests or specific support groups. However, it should be noted that people living with dementia did not see themselves as only receiving support but considered the relationship reciprocal and saw value in being a source of support and providing for their friends and family through activities such as cooking and caring. Through this coding category, we were able to highlight that cognition was viewed in concrete and practical terms. Excerpts relating to the health care system revealed diverse experiences of the diagnostic pathway; many felt their care had been *disjointed* and left them feeling isolated.

Technology

A variety of views were expressed about technology, its potential role in improving quality of life, and the possibility of digital assessment methods for cognitive impairments. Most of our participants enjoyed cognitively stimulating activities, valued the objective measures of their performance on cognitive tasks, and appreciated the instant feedback afforded by technology.

Cognitive Offloading

Cognitive offloading strategies, such as diaries and reminders, were also mentioned, although some worried that this may actually lead to deskilling:

> *Even though we've lost our memories, there is still the ability to put the thing in the telephone and get the answer. Which I must say, we do a lot of!* [C3]

> *...the phone become our memory. I mean I used to have x number of telephone numbers in my head, and now I can just about remember mine. So, I think, in a way, you know, we're de-skilling ourselves, certainly de-skilling our memory.* [F1]

Ease of Use

Workshop attendees were interested and enthusiastic about understanding the possible benefits of technology for people living with dementia and were optimistic about technology becoming easier to use and more accessible to a wider variety of people:

> *I am curious as I want to see how the technology can possibly help somebody who is a real technophobe and my husband as well. I am really looking forward to seeing what might be forthcoming.* [C2]

Gaming as a Waste of Time

As mentioned previously, participants also discussed playing games on mobile devices and often enjoyed games that they considered a challenge or games that involved some mental stimulation. However, in general, computer games that were not seen as providing any additional benefit were viewed negatively by members of the group:

> *I can remember playing Angry Birds once and after about 10 goes I thought this is such a waste of time.* [P1]

Overall, views on technology were largely positive, and there was general optimism about how technology could improve the lives of people living with dementia, leading to better health outcomes overall. However, as mentioned previously, other views expressed concerns about privacy and deskilling.

Family, Friends, and Support

Family, friends, and close personal relationships were particularly important in managing cognitive impairment. Family and friends were seen as sources of support and enablement and were also those best placed to understand changes in cognition.

Detecting Cognitive Change

One of the participants indicated that their spouse was able to detect changes in cognition before they became aware of any cognitive issues:

> *My husband knew before me...* [P3]

However, it should be noted that those who are in contact with a person living with dementia every day may be less likely to notice gradual changes in cognition, and those who visit more infrequently would be better placed to notice differences:

> *It is difficult to tell really, I think if you are living with someone, and you are there all the time unless there is a massive change you don't notice... you only see somebody say once week or however often it is you might notice something more whereas the person living with them wouldn't.* [C3]

Support and Being Supported

Friends and family were also discussed in the context of helping facilitate activities and hobbies and providing emotional support. However, it should also be noted that this relationship was often seen as reciprocal, and people living with dementia also discussed providing for their friends and family through cooking and housework:

> *I cook for my daughter...* [P1]

Work and Hobbies

Overview

Hobbies and interests were very important to the participant groups. Several participants mentioned discovering new hobbies later in life, and many mentioned actively seeking out groups and activities such as *Singing for the Brain*:

> *I see a great difference in this as my wife was able to sing and she loved the singing side, which is very well known with dementia.* [C3]

Hobbies as Beneficial

Hobbies that were seen as providing additional benefits to either oneself or others were highly valued. Session attendees frequently mentioned puzzles such as number and word games on their mobile devices but were critical of activities they saw of little value to themselves or other people:

> *You've got to give yourself an interest. Gardening, I've found is ideal if you've got an allotment, you've got a community...* [C1]

> *I have been hanging on to that because I have been learning Spanish for 10 years.* [C2]

> *For me if it was actually shown to improve something it would make me do it.* [P1]

The advantages of asynchronous competitive games were also mentioned:

> *You can play it at any time so you don't have to sit and play I suppose like the normal game of scrabble, but I suppose you can just abandon it and come back to it later.* [C2]

Technology in the Workplace

Several (but not all) participants were employed when computers and other digital artifacts were introduced in their workplace. However, some mentioned that they were unable to keep up with the pace of change:

> *We...we had no choice at work, they brought in computers so anything you were booking in or out had to be done on a computer, so you had to learn that. I was never happy with them, but I could find my way round. The one thing I never liked was mobile phones and I still don't. I call them the curse of the twenty-first century!* [C1]

Health Care System

Attendees were, on the whole, very eager to take advantage of therapeutic activities offered by health care and third-party providers, although there was great variability in the support offered:

> *I think we were lucky because at our surgery as they teach in the practice, so we had loads of stuff coming through like, "This programme is on, that programme is on and are you interested?"* [C2]

> *"Go home." That's it. No support, nothing, just gave me the books and, "Go home."* [P1]

Digital technologies were seen as potentially transformative to health care systems. However, group members did not see technological developments within health care as being able to solve all issues related to dementia. They worried that the health care service was no longer integrated, which could lead to inconsistent treatment, a lack of support, and a breakdown of trust. In addition, there were mixed opinions about the possibility and accuracy of digital cognitive assessments without the presence of qualified clinicians:

> *a computer will be able to diagnose what we've got better than a surgeon or a doctor. That's the way it's moving.* [C3]

> *I just think it's dangerous to see the test and you in some sort of a vacuum as though that's all that there is.* [F1]

Cognition

Overview

Participants had an intuitive understanding of changes in their cognition and saw treatments and therapies as a way of reducing the rate of cognitive decline rather than lessening their symptoms:

> *It only goes one way. If it goes back up again that is not positive because I can assure you it goes that way.* [P1]

Functional Measures of Cognition

Objective, functional, and performance-based measures of cognition were seen as the most appropriate indicators of cognitive issues, and dyads thought that difficulty in learning new things might be an indicator of cognitive decline. There was a general group consensus that other people (rather than the person with dementia themselves) were better able to detect changes in cognition:

> *...he suddenly realised he had got the map upside down and that would never have happened, it was so out of character and there must of be lots of other little things, but I just thought there is something going on here that is not quite right.* [C2]

Diagnosis and Self-image

A diagnosis of cognitive impairment, dementia, or the risk of these could lead to a lack of confidence and increased anxiety:

> *Because if you're given numbers about yourself, and one day the numbers are twenty points lower, whatever those numbers are, if they're twenty points lower than they were the day before, that could possibly lead to anxiety.* [P2]

> *It robbed me of my confidence, and I do not know how to get it back still.* [P1]

Interrater Reliability

Coded excerpts were checked for interrater reliability initially by determining the proportion of observed agreement (OPA) and then finally by calculating Cohen κ: capability (OPA=0.762; Cohen κ=0.491), opportunity (OPA=0.683; Cohen κ=0.368), motivation (OPA=0.746; Cohen κ=0.451), technology (OPA=0.921; Cohen κ=0.838), friends, family, and support (OPA=0.841; Cohen κ=0.624), work and hobbies (OPA=0.794; Cohen κ=0.566), cognition (OPA=0.723; Cohen κ=0.408), health care system (OPA=0.905; Cohen κ=0.670).

Relationship Between Themes

Table 1 indicates that in general, there was a good level of independence between themes within the same category, with the highest positive ϕ coefficient between 2 themes within the same framework of 0.091 (technology and health care system). There were positive, statistically significant ϕ coefficients between themes belonging to different coding categories (capability and cognition; opportunity and technology; opportunity and health care system; motivation and technology; and motivation, work, and hobbies), indicating the extent to which the COM-B elements mapped onto our data-driven themes.

Table 1. The ϕ coefficient between themes to look for relationships.

Themes	Capability	Opportunity	Motivation	Technology	Friends, family, and support	Work and hobbies	Cognition
Opportunity	−0.231[a,b]	—[c]	—	—	—	—	—
Motivation	−0.229[a,b]	−0.279[a,b]	—	—	—	—	—
Technology	−0.234[b]	0.153[a,b]	.0815[d]	—	—	—	—
Friends, family, and support	−0.017	0.073	0.027	−0.122[a,d]	—	—	—
Work and hobbies	0.019	−0.061	0.147[b]	−0.237[a,b]	0.019[a]	—	—
Cognition	0.468[b]	−0.083[d]	−0.184[b]	−0.110[a,d]	−0.057[a]	−0.220[a,b]	—
Health care system	−0.219[b]	0.282[b]	−0.137[b]	0.091[a,d]	−0.151[a,b]	−0.333[a,b]	−0.034[a]

[a]Coefficients between same framework.

[b]*P*<.001.

[c]Not applicable.

[d]*P*<.05.

Discussion

Qualitative Findings

Several interesting findings came from the workshop sessions, with implications for technology designers and policy makers involved in developing serious games and cognitive assessments for people living with dementia.

Technology designers should be considerate of the fact that cognitive assessments might be stressful for some people who worry about cognitive decline. Presenting the results from cognitive assessments without the support of clinicians may cause undue worry about cognition; therefore, this finding agrees with earlier work on the experiences of cognitive screening [33].

Our workshops indicated that people had widely differing experiences of the diagnostic pathway, with some feeling supported and others feeling isolated. Inequalities in the provision of care and care outcomes in dementia are known; however, their determinants are underresearched [34].

An unexpected diagnosis of dementia or cognitive impairment can lead to uncertainty and raise questions regarding identity and autonomy. Previous research indicates that this experience is not uncommon and is often shared by partners of those who have been diagnosed [35]. Therefore, it is vitally important that cognitive screening technology offers options for further support and provides resources that may help counter the potential feeling of disempowerment. Digital technology may offer a way of countering some of these existing inequalities in dementia care; however, equality itself is not guaranteed by the use of digital technologies [36].

Confirming a dementia diagnosis is often a lengthy process [37]; screening technology developers should consider the role they play in the overall diagnostic pathway and how screening may fit into existing health care systems. Technology designers and policy makers must ensure that the results of cognitive assessments taken on at-home devices signpost or link to appropriate clinical support and provide an environment that guides users through the appropriate next steps.

Participants were in favor of using performance-based and objective measures of cognition but raised concerns about the results of these assessments being taken in isolation. However, older patients who are hospitalized and undergo cognitive screening often report being unaware of the significance of screening tests and feel stressed because of the pressure to perform, sometimes evoking feelings of shame and irritation [16].

People were also generally worried that their results would only be a snapshot of their performance at the moment and not necessarily indicative of their cognition in general, and their results could be affected by a variety of factors, not necessarily indicative of dementia. Conversely, it is known in dementia, particularly Lewy bodies and Alzheimer, that cognition does indeed fluctuate, and this may have an impact on clinical diagnosis [38]. Therefore, it is advisable that any cognitive screening or training technologies attempt to take a more holistic and individualized approach, which takes account of personal circumstances and cognitive fluctuation, ensuring that the screening takes into consideration a broad range of factors and not just a single example of performance at any particular time.

Regarding cognitive training aspects, participants were particularly motivated to engage in activities that they felt would be of some benefit to themselves or others; this included cognitively stimulating games and apps. Despite this, participants expressed skepticism about gaming in general. Therefore, cognitive games should relate to and emphasize the potential real-world applications of cognitive training, which may be applicable to other areas of life and should arguably be presented as distinct from games. Equally, people living with dementia and their carers placed great importance on their ability to perform practical tasks as an indicator of cognition. Therefore, cognitive assessments that more easily relate to the activities of daily living (such as the instrumental activities of daily living [39]) may be seen as more acceptable to people at risk of dementia.

When discussing their own hobbies and motivations, the social aspect of pastimes and the importance of community building were frequently mentioned. Most attendees had taken up new hobbies and skills at an older age, often viewing the acquisition of new skills as a way of keeping their mind and body active but also to form new community groups in retirement. Several participants mentioned gaining enjoyment from the socially competitive element of games and the idea of playing asynchronous games against people they knew. Participants mentioned that turn-based competitive games could fit into their own lifestyles without the pressure of an immediate response. App developers may wish to further emphasize and develop the community-building aspects of cognitive training technologies. However, it should be noted that communication abilities may decline with the progression of dementia [40,41]; hence, the social aspects of gaming may not be as relevant to those with severe clinical symptoms.

Quantitative Findings

When considering the quantitative aspects of our work, we found that the theory-driven analysis framework (COM-B) resulted in less overall interrater reliability than our data-driven framework. Although the framework was a useful taxonomy for understanding the broad theoretical determinants of behaviors related to brain-training and cognitive screening technologies, it resulted in a lower overall interrater agreement. We considered interrater reliability as measured by an observed proportional agreement of ≥0.6 to be low but acceptable for our purposes, 0.7 to be good interrater reliability, and >0.8 to be very good interrater reliability. Using Cohen κ benchmarks put forward by Landis and Koch [42], we observed that agreement on the technology theme was *almost perfect* and that there was substantial agreement on coding of the health care system and friends and family themes; moderate agreement on capability, motivation, and work and hobbies themes; and only fair agreement on the opportunity theme. The lower interrater reliability for the COM-B categories applied to data versus data-driven themes is not surprising.

With reference to the naming conventions suggested by Rea and Parker [43] for analyzing the strengths of association in cross-tabulated data, we took <0.1 to indicate a negligible association, ≥0.1 and <0.2 to indicate a weak association, ≥0.2 and <0.4 to indicate a moderate association, and ≥0.4 and <0.6 to indicate relatively strong association. Capability was most strongly related to cognition (0.468), and this was highly significant (P<.001), indicating a relatively strong relationship between participants' conceptions of the two, a relationship that is well established in the literature [44,45]. Opportunity was most strongly related to the health care system (0.282) and moderately associated; opportunity also had a weak association with technology (0.153), and these relationships were highly significant (P<.001), further indicating that participants viewed technology as being able to provide additional opportunities. Motivation was most strongly related to work and hobbies (0.147; P<.001), emphasizing the importance of hobbies and leisure activities as motivating factors in the lives of our participants, although opportunity showed a lower correlation with any single emergent theme than either capability or opportunity. This may be because capability is often seen as more individualistic (*within* the person), and motivation is related more to the habits, preferences, and long-term goals of individuals. The strongest positive relationship between the friends and family subtheme and any of the COM-B themes was with opportunity, although this relationship was negligible (0.073) and insignificant. This may be because friends and family were often discussed in the context of health care or shared activities, providing part of a supportive role rather than being the subject of the conversation. Overall, we observed a good level of independence between themes within the 2 coding categories; however, we observed a statistically significant relationship between the health care system and technology themes, although the strength of this association was negligible (0.091).

Strengths

Using the method described for the PPI workshops, we found that participants were able to talk about the issues presented to them in great detail, with reference to their own personal lives and those close to them. Patient and carer dyads provided added value in understanding by allowing a shared perspective of similar incidents [46], which was a view shared by the researchers, facilitators, and participants involved in this study. In practice, this often meant that carers would add context or fill in details. Workshop attendees were supportive of each other and were willing to openly share their experiences and attitudes about living with dementia, expressing a wide range of views on issues surrounding diagnosis and the potentially transformative impact of technology. Before the workshops, facilitators helped create and moderate the workshop agenda and assisted in reviewing the wording of communications with workshop participants. Facilitators also ensured that all participants were comfortable, both physically and emotionally, throughout the sessions and that everyone was able to contribute.

Using a 2-stage coding system, we were able to map our data-driven themes onto what we considered the key determinants of behavior; however, we acknowledge that even our *data-driven* themes were also, in part, influenced by our research considerations.

Finally, facilitators sense checked our work to improve the dependability of our results [47]. Interrater reliability checks served to establish the credibility and confirmability of our findings [47].

Limitations

Although the participants reflected a range of different cultural backgrounds, the findings from relatively small workshops may not be generalizable to the population as a whole; caution must be taken in regarding these findings as universally applicable. Although the number of participants was relatively small, we believe that a smaller group setting was beneficial in allowing each participant to express themselves without fear of interruption. The data set that we gained from these workshops also reflected a variety of attitudes, opinions, and circumstances. We do not claim to have reached data saturation on these issues, and there may be other themes related to each of our workshop topics that are yet to be identified. However, we believe that the depth of our discussions (over 10 hours 24 minutes of interview materials), the variety within our workshops (4 separate topics, each split into 2 sessions), and the interpretive status of evidence (confirmed by our checks on interrater reliability) demonstrate that the materials have met data adequacy [48] and that our findings are a reflection of genuine attitudes held among key user and patient groups relating to cognitive training and screening technologies.

Owing to practical considerations of room size and considerations of allowing all participants to contribute to the discussion, we were limited in the number of additional participants that we could recruit. Although we were prepared to have different participants on the 4 days, we encouraged participants who had been involved in earlier workshops to return to subsequent ones; as such, the same group attended all of them. Working with a different group in each of the 4 workshops might have increased the breadth of the viewpoints. However, we believe that this was ultimately a strength of our approach, as participants developed a rapport with each other during the sessions and were consequently willing to share personal experiences about sensitive topics in our later sessions.

The used recruitment method may have led to a selection bias that favored those already interested in dementia research and, hence, people who may be more likely to view these interventions favorably. Involving carers in the sessions raised the possibility of them speaking on behalf of those living with dementia. It could be argued that recruiting dyads limited the representativeness and transferability of our sample. However, in practice, we found that carers would often provide additional contextual information, offer a different perspective on a similar incident, or else provide an example that could be expounded upon. In general, carers were supportive and helped explain things that their spouses may have been struggling with, as well as helping with practical issues such as transportation to the venue.

Areas for Future Work

Research on serious games for people living with or at risk of cognitive impairment needs to further explore the consequences of technology in relation to the quality of life, digital rights, and overall well-being to facilitate better usability and acceptability [49]. Although we showed participants a brain-training application to contextualize the discussion, the workshop attendees had not themselves been recently engaged in a cognitive training regime and, therefore, could only talk in general terms about the concept and how it would relate to their own experiences. Therefore, future work should aim to elicit the responses of people at risk of dementia who have experienced digitized cognitive training and assessments to garner more specific issues that are likely to emerge because of these technologies.

Conclusions

Motivation and user attitudes toward cognitive gaming is an underresearched area, and our workshop study, with analysis using the COM-B model and data-driven themes, has revealed a variety of opinions about both cognitive training and digital assessment technologies. More broadly, participants were able to express their opinions well on the opportunities and potential shortcomings of the digital health care provision. Potential facilitators included close support networks as a way of increasing motivation in meaningful activities in the context of cognitive training and, for cognitive assessments, the provision of instant feedback. The growing acceptability and use of technology facilitate both of these activities. End user perceptions of potential benefits included the ability of timely and accurate diagnosis; however, a potential shortcoming was that digital assessments may not take into account the context or the fluctuation of cognition.

This study identified some negative opinions regarding playing games for cognitive health. A potential solution is the enhancement of real-world, social, or community aspects of cognitive gaming. Regarding assessment technologies, concerns were raised about the lack of integration and consistency within the health care system and the lack of support following diagnosis. There were mixed opinions about the utility of cognitive assessment technologies for identifying dementia risk. There were also concerns about privacy and potential misuse of personal data.

Further user-centered research, including PPI activities, will help optimize the design of technologies that promise to improve cognitive health and well-being.

Acknowledgments

The authors would like to thank the workshop participants and group facilitators who contributed to this research. This work was supported by the European Commission Horizon 2020 EIC-FTI-2018-2020 under grant 820636, coordinated by Brain+ ApS (Denmark). The research reported in this paper was also supported by the National Institute for Health and Care Research (NIHR) MindTech MedTech Co-operative and the NIHR Nottingham Biomedical Research Centre. The views represented are those of the authors alone and do not necessarily represent the views of the Department of Health and Social Care in England, the National Health Service, or the NIHR.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Workshop activities and materials.
[DOCX File , 1126 KB - games_v10i2e32489_app1.docx]

Multimedia Appendix 2
The Capability, Opportunity, and Motivation–Behavior model themes.
[DOCX File , 24 KB - games_v10i2e32489_app2.docx]

Multimedia Appendix 3
Emergent themes.
[DOCX File , 23 KB - games_v10i2e32489_app3.docx]

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2021-11-26]
2. Rolandi E, Zaccaria D, Vaccaro R, Abbondanza S, Pettinato L, Davin A, et al. Estimating the potential for dementia prevention through modifiable risk factors elimination in the real-world setting: a population-based study. Alzheimers Res Ther 2020 Aug 07;12(1):94 [FREE Full text] [doi: 10.1186/s13195-020-00661-y] [Medline: 32767997]
3. Ben-Sadoun G, Manera V, Alvarez J, Sacco G, Robert P. Recommendations for the design of serious games in neurodegenerative diseases. Front Aging Neurosci 2018;10:13 [FREE Full text] [doi: 10.3389/fnagi.2018.00013] [Medline: 29456501]
4. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. Eur Heart J 2016 May 07;37(18):1428-1438 [FREE Full text] [doi: 10.1093/eurheartj/ehv770] [Medline: 26873093]
5. Cudd P, Bolton E, Gallant Z, Greasley P. The person living with dementia, their carer and their digital technology. In: Proceedings of the 12th biennial European conference of the Association for the Advancement of Assistive Technology. 2013 Presented at: 12th biennial European conference of the Association for the Advancement of Assistive Technology; Sep 19-22, 2013; Vilamoura, Portugal. [doi: 10.1108/JAT-10-2014-0022]
6. Gibson G, Dickinson C, Brittain K, Robinson L. The everyday use of assistive technology by people with dementia and their family carers: a qualitative study. BMC Geriatr 2015 Jul 24;15:89 [FREE Full text] [doi: 10.1186/s12877-015-0091-3] [Medline: 26205957]
7. Larson EB, Wang L, Bowen JD, McCormick WC, Teri L, Crane P, et al. Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. Ann Intern Med 2006 Jan 17;144(2):73-81 [FREE Full text] [doi: 10.7326/0003-4819-144-2-200601170-00004] [Medline: 16418406]
8. Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. J Int Neuropsychol Soc 2002 May 16;8(3):448-460. [doi: 10.1017/s1355617702813248]

9. Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol 2012 Nov;11(11):1006-1012 [FREE Full text] [doi: 10.1016/S1474-4422(12)70191-6] [Medline: 23079557]
10. Bonnechère B, Langley C, Sahakian BJ. The use of commercial computerised cognitive games in older adults: a meta-analysis. Sci Rep 2020 Sep 17;10(1):15276 [FREE Full text] [doi: 10.1038/s41598-020-72281-3] [Medline: 32943742]
11. Wang G, Zhao M, Yang F, Cheng LJ, Lau Y. Game-based brain training for improving cognitive function in community-dwelling older adults: a systematic review and meta-regression. Arch Gerontol Geriatr 2021;92:104260. [doi: 10.1016/j.archger.2020.104260] [Medline: 32980574]
12. Simons DJ, Boot WR, Charness N, Gathercole SE, Chabris CF, Hambrick DZ, et al. Do "Brain-training" programs work? Psychol Sci Public Interest 2016 Oct;17(3):103-186. [doi: 10.1177/1529100616661983] [Medline: 27697851]
13. Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. Cochrane Database Syst Rev 2013 Jun 05(6):CD003260 [FREE Full text] [doi: 10.1002/14651858.CD003260.pub2] [Medline: 23740535]
14. Fanfarelli J. Games and dementia: evidence needed. In: Video Game Influences on Aggression, Cognition, and Attention. Cham: Springer; 2018:163-171.
15. Harvey PD, McGurk SR, Mahncke H, Wykes T. Controversies in computerized cognitive training. Biol Psychiatry Cogn Neurosci Neuroimaging 2018 Nov;3(11):907-915 [FREE Full text] [doi: 10.1016/j.bpsc.2018.06.008] [Medline: 30197048]
16. Krohne K, Slettebø A, Bergland A. Cognitive screening tests as experienced by older hospitalised patients: a qualitative study. Scand J Caring Sci 2011 Dec;25(4):679-687. [doi: 10.1111/j.1471-6712.2011.00878.x] [Medline: 21362007]
17. Bureš V, Čech P, Mikulecká J, Ponce D, Kuca K. The effect of cognitive training on the subjective perception of well-being in older adults. PeerJ 2016;4:e2785 [FREE Full text] [doi: 10.7717/peerj.2785] [Medline: 28028465]
18. Bennett B. Technology, ageing and human rights: challenges for an ageing world. Int J Law Psychiatry 2019;66:101449. [doi: 10.1016/j.ijlp.2019.101449] [Medline: 31706404]
19. Westera W. Why and how serious games can become far more effective. Educ Technol Soc 2019;22(1):59-69 [FREE Full text]
20. 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 Jul 14;24(5):453-498. [doi: 10.1007/s11257-014-9149-8]
21. Benabou R, Tirole J. Intrinsic and extrinsic motivation. Rev Econ Stud 2003 Jul;70(3):489-520. [doi: 10.1111/1467-937X.00253]
22. Lepper MR, Corpus JH, Iyengar SS. Intrinsic and extrinsic motivational orientations in the classroom: age differences and academic correlates. J Educ Psychol 2005;97(2):184-196. [doi: 10.1037/0022-0663.97.2.184]
23. Alzheimer's Society's view on public health, prevention and dementia. Alzheimer's Society. URL: https://www.alzheimers.org.uk/about-us/policy-and-influencing/what-we-think/public-health-prevention-dementia [accessed 2021-12-13]
24. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
25. Hassan L, Swarbrick C, Sanders C, Parker A, Machin M, Tully MP, et al. Tea, talk and technology: patient and public involvement to improve connected health 'wearables' research in dementia. Res Involv Engagem 2017;3:12 [FREE Full text] [doi: 10.1186/s40900-017-0063-1] [Medline: 29062537]
26. Brett J, Staniszewska S, Mockford C, Herron-Marx S, Hughes J, Tysall C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expect 2014 Oct;17(5):637-650 [FREE Full text] [doi: 10.1111/j.1369-7625.2012.00795.x] [Medline: 22809132]
27. Pickett J, Murray M. Editorial: patient and public involvement in dementia research: setting new standards. Dementia (London) 2018 Nov;17(8):939-943. [doi: 10.1177/1471301218789290] [Medline: 30373464]
28. Public involvement in clinical trials. NIHR Involve. URL: http://www.invo.org.uk/posttypepublication/public-involvement-in-clinical-trials/ [accessed 2021-11-13]
29. Gove D, Diaz-Ponce A, Georges J, Moniz-Cook E, Mountain G, Chattat R, European Working Group of People with Dementia. Alzheimer Europe's position on involving people with dementia in research through PPI (patient and public involvement). Aging Ment Health 2018 Jun;22(6):723-729. [doi: 10.1080/13607863.2017.1317334] [Medline: 28513210]
30. UnBias Fairness Toolkit. UnBas. URL: https://unbias.wp.horizon.ac.uk/2018/09/06/unbias-fairness-toolkit/ [accessed 2021-07-25]
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
32. Harrington K, Brown M, Pinchin J, Sharples S. Decision-making within missing person search. Cogn Tech Work 2018 Aug 30;20(4):665-680. [doi: 10.1007/s10111-018-0515-x]
33. Martin S, Kelly S, Khan A, Cullum S, Dening T, Rait G, et al. Attitudes and preferences towards screening for dementia: a systematic review of the literature. BMC Geriatr 2015 Jun 16;15(1):66 [FREE Full text] [doi: 10.1186/s12877-015-0064-6] [Medline: 26076729]
34. Farina N, Hicks B, Baxter K, Birks Y, Brayne C, Dangoor M, et al. DETERMinants of quality of life, care and costs, and consequences of INequalities in people with Dementia and their carers (DETERMIND): a protocol paper. Int J Geriatr Psychiatry 2020 Mar 25;35(3):290-301. [doi: 10.1002/gps.5246] [Medline: 31876069]

35. Robinson L, Clare L, Evans K. Making sense of dementia and adjusting to loss: psychological reactions to a diagnosis of dementia in couples. Aging Ment Health 2005 Jul;9(4):337-347. [doi: 10.1080/13607860500114555] [Medline: 16019290]
36. Dai R, Spector A, Wong G. e-Mental health care for people living with dementia: a lesson on digital equality from COVID-19. Alzheimers Dement (Amst) 2020;12(1):e12100 [FREE Full text] [doi: 10.1002/dad2.12100] [Medline: 33072845]
37. Speechly CM, Bridges-Webb C, Passmore E. The pathway to dementia diagnosis. Med J Aust 2008 Nov 03;189(9):487-489. [doi: 10.5694/j.1326-5377.2008.tb02140.x] [Medline: 18976188]
38. Lee DR, McKeith I, Mosimann U, Ghosh-Nodial A, Grayson L, Wilson B, et al. The dementia cognitive fluctuation scale, a new psychometric test for clinicians to identify cognitive fluctuations in people with dementia. Am J Geriatr Psychiatry 2014 Sep;22(9):926-935. [doi: 10.1016/j.jagp.2013.01.072] [Medline: 24332982]
39. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969 Sep 01;9(3 Part 1):179-186. [doi: 10.1093/geront/9.3_part_1.179]
40. Dooley S, Walshe M. Assessing cognitive communication skills in dementia: a scoping review. Int J Lang Commun Disord 2019 Sep 28;54(5):729-741. [doi: 10.1111/1460-6984.12485] [Medline: 31250524]
41. Hendryx-Bedalov PM. Alzheimer's dementia. Coping with communication decline. J Gerontol Nurs 2000 Aug;26(8):20-4; quiz 52. [doi: 10.3928/0098-9134-20000801-06] [Medline: 11276608]
42. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]
43. Rea L, Parker A. Designing and Conducting Survey Research: A Comprehensive Guide, 4th Edition. Hoboken, New Jersey, United States: Wiley; 2014:219.
44. Ball KK, Ross LA, Roth DL, Edwards JD. Speed of processing training in the ACTIVE study: how much is needed and who benefits? J Aging Health 2013 Dec;25(8 Suppl):65S-84S [FREE Full text] [doi: 10.1177/0898264312470167] [Medline: 24385640]
45. Barberger-Gateau P, Commenges D, Gagnon M, Letenneur L, Sauvel C, Dartigues JF. Instrumental activities of daily living as a screening tool for cognitive impairment and dementia in elderly community dwellers. J Am Geriatr Soc 1992 Nov;40(11):1129-1134. [doi: 10.1111/j.1532-5415.1992.tb01802.x] [Medline: 1401698]
46. Braun M, Scholz U, Bailey B, Perren S, Hornung R, Martin M. Dementia caregiving in spousal relationships: a dyadic perspective. Aging Ment Health 2009 May;13(3):426-436. [doi: 10.1080/13607860902879441] [Medline: 19484607]
47. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 4: trustworthiness and publishing. Eur J Gen Pract 2018 Dec 05;24(1):120-124 [FREE Full text] [doi: 10.1080/13814788.2017.1375092] [Medline: 29202616]
48. 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 Nov 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]
49. Holthe T, Halvorsrud L, Karterud D, Hoel K, Lund A. Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: a systematic literature review. Clin Interv Aging 2018;13:863-886 [FREE Full text] [doi: 10.2147/CIA.S154717] [Medline: 29765211]

Abbreviations

COM-B: Capability, Opportunity, and Motivation–Behavior
JDR: Join Dementia Research
NIHR: National Institute for Health and Care Research
POA: proportion of observed agreement
PPI: patient and public involvement

Edited by N Zary; submitted 29.07.21; peer-reviewed by A Garcia, L Ferreira, N Choudhury, LJ Cheng; comments to author 30.10.21; revised version received 22.12.21; accepted 17.02.22; published 20.06.22.

Please cite as:
Harrington K, Craven MP, Wilson ML, Landowska A
Perceptions of Cognitive Training Games and Assessment Technologies for Dementia: Acceptability Study With Patient and Public Involvement Workshops
JMIR Serious Games 2022;10(2):e32489
URL: https://games.jmir.org/2022/2/e32489
doi:10.2196/32489
PMID:35723912

Original Paper

Feedback on Trunk Movements From an Electronic Game to Improve Postural Balance in People With Nonspecific Low Back Pain: Pilot Randomized Controlled Trial

Anita Meinke[1], MSc; Rick Peters[2], MAS; Ruud H Knols[2,3,4], PhD; Jaap Swanenburg[3,5,6*], PhD; Walter Karlen[1,7*], PhD

[1]Mobile Health Systems Lab, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

[2]Department of Physiotherapy Occupational Therapy, University Hospital Zurich, Zurich, Switzerland

[3]Directorate of Research and Education, Physiotherapy Occupational Therapy Research Center, University Hospital Zurich, Zurich, Switzerland

[4]Institute of Human Movement Sciences and Sport, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

[5]Faculty of Medicine, University of Zurich, Zurich, Switzerland

[6]Integrative Spinal Research, Department of Chiropractic Medicine, Balgrist University Hospital, Zurich, Switzerland

[7]Institute of Biomedical Engineering, University of Ulm, Ulm, Germany

[*]these authors contributed equally

Corresponding Author:
Anita Meinke, MSc
Mobile Health Systems Lab
Department of Health Sciences and Technology
ETH Zurich
Balgrist Campus, BAA
Lengghalde 5
Zurich, 8008
Switzerland
Phone: 41 446337754
Email: meinke_a@outlook.de

Abstract

Background: Postural balance is compromised in people with low back pain, possibly by changes in motor control of the trunk. Augmenting exercising interventions with sensor-based feedback on trunk posture and movements might improve postural balance in people with low back pain.

Objective: We hypothesized that exercising with feedback on trunk movements reduces sway in anterior-posterior direction during quiet standing in people with low back pain. Secondary outcomes were lumbar spine and hip movement assessed during box lift and waiter bow tasks, as well as participant-reported outcomes. Adherence to the exercising intervention was also examined.

Methods: A randomized controlled trial was conducted with the intervention group receiving unsupervised home exercises with visual feedback using the Valedo Home, an exergame based on 2 inertial measurement units. The control group received no intervention. Outcomes were recorded by blinded staff during 4 visits (T1-T4) at University Hospital Zurich. The intervention group performed 9 sessions of 20 minutes in the 3 weeks between T2 and T3 and were instructed to exercise at their own convenience between T3 and T4. Postural balance was assessed on a force platform. Lumbar spine and hip angles were obtained from 3 inertial measurement units. The assessments included pain intensity, disability, quality of life, and fear of movement questionnaires.

Results: A total of 32 participants with nonspecific low back pain completed the first assessment T1, and 27 (84%) participants were randomized at T2 (n=14, 52% control and n=13, 48% intervention). Intention-to-treat analysis revealed no significant difference in change in anterior-posterior sway direction during the intervention period with a specified schedule (T2-T3) between the groups (W=99; P=.36; r=0.07). None of the outcomes showed significant change in accordance with our hypotheses. The intervention group completed a median of 61% (55/90; range 2%-99%) of the exercises in the predefined training program. Adherence was higher in the first intervention period with a specified schedule.

Conclusions: The intervention had no significant effect on postural balance or other outcomes, but the wide range of adherence and a limited sample size challenged the robustness of these conclusions. Future work should increase focus on improving adherence to digital interventions.

Trial Registration: ClinicalTrials.gov NCT04364243; https://clinicaltrials.gov/ct2/show/NCT04364243

International Registered Report Identifier (IRRID): RR2-10.2196/26982

(JMIR Serious Games 2022;10(2):e31685) doi:10.2196/31685

KEYWORDS

low back pain; postural balance; exergame; postural feedback; motor control; kinesiophobia; inertial measurement unit; randomized controlled trial

Introduction

Background

Low back pain (LBP) contributed to most years lived with disability to the global burden of disease in 1990 as in 2017 [1]. The impact of LBP ranges from causing minor inconvenience to substantial restrictions in daily activities and, in extreme cases, disability and early retirement. Although there may be improvements when considering population age, the overall years lived with disability from LBP is rising and needs to be addressed [1]. Standard treatment recommendations for LBP often incorporate exercising and advice regarding physical activity [2], and it has been demonstrated that exercises for chronic LBP improve outcomes such as pain or disability to a certain degree [3-5]. Differences in effects between exercises with a distinct training focus have been appraised as negligible [4,5]. The limited effectiveness motivates the exploration of new ways for enhancing these treatments, as it has already been outlined by other authors [6]. Considering that changes of motor control of the lumbar region are discussed as a plausible cause for recurrence of LBP [7] and given that feedback plays a central role in motor learning [8], digital tools that make movement patterns more visible could be one such way to enhance exercise treatments. These interventions could be used together with other treatments as implemented in previous studies [6,9-11] or independently as needed, as a form of self-management. Many people with LBP do not request treatment, especially those with mild disability [12]. Therefore, supportive technology that can provide some degree of guidance at home, while maintaining independence, and could be especially interesting for this group of people.

Motor control can be described "...as the way in which the nervous system controls posture and movement to perform a specific motor task, and includes consideration of all the associated motor, sensory, and integrative processes" [7]. Physical characteristics and movement behaviors assessed to derive insights into deviations in people with LBP concerning these processes have revealed many new insights but still demand further clarification [7,13]. Examples of movement differences such as the limitations in range of motion (ROM) of the lumbar spine were found in all movement planes [14] and limited ROM of the trunk in the frontal plane but not in the sagittal plane seem to precede the occurrence of LBP [15]. Differences in the trunk region are thought to relate to differences in postural balance [7], which have been found by many studies [13,16,17]. Consequently, practicing movement tasks that focus on movement of the lumbar spine and hip could have the potential to enhance postural balance. Nevertheless, as highlighted earlier, the associations of LBP and movement behavior have not been fully clarified [7].

Altered movement behaviors seem to extend to tracing tasks, which work with feedback on trunk movements and have been used as an indicator for motor control in laboratory settings [18,19]. These studies used tasks that required participants tracing a circular pattern [18] or performing flexion movements [19] with their trunk, while receiving concurrent feedback. Results regarding the accuracy were conflicting, as one study found a difference between people with and without LBP in the accuracy of the tracing [18], whereas the other did not [19]. However, the latter study confirmed differences with respect to timing relative to the feedback between the groups [19]. Similar tasks may serve not only as a proxy measure of trunk motor control but also as training opportunity. It was recently found that practice to keep the lumbar spine constantly neutral during a box lift task was more successful when participants obtained digital feedback than when the participants used a mirror [20]. As described earlier, people with LBP seem to show movement patterns that deviate from those in people without LBP in different manners, such as displaying a high degree of rigidity or little control at all on their movement [7]. In tracing tasks, where movement becomes visible in relation to a target, people with LBP may develop well-coordinated trunk movements that are neither rigid nor loose. Furthermore, proprioception seems to be affected in people with LBP [21], which might relate to difficulties with fine-tuned trunk movements [18]. Early results indicate that there may be improvements in proprioception through interventions using feedback on trunk movement in people with LBP [22]. In short, we assume that such training may reduce postural balance impairments by restoring more adequate movement behaviors of the trunk.

The effects of exercising interventions directly on postural balance have been studied previously. Meta-analyses on intervention studies with older people suggest that balance training [23] and Pilates [24] but not programs focusing on strength or mixing different kinds of exercises [23] can enhance postural balance. Although these results originate from older participants, studies in people with LBP imply similar development of pain [25] and treatment results [26], largely independent of age. Furthermore, altered trunk characteristics have also been reported in younger people with LBP [27]. Multiple interventional studies with people with LBP found an impact of exercising interventions on at least one of the investigated criteria describing balance [28-30], whereas in another study no differences in postural balance were detected [31]. However, different tasks with varying requirements were used; for example, standing on moving ground [28], standing on a single leg [29], and assessments in squat positions [30].

Reviews so far revealed mixed success of digital tools for exercising [32] and encouraging early results for virtual reality applications [33] in people with LBP. The present literature has been described as heterogeneous with respect to the interventions [32] and methods [33]. Technical progress is considered as a factor to possibly impact intervention effectiveness, although so far, no clear difference for later studies was observed [32]. Thus, consistent research effort is necessary. Digital tools can be designed to influence the way movements are performed during exercises, for instance, to vary the speed [34]. Investigated tools include exergames readily available in the market; for example, exercising with the well-known Wii balance board (Nintendo of America, Inc) [35,36] or the Valedo Motion (Hocoma AG) [6,11]. Sensor technology has been used to intervene on movement characteristics of the trunk directly; for example, in studies [37,38] where warning participants from performing extreme back movements during everyday work was investigated. Other interventions have specifically encouraged movement of the lumbar spine in an exercising context [9,11,22,39] or are otherwise dedicated to providing feedback on lumbar spine movement [9-11,22]. For different kinds of tools used, only a small amount of research has been conducted [32]. Therefore, such digitally supported training modalities should further be investigated. Different systems and technological setups have been explored, for instance, cameras [40], wearable sensors [6,41], and sensors readily available in mobile phones [42], sometimes in combination with virtual reality headsets [41,42]. However, only few studies provide first insights in the effects of these interventions on movement quality in people with LBP [9-11]. In all, 2 studies suggested such interventions might have positively affected trunk ROM, but in one study, it remained ambiguous whether there was a significant difference to the standard care control group [9] and in the other study only a single group was investigated [22]. In a third study no effect on ROM was found [10]. Motor control impairment was not different in a study, where patients in the intervention group received access to additional exercises with sensor-based feedback other than the control group [11]. To our knowledge, the effect of such exercises on postural balance in people with LBP has not yet been investigated.

Objectives

The primary aim of this study was to examine whether exercising with feedback on trunk movements can enhance postural balance, indicated by the change in anterior-posterior (AP) postural sway between the assessments before and after the intervention. Additional parameters to quantify postural balance were explored. As secondary outcomes, movement of the lumbar spine and hip during 2 different movement tasks and participant-reported outcomes were included. A further aim was to analyze adherence to the intervention.

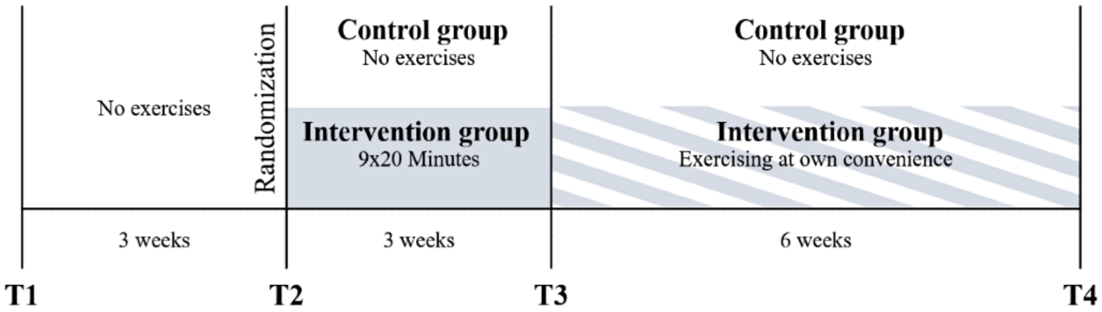
Methods

The completed CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist [43] is provided in Multimedia Appendix 1 [2-12,20,22,32,39-42,44-50]. The intervention is described according to the TIDieR (Template for Intervention Description and Replication) checklist [51].

Study Design

This manuscript was based on a study protocol [44] that included a 2-arm randomized controlled trial. Figure 1 shows the assessment and intervention schedule. The study took place at University Hospital Zurich between May 2019 and October 2020. Except for an extension of the study period of 3 Months to compensate for a pause due to the COVID-19 pandemic, the study was completed as planned, and interim analyses of intervention effects were not conducted. Outcomes were assessed twice at T1 and T2, before an intervention was given. Further assessments were taken after another 3-week period with a fixed exercising schedule for the intervention group (T3) and a subsequent 6-week period without specified exercising schedule (T4). Participants were randomized during the assessments at T2, and those assigned to the intervention group received an introduction to the exercising program right after the assessment. After T3, participants in the intervention group retained the Valedo Home exercising system (Hocoma AG), without being required to follow a specific schedule or to complete any exercises at all. This period was introduced to observe further adherence to the exercising program, without commitment to a schedule provided by a therapist or to complete a schedule for research purposes. Participants who were randomized to the control group did not receive a sham intervention.

Figure 1. Overview of the study schedule showing assessment visits T1 to T4 and the interventions.



Block randomization (blocks of 2 and 4), stratification by body height, and 1:1 allocation were implemented through the randomization tool in REDCap (Research Electronic Data Capture; Vanderbilt University) [45] hosted at Eidgenössische Technische Hochschule Zurich. AM generated random sequences with the dedicated R package blockrand (version 1.3; [49]) and randomized the participants using REDCap. The staff conducting assessments of the outcomes was blinded, and randomization occurred as late as possible (at T2) to reduce the risk of accidental unblinding.

The published study protocol [44] contained a further research question involving an additional patient group. We intended to compare the effect of the intervention between a group of patients and other participants who did not receive other treatments than the exercise intervention with postural feedback. In this manuscript, we report only the research questions that could be investigated based on the collected data, as insufficient patients were enrolled at the study site.

Participant Recruitment

Participants were recruited through different web-based bulletin boards, websites, distribution of flyers, and personal communication. Recruitment was completed 3 months before the planned end date, to allow all participants to finish in time. Eligibility was ascertained in an interview-like setting that allowed the participants to describe their situation. Participants were considered eligible if they provided informed consent, they were aged at least 18 years, reported to the investigator that they had nonspecific LBP, and did not receive therapy or medical treatment for LBP within the past 6 months. There was no questionnaire-based assessment or formalized cutoff scores for pain intensity during the eligibility check. If it was not possible to clarify the eligibility of a participant by asking standard questions alone, one of the trained physiotherapists of the team decided on the criterion in question. The criterion of no recent treatment was relaxed from 12 months to 6 months during the study to improve recruitment rates. Participants reporting specific LBP or radicular syndrome were excluded from participation. Participants were also excluded if they indicated to the investigator that they would not be able to complete the movements required by the exercise intervention owing to high pain. Other reasons for exclusion were pregnancy, taking medication that impairs postural balance, severely impaired vision, allergic reactions to adhesive strips, and insufficient proficiency in German or English.

Ethics Approval

Participants were not compensated and provided informed consent in writing before any study procedure was started. The trial was approved by the Cantonal Ethics Committee Zurich (BASEC: 2018-02132) and registered in ClinicalTrials.gov (NCT04364243).

Outcomes and Procedures

Postural Balance

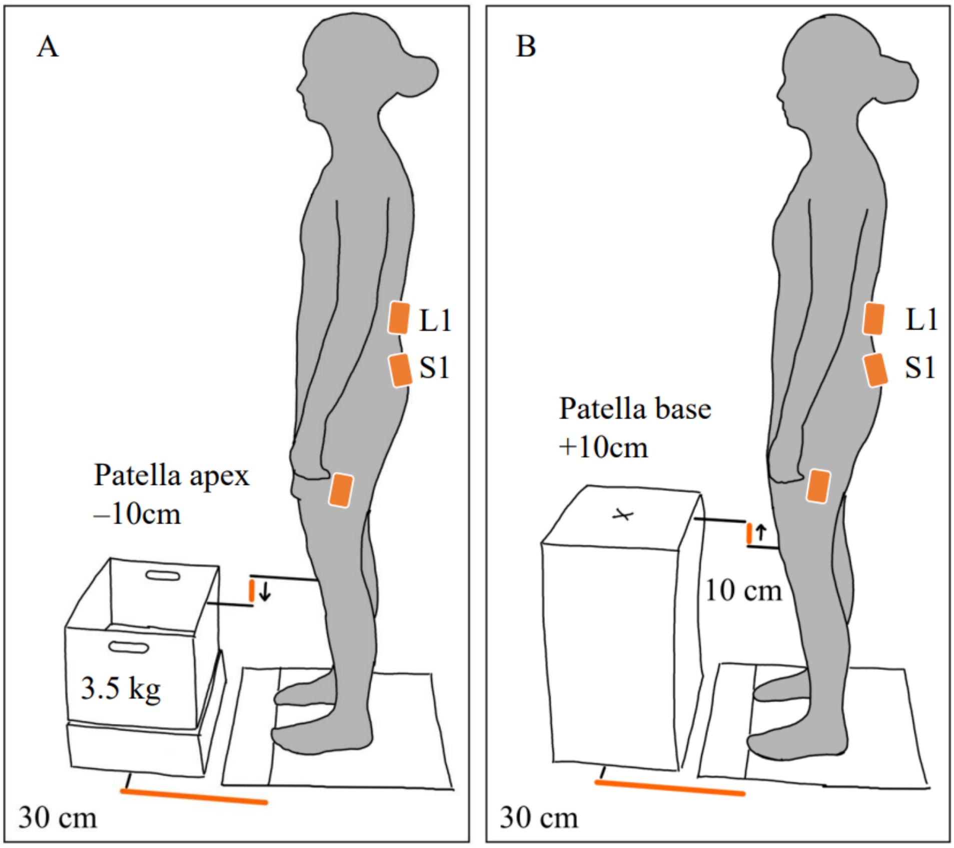
Records of center of pressure (COP) during quiet standing on a stable force platform (AMTI, Accusway Plus) were used to quantify postural balance. Specifications of the number of repetitions, duration, instructions, sampling rate, and filter cutoff frequency were based on relevant literature [52] and are described in detail in this section. During the assessment, the participants stood as quietly as they could, with the arms relaxed at the side and eyes closed, while wearing opaque goggles. Each participant selected an individually comfortable, *usual* foot position. To keep the stance consistent for each participant during all balance assessments, the foot position was recorded on a plastic foil. Participants wore socks but no shoes on the platform. A total of 4 postural balance trials of a duration of 120 seconds were recorded with a sampling rate of 100 Hz at each assessment visit. The data were filtered using a fourth order low-pass Butterworth filter with a cutoff frequency of 10 Hz. Data from the first and last 5 seconds were removed from the records to permit a stabilization phase at the beginning and to assure that any effects of a lateral leaning movement used for time synchronization with additional sensor data were removed with a safety margin. Thus, parameter estimates for each repetition were based on segments of 110 seconds. Mechanisms regulating balance in AP and medio-lateral direction differ [53]; therefore, directional measures were used to quantify balance in addition to measures irrespective of the direction on the 2D plane (global). The trajectory of the COP was described by the mean absolute displacement from the mean COP (AP, medio-lateral, and global), and by corresponding velocities, graphically represented in Multimedia Appendix 2 [54]. Change in displacement in AP direction (T3-T2) was a priori defined as the primary outcome. The data were reported on a mm and mm/s scale, and reduction in displacement and velocity were the favorable outcomes.

Movement Tasks

Further assessments during movement tasks were performed to see whether the participants were able to follow the instruction to limit movement of the lumbar spine and perform movements on the sagittal plane by bending the hip joint instead. The protocol and setup of these assessments were adopted from the study by Matheve et al [46]. The assessments taken during box lift and waiter bow tasks were shown to reliably determine the change of the lumbar spine and hip posture during these tasks [46]. Similar versions of these tasks have been used elsewhere [20]. Figure 2 shows the setting and adaptation of the tasks to the individual participants. Lumbar spine and hip angles were used to describe the performance during these tasks. The Valedo Pro (Hocoma AG), consisting of 3 inertial measurement units (IMU) and dedicated software, was used for the assessments. The IMUs were placed with medical adhesive strips at the height of the spinal process of the S1 and L1 vertebrae, and 1 IMU was placed on the left leg, 20 cm from the lateral femoral condyle. Sensor positions were identified by palpation. Different from the aforementioned study [46], we did not alter the participants natural spinal posture before the tasks were performed. We assumed the tasks would otherwise be selectively more difficult to perform for participants who received more intense corrections to their posture. In addition, we allowed only 1 practice trial before the 5 repetitions of each task to keep learning effects minimal. By tracing the position of the feet to a foil, the position was standardized across assessments.

The box lift task required the participants to lift a box and hold it during upright standing and put the box down again and return to the standing position. For the waiter bow task, the participants were asked to touch a marked spot positioned in front of them with the fingers by bending from the hip joints and return to upright stance. The central instruction was to not change the alignment of the lumbar spine while performing the tasks. During the waiter bow task, participants were instructed to keep their knees at the original angle. The participants stood with their feet parallel, at a self-selected width, in a predefined distance to the task materials. Correct task execution and possible mistakes were shown to the participants by the outcome assessor. The order of tasks was randomized for each assessment visit. The data were collected at a sampling frequency of 50 Hz and change in lumbar spine angle was calculated by subtracting the rotation of the S1 sensor on the sagittal plane from the rotation of the L1 sensor on the sagittal plane. The obtained data were filtered using a moving average of 0.2 seconds, and the maximum absolute departure of the position at task beginning was used as the end point. Hip angles were obtained analogously, using data from the IMU at S1 and IMU at the thigh.

Figure 2. Setup of the movement tasks, inertial measurement unit positions (orange markers), and task material adaptation for A: box lift task and B: waiter bow task. The specifications have been adopted from Matheve et al [46].



Participant-Reported Outcomes

Before the movement assessment at each visit, the participants completed a questionnaire in English or German, on a laptop. Considering the recommendations regarding relevant outcome assessments for studies on LBP [55], we included questionnaires covering pain intensity, disability associated with LBP, and quality of life (QOL). A 11-point numeric rating scale (NRS) asking participants to rate their pain intensity during the previous week, with the anchors *no pain* and *worst imaginable pain*, was applied additionally at the first assessment [56].

The Roland Morris disability questionnaire (RMDQ) was used to measure disability [57,58]. Respondents selected those of the 24 statements, which they experienced on the date of assessment, resulting in scores from 0 to 24 [57,58]. The RMDQ is an established questionnaire with adequate psychometric performance [59].

The World Health Organization Quality of Life Questionnaire-short version (WHOQOL-Bref) includes 26 items, which cover different aspects of QOL: physical health, psychological QOL, social relationships, and environmental factors [60]. The score of the physical health subscale is calculated by averaging the responses of 7 items (5 response options per item multiplied by 4) [60]. The selection of questions for the WHOQOL-Bref was based on data from international samples [60], and it was found to be reliable and valid [60,61].

The Tampa Scale for Kinesiophobia, 11- item version, was used to measure fear of movement, and the sum scores (4 response options: 11 to 44) were analyzed [62,63]. The English and German versions were found to generate reliable and valid data [62,63].

Baseline Characteristics and Adherence

The questionnaire at T1 contained questions regarding the participants age at the first occurrence of LBP, days with LBP during the previous month, and the average LBP intensity (*When you have back pain, how would you rate your average low back pain intensity in general?*), using labels of *no pain* and *worst imaginable pain* to describe the minimum and maximum values of 0 and 10, respectively. In addition, demographic data were collected. Weight and height were assessed at the study site. The exercises that were performed at home and the matching time stamps were extracted from the Valedo Home app.

Intervention

The intervention is described in Textbox 1. Multimedia Appendix 3 contains a video on the exercises used in this study (published with permission from Hocoma AG).

Textbox 1. The TIDieR (Template for Intervention Description and Replication) checklist items [51] for exercising, with postural feedback.

Item and intervention description

- Brief name: exercising with postural feedback on trunk movements using the Valedo Home system (Hocoma AG).
- Rationale: postural balance deficits in people with low back pain may stem from disturbed coordination of the trunk. We assume that practicing trunk movements with a feedback system helps participants to learn to control their trunk precisely. This improved control of the trunk could in turn affect how well balance can be controlled in people with low back pain.
- Materials: the Valedo Home system and belts or medical adhesive strips that were used for attaching the sensors to the chest and lower back. A tablet (Huawei Media Pad T5) with the Valedo app, a paper document summarizing the instructions, and the user manual [47] were provided.
- Procedures: participants randomized to the intervention group were instructed how to use the Valedo system and performed 1 exercise under supervision at T2. During this training session, the participants learned how to place the sensors correctly and to use the tablet and the Valedo app. At each of 9 home exercising sessions, the participants performed 10 exercises. Multimedia Appendix 3 contains a video showing all the exercises. The participants practiced moving their trunk and pelvis precisely to guide an avatar along a specified path with their movements through a virtual world. The exercises consist of movements of the upper body or the pelvis. Trunk movements are performed on the sagittal, frontal, and transversal plane, and hip movements are performed on the sagittal and frontal plane. Participants see on the display how well they match the specified movement trajectory while playing, and further auditory feedback is provided. At the end of the game, a ranking of the current and previous performance in the game is provided. After the assessment at T3, the participants in the intervention group were informed that they could keep using the system until T4 and that there was no specific schedule to complete, and they could use the system at their own convenience.
- Provider: the exercises were delivered by the Valedo Home system. AM trained the participants and acted as the contact person during the study. The participants were encouraged to contact AM, if any questions or technical difficulties should occur.
- Mode of delivery: each participant was instructed individually. Exercises were guided by the Valedo Home system.
- Location: instructions were provided at University Hospital Zurich, and the regular exercises were performed by the participants at home.
- Frequency and duration: the participants completed 10 exercises repeatedly, with an effective duration of 20 minutes, in 9 sessions until T3. Participants were told to space out the exercising sessions approximately equally between the appointments, but the exact dates were not defined. After T3, the participants could choose the exercises and duration by themselves.
- Tailoring: the exercises are adapted to the range of motion of the participant, which is measured as part of the user profile setup. Participants could repeat this assessment at any time. Progress and difficulty were determined by the Valedo Home app.
- Modifications: to improve the attractivity of the study and recruitment, starting September 2019, the participants in the control group could borrow the Valedo Home and tablet for 3 weeks after completion of T4.
- Adherence measures: the exercises performed by the participants were automatically recorded on the tablet.
- Actual adherence: reported in the *Results* section.

Data Preparation and Statistical Analysis

Data preparation and analysis were conducted in MATLAB R2018a (The MathWorks, Inc) and R (version 4.0.4; R Foundation for Statistical Computing). The simultaneously recorded data of the force platform and the IMUs were time-synchronized based on aligning a sideways leaning movement of the participants, shifting their weight to the left and the right, which was performed before and after each repetition. For this time synchronization, the movement had to be clearly distinct from the tasks and identifiable in both sources of data. This parallel recording was important for comparing force plate and IMU data. The beginning and the end of each balance and movement task were defined based on marker time stamps set in the IMU data during the assessment. The markers were inspected visually and corrected by hand before further analysis, as placement during the assessment was sometimes not optimal and occurred too early during the time-synchronization movement or too late during the task.

To assess the equivalence among the treatment groups at study entry (T1), participant characteristics were compared. Welch *t* test (2-tailed) or alternatively Wilcoxon rank-sum tests were used, if the data appeared to be not normally distributed based on Normal QQ plots or Shapiro-Wilk tests within groups. Dependent group *t* tests (2-tailed) were used to test whether change had occurred between T1 and T2, or if the assumptions were not met, Yuen tests, as provided by the R package WRS2 [64], were used. The hypotheses regarding the intervention effects were tested by comparing the change of the respective outcome (Δ outcome: T3-T2) between the intervention and the control group, predicting the more favorable outcome for the intervention group. These comparisons were performed each as intention-to-treat (ITT) and per-protocol (PP) analyses. In the ITT analyses, all participants who had been randomized at

T2 were included. Missing values at T2 and T3 were replaced with the mean of the previous 2 assessments (T1 and T2) of the participant. For the PP analyses, participants who either had incomplete data or had been randomized to the intervention group but exercised <1 hour, between T2 and T3, were excluded. Comparisons were performed using independent group *t* tests (1-tailed), when the data were normally distributed according to Shapiro-Wilk tests, and a Levene test did not show heterogeneity of variances. Otherwise, Wilcoxon rank-sum tests were used. The a priori power analysis is reported within the study protocol [44]. We calculated post hoc power for the primary comparison using G*Power (Heinrich Heine Universität Düsseldorf) [65]. The effect size *r* observed for the ITT comparison of the primary outcome was converted to *d* using a web-based tool [66]. Power was calculated for a directed Wilcoxon rank-sum test with an α level set to .05 and the distribution menu set to *normal*.

Additional exploratory analyses to compare the absolute scores across all assessment visits including the second intervention period were conducted using mixed 2-way ANOVA. Only participants who completed the study (n=20) were included in these analyses. Missing data were replaced by mean scores of the previous assessments of that participant. Generalized η^2 was used as effect size [48], and calculations were made using the r package rstatix (version 0.7.0; [50]). Shapiro-Wilk tests and Levene tests were used to test the assumptions of normality and homogeneity of variances. If the data did not fulfill the assumptions of normality and homogeneity of variances, different data transformations were explored. In cases where no suitable transformation was found, Friedman ANOVA was conducted across the assessment visits for each group separately, and group differences were compared at each assessment visit using the Bonferroni-corrected Wilcoxon rank-sum tests. Data on adherence were analyzed using descriptive statistics and graphs.

Results

Data Cleaning and Preparation

On the basis of visual inspection, orientation data from the IMU sensors were corrected in 2 cases where axes were flipped (15 trials of 2 participants). The data from a participant at T1 and another participant in the control group at T2 were discarded, because misplacement of the sensors was suspected. For a participant in the intervention group, no data for the T3 assessment were available, as the sensors had not been sufficiently charged. In the ITT analysis, all randomized participants were analyzed and missing values were replaced as described in the *Methods* section. For the ITT analysis of the balance and questionnaire data, 6 replacements (5 control and 1 intervention) were made for the T3 assessments. For the movement tasks, 1 replacement in the control group was made for the T2 assessments, and 7 replacements (5 control and 2 intervention) were made for the T3 assessments. For the PP analysis, participants for whom replacements had to be made were removed from the analysis. In addition, data from the participants in the intervention group (3 balance and questionnaires; 2 movement tasks), who had exercised <1 hour within the 3-week period and were excluded from the PP analysis, if this was not already the case when the participant was also a dropout, or the data had already been removed owing to insufficiently charged sensors.

During data analysis, it was discovered that some of the items of the WHOQOL-Bref at T3 in the German version had been collected with response options ranging from 1 to 4 instead of 1 to 5. Data collected with the affected items (items 3-9) were discarded for all assessment visits, and the scores of the scales were calculated without those items.

As not all data fulfilled the requirements for the 2-way mixed ANOVA for the analysis across all 4 assessment visits, the data were transformed where necessary. Transformations applied are reported in Table 1.

Table 1. Transformations applied to satisfy requirements for the analysis including all assessment visits in the two-way mixed ANOVA.

Outcome	Transformation
Mean anterior-posterior displacement	max(1 / (x + 1)) – (1 / (x + 1))
Mean medio-lateral displacement	log(x + 1)
Mean global displacement	log(x + 1)
Mean anterior-posterior velocity	None necessary
Mean medio-lateral velocity	max(1 / (x + 1)) – (1 / (x + 1))
Mean global velocity	None necessary
Box lift lumbar spine	log((x / 5) + 1)
Box lift hip	log((x / 5) + 1)
Waiter bow lumbar spine	log((x / 5) + 1)
Waiter bow hip	log((x / 5) + 1)
Pain intensity numeric rating scale	log(x + 1)
Roland Morris disability questionnaire	No suitable found
Quality of life physical subscale	log(x + 1)
Tampa Scale for Kinesiophobia, 11-item version	None necessary

Duration Between Assessment Visits

The effective duration between T1 and T2 had a median of 21 days (IQR 5; minimum 17, maximum 97); between T2 and T3, 23 days (IQR 3; minimum 19, maximum 36); and between T3 and T4, 44 days (IQR 7.75; minimum 38, maximum 112). For a participant, the time span between T1 and T2 was extended to 97 days, and for 2 participants, the period between T3 and T4 was extended to 112 and 99 days, respectively, because of an interruption in the study owing to the COVID-19 pandemic. The period between T2 and T3 was not affected by extensions due to the COVID-19 pandemic.

Participants and Baseline Characteristics

As presented in Figure 3, a total of 93 participants made an initial contact and requested information regarding the study. Of those 93 participants, 38 (41%) provided written informed consent. At T1, a total of 32 participants, without recent treatment for LBP, were eligible for the study. In all, of 32 participants, 5 (16%) dropped out before randomization at T2 (n=27, 84%). Tables 2 and 3 show the participant characteristics at baseline. Between the participants randomized to the intervention and the control group, there were no significant differences in any outcome measure at T1 or T2, but analyses of change between T1 and T2 revealed that there was a significant reduction in pain intensity across participants who had been randomized (T1: median 3.00; mean 3.26, SD 1.56 and T2: median 2.00; mean 2.59, SD 1.34). Descriptive statistics on all outcomes at all assessment visits are reported in Table S1 in Multimedia Appendix 4 and comparisons of outcomes at T1 and T2 are shown in Table S2 in Multimedia Appendix 4.

Figure 3. Participant flow through the study. Numbers of analyzed participants are reported in the text. LBP: low back pain.

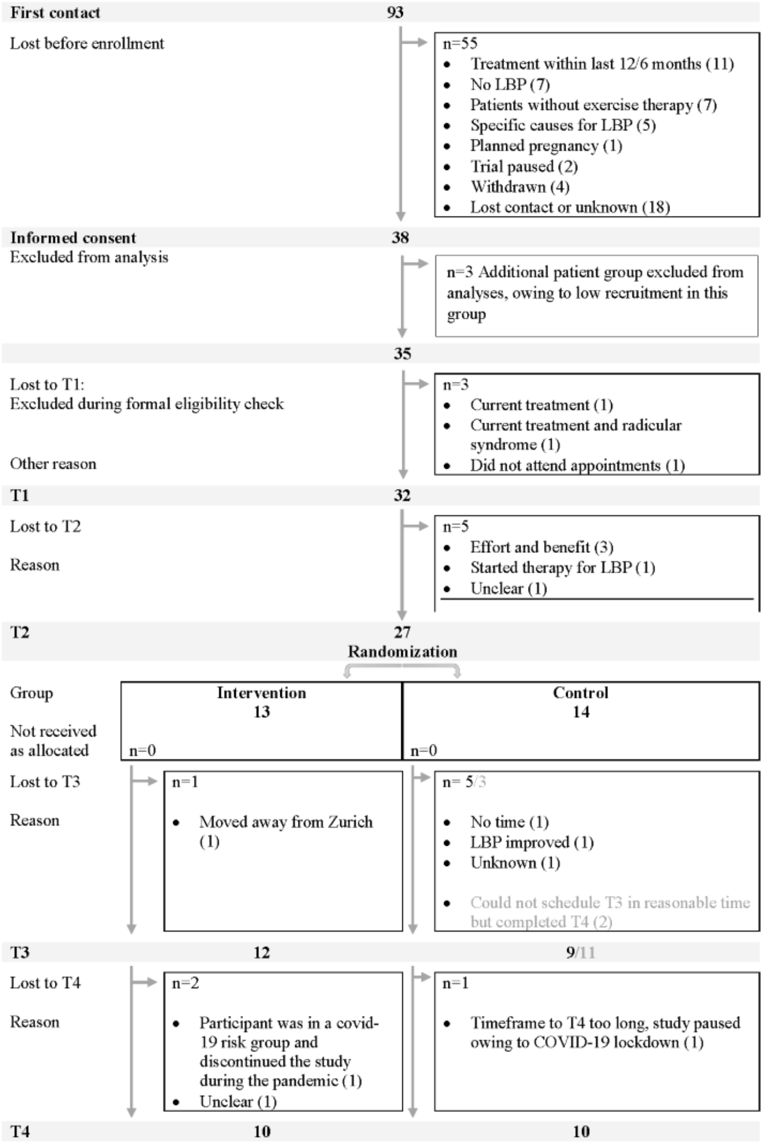


Table 2. Participant characteristics at T1.

Characteristics	Control (n=14)		Intervention (n=13)		Comparison
	Median (range)	Mean (SD)	Median (range)	Mean (SD)	*P* value
Age (years)	37.50 (38.00)	40.14 (12.38)	34.00 (40.00)	40.85 (15.15)	.92[a]
Height (cm)	174.25 (36.00)	173.27 (8.61)	170.50 (18.50)	170.73 (6.55)	.39
Weight (kg)	73.55 (42.10)	76.01 (11.97)	74.10 (36.40)	72.56 (9.57)	.41
Age at first time LBP[b] (years)	24.50 (33.00)	26.50 (10.51)	20.00 (29.00)	24.00 (9.52)	.52
LBP previous month (days)	10.00 (18.00)	9.43 (6.16)	11.00 (28.00)	5.46 (10.08)	.16[a]
Average pain intensity (0-10)	4.00 (5.00)	4.07 (1.33)	4.00 (4.00)	3.85 (1.07)	.63

[a]Wilcoxon rank-sum test.

[b]LBP: low back pain.

Table 3. Gender and language data of participants at T1.

Characteristic	Values
Gender (female/male)	
Control	9/5
Intervention	8/5
Language (German/English)	
Control	11/3
Intervention	11/2

Change in Outcomes During the Intervention Period With Predefined Schedule

Overview

The change in the outcome variables between (T3-T2) was compared between both groups, for all outcome variables. All comparisons were performed as ITT and PP analyses. ITT analyses were performed with 14 participants in the control and 13 participants in the intervention groups, as randomized. PP analyses were conducted with 9 participants in each group, except for the movement tasks, where data from only 8 participants were available in the control group. Descriptive statistics and T2 and T3 scores for the ITT and PP analysis are reported in Table S3 in Multimedia Appendix 4.

Postural Balance

The primary outcome, change between T2 and T3 in mean AP displacement, did not differ among groups in the ITT analysis (control: median 0.01, range 3.91; mean 0.32, SD 0.95 and intervention: median 0.18, range 2.76; mean 0.31, SD 0.77; comparison: W=99; P=.36; r=0.07) and neither in the PP analysis (control: median 0.03, range 3.91; mean 0.45, SD 1.17 and intervention: median 0.05, range 2.51; mean 0.17, SD 0.69; comparison: t_{16}=0.64; P=.73; r=0.16). The post hoc power for detecting the small effect observed in the ITT analysis, r=0.07 (equivalent to d=0.14), of the primary outcome was 0.1. No group differences in the ITT or PP analyses were found for the other postural balance parameters explored (Table 4). Multimedia Appendix 5 shows a graph of the postural balance data as it was used for the ITT analysis.

Table 4. Directed group comparisons of change in the balance outcomes between T2 and T3.

Outcome and analysis	Control		Intervention		Comparison			
	Median (range)	Mean (SD)	Median (range)	Mean (SD)	*t* test (*df*)	*W*	*P* value	*r*
Δ mean medio-lateral displacement								
ITT[a]	–0.22 (3.64)	–0.36 (0.84)	–0.15 (1.72)	–0.36 (0.58)	N/A[b]	89	.55[c]	0.02
PP[d]	–0.18 (3.64)	–0.37 (1.05)	–0.29 (1.72)	–0.47 (0.61)	0.27 (16)	N/A	.40	0.07
Δ mean global displacement								
ITT	–0.14 (4.23)	–0.56 (1.12)	–0.18 (2.35)	–0.46 (0.80)	N/A	83	.66[c]	0.07
PP	–0.13 (4.23)	–0.68 (1.39)	–0.18 (2.32)	–0.39 (0.73)	N/A	39	.57[c]	0.03
Δ mean anterior-posterior velocity								
ITT	–0.21 (5.48)	–0.49 (1.26)	–0.41 (4.32)	–0.83 (1.29)	0.70 (25)	N/A	.25	0.14
PP	–0.66 (5.48)	–0.76 (1.53)	–0.22 (3.38)	–0.87 (1.17)	0.18 (16)	N/A	.43	0.05
Δ mean medio-lateral velocity								
ITT	–0.19 (2.37)	–0.33 (0.66)	–0.17 (4.12)	–0.23 (0.97)	–0.31 (25)	N/A	.62	0.06
PP	–0.26 (2.37)	–0.38 (0.73)	–0.17 (1.53)	–0.26 (0.52)	–0.41 (16)	N/A	.66	0.10
Δ mean global velocity								
ITT	–0.48 (6.39)	–0.65 (1.50)	–0.29 (5.99)	–0.92 (1.73)	0.44 (25)	N/A	.33	0.09
PP	–0.78 (6.39)	–0.93 (1.81)	–0.23 (3.64)	–0.98 (1.35)	0.06 (16)	N/A	.47	0.02

[a]ITT: intention-to-treat.

[b]N/A: not applicable.

[c]Wilcoxon rank-sum test.

[d]PP: per-protocol.

Movement Tasks

Comparisons of change between T2 and T3 in lumbar and hip movement during the movement tasks are shown in Table 5 and Figure 4. There was no significant difference in either the ITT or the PP comparisons in accordance with our hypotheses. However, for the lumbar spine there were small decreases in the deviation from the starting position during task performance in the control group and small increases in the intervention group. Thus, the results descriptively showed a trend opposing our predictions with respect to the lumbar spine for both the box lift and waiter bow tasks with moderate effect sizes.

Table 5. Directed group comparisons of change in the movement tasks between T2 and T3.

Outcome and analysis	Control		Intervention		Comparison			
	Median (range)	Mean (SD)	Median (range)	Mean (SD)	*t* test (*df*)	*W*	*P* value	*r*
∆ box lift lumbar spine								
ITT[a]	–3.05 (27.86)	–3.00 (8.61)	3.37 (45.08)	3.25 (12.10)	–1.56 (25)	N/A[b]	.93	0.30
PP[c]	–5.37 (27.80)	–5.05 (10.31)	6.69 (45.08)	6.03 (13.00)	–1.93 (15)	N/A	.96	0.45
∆ box lift hip								
ITT	0.31 (31.94)	–0.14 (8.52)	1.10 (42.64)	0.43 (12.04)	–0.14 (25)	N/A	.44	0.03
PP	2.48 (31.94)	0.84 (10.21)	–2.27 (42.09)	–2.07 (12.11)	–0.53 (15)	N/A	.70	0.14
∆ waiter bow lumbar spine								
ITT	–1.12 (15.20)	–2.50 (5.22)	1.91 (28.04)	3.16 (8.22)	–2.15 (25)	N/A	.98	0.40
PP	–1.12 (15.20)	–2.62 (5.51)	1.91 (24.28)	3.07 (7.14)	–1.82 (15)	N/A	.96	0.42
∆ waiter bow hip								
ITT	–0.85 (25.43)	1.50 (6.83)	1.32 (37.10)	–0.48 (10.33)	N/A	92	.53[d]	0.01
PP	–1.81 (25.43)	2.46 (8.93)	1.32 (27.99)	0.41 (7.94)	N/A	33	.41[d]	0.07

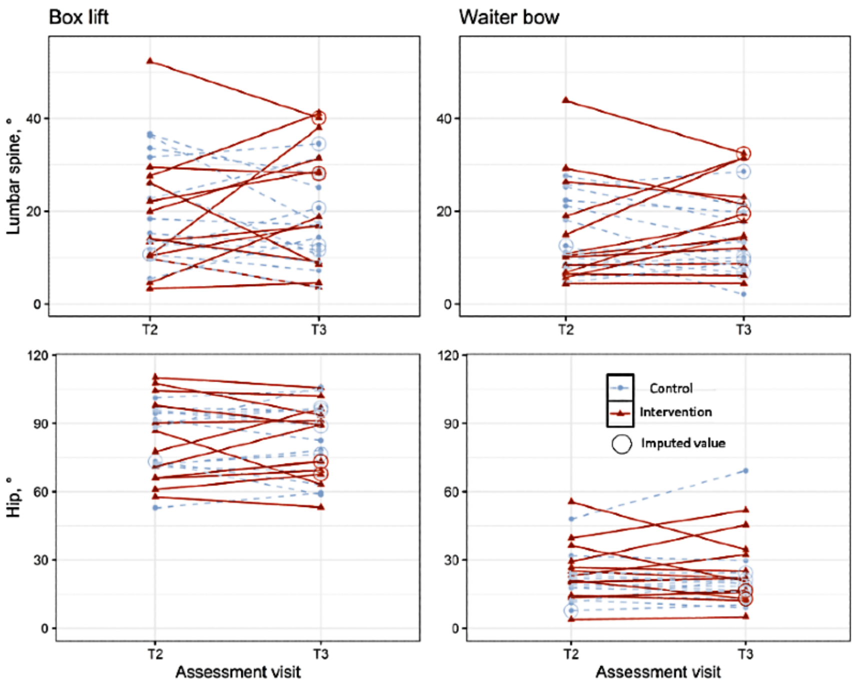
[a]ITT: intention-to-treat.

[b]N/A: not applicable.

[c]PP: per-protocol.

[d]Wilcoxon rank-sum test.

Figure 4. Lumbar spine and hip movement in degrees during the box lift and waiter bow task at T2 and T3. Data as included in the intention-to-treat analysis (control: n=14 and intervention n=13). Red triangles and solid lines show data of participants in the intervention group. Blue points and dashed lines show data of participants in the control group. Circled values represent imputed data.



Participant-Reported Outcomes

The groups did not significantly differ in the change of scores in participant-reported outcomes in the ITT or PP analysis (Table 6 and Figure 5).

Table 6. Directed group comparisons of change in the participant-reported outcomes between T2 and T3.

Outcome and analysis	Control		Intervention		Comparison			
	Median (range)	Mean (SD)	Median (range)	Mean (SD)	*t* test (*df*)	*W*	*P* value	*r*
Δ pain intensity numeric rating scale								
ITT[a]	0.00 (5.00)	0.14 (1.18)	0.00 (4.00)	–0.12 (1.12)	0.58 (25)	N/A[b]	.28	0.12
PP[c]	0.00 (5.00)	0.44 (1.33)	0.00 (3.00)	–0.44 (0.88)	1.67 (16)	N/A	.06	0.38
Δ Roland Morris disability questionnaire								
ITT	0.00 (8.50)	–0.25 (1.90)	0.00 (10.00)	0.12 (2.26)	N/A	88	.57[d]	0.03
PP	0.00 (4.00)	0.33 (1.41)	–1.00 (6.00)	–0.55 (1.67)	1.22 (16)	N/A	.12	0.29
Δ Tampa Scale for Kinesiophobia, 11-item version								
ITT	–0.75 (9.00)	–0.04 (2.63)	–1.00 (11.50)	–0.88 (3.18)	0.76 (25)	N/A	.23	0.15
PP	–1.00	0.11 (3.18)	–1.00	–1.22 (3.46)	0.85 (16)	N/A	.20	0.21
Δ quality of life physical subscale								
ITT	0.20 (7.20)	0.23 (1.83)	1.60 (3.20)	1.11 (1.06)	N/A	63.5	.09[d]	0.26
PP	1.60 (7.20)	0.53 (2.23)	1.60 (2.40)	1.60 (0.80)	N/A	27.5	.13[d]	0.27
Δ quality of life psychological subscale								
ITT	0.00 (12.00)	–0.19 (2.65)	0.00 (4.00)	–2.56 (1.26)	N/A	101.5	.71[d]	0.10
PP	1.33 (12.00)	–0.15 (3.36)	0.00 (4.00)	–0.44 (1.33)	N/A	50	.81[d]	0.20
Δ quality of life social subscale								
ITT	0.00 (5.33)	0.62 (1.35)	0.00 (6.67)	–0.15 (1.81)	N/A	114.5	.89[d]	0.23
PP	0.00 (5.33)	0.89 (1.63)	0.00 (6.67)	0.30 (1.86)	0.72 (16)	N/A	.76	0.18
Δ quality of life environment subscale								
ITT	0.00 (5.33)	0.45 (1.32)	0.00 (4.00)	0.44 (1.58)	N/A	94	.57	0.03
PP	0.00 (5.33)	0.67 (1.60)	0.00 (4.00)	0.22 (1.60)	N/A	50	.82[d]	0.20

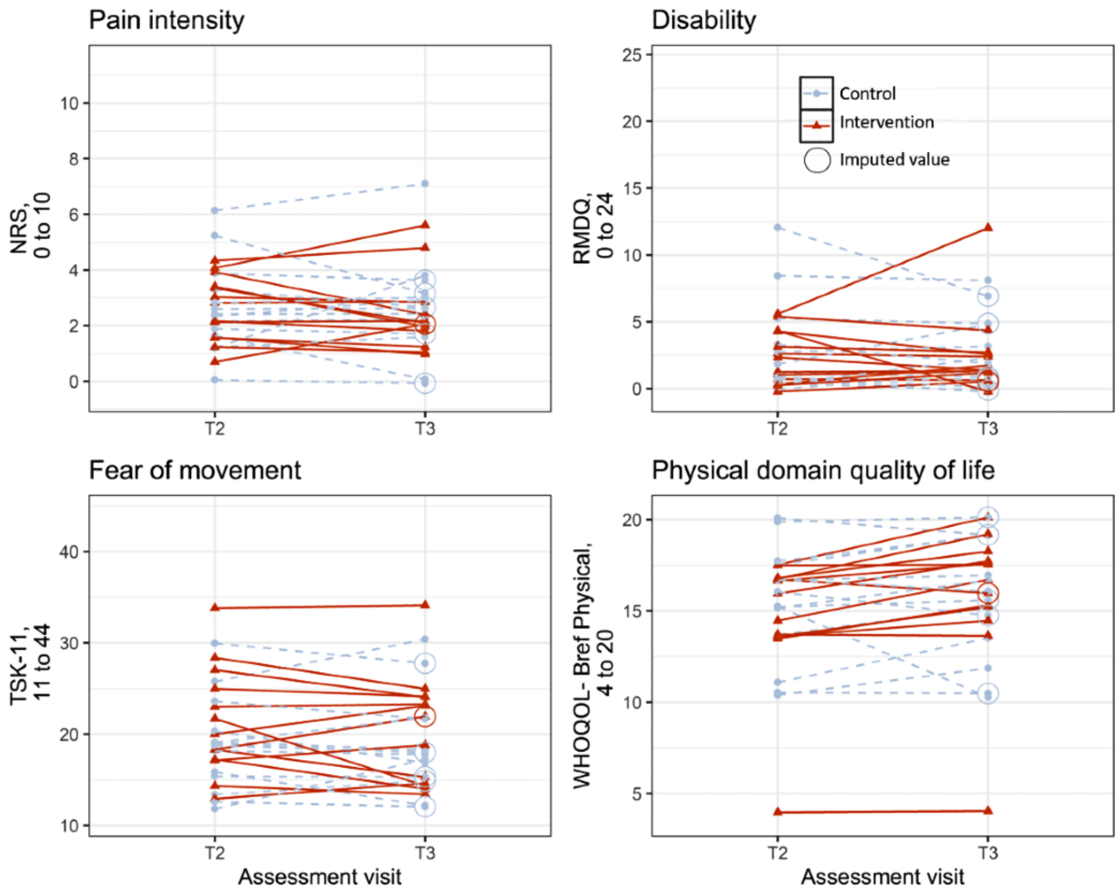
[a]ITT: intention-to-treat.

[b]N/A: not applicable.

[c]PP: per-protocol.

[d]Wilcoxon rank-sum test.

Figure 5. Scores of participant-reported outcomes for the assessment visits T2 and T3. Data as included in the intention-to-treat analysis (control: n=14 and intervention: n=13) are displayed. To represent all data in the graph despite exact overlap, small random values were added for the graphical representation of the data. Red triangles and solid lines show data of participants in the intervention group. Blue points and dashed lines show data of participants in the control group. Circled values represent imputed data. NRS: numeric rating scale; RMDQ: Roland Morris disability questionnaire; TSK-11: Tampa Scale for Kinesiophobia, 11-item version; WHOQOL-Bref Physical: Physical domain World Health Organization Quality of Life Questionnaire-short version.



Exploratory Comparisons Across All Assessment Visits

Exploratory analyses were conducted across all 4 assessment visits among a subset of participants who remained in the study until T4 (n=20).

Postural Balance

There were no main effects of group or significant interaction effects (group by time) for any of the postural balance variables (Table 7). For AP velocity and global velocity, there was each a significant main effect of assessment visit, but none of the post hoc comparisons for the individual assessment visits showed significant differences. Descriptively displacement and velocity parameters increased between T1 and T2 and decreased from T2 to T3. This is surprising, as we did not expect to see such fluctuations in balance across time for the entire group of participants.

Table 7. Effects of Group and Assessment Visit on Postural Balance parameters within two-way mixed ANOVA.

Outcome	Group			Assessment visit			Group×assessment visit		
	F test (*df*)	*P* value	η^2_G[a]	*F* test (*df*)	*P* value	η^2_G	*F* test (*df*)	*P* value	η^2_G
Mean anterior-posterior displacement	0.25 (1,18)	.63	0.01	1.51 (3,54)	.22	0.02	1.21 (3,54)	.32	0.01
Mean medio-lateral displacement	1.21 (1,18)	.29	0.05	2.52 (3,54)	.07	0.03	0.10 (3,54)	.96	0.00
Mean global displacement	0.71 (1,18)	.41	0.03	2.59 (3,54)	.06	0.02	0.51 (3,54)	.68	0.01
Mean anterior-posterior velocity	1.60 (1,18)	.22	0.07	3.51 (3,54)	.02	0.03	0.28 (3,54)	.84	0.00
Mean medio-lateral velocity	0.03 (1,18)	.87	0.00	2.07 (3,54)	.12	0.01	0.21 (3,54)	.89	0.00
Mean global velocity	0.50 (1,18)	.49	0.02	3.61 (3,54)	.02	0.03	0.12 (3,54)	.95	0.00
Box lift lumbar spine	0.00 (1,18)	.99	0.00	1.32 (3,54)	.28	0.02	1.56 (3,54)	.21	0.03
Box lift hip	0.06 (1,18)	.81	0.00	0.27 (3,54)	.85	0.00	0.91 (3,54)	.44	0.01
Waiter bow lumbar spine	0.54 (1,18)	.47	0.02	0.22 (1.89,34.01)	.79[b]	0.00	0.82 (1.89,34.01)	.44[b]	0.02
Waiter bow hip	0.15 (1,18)	.71	0.01	1.40 (3,54)	.24	0.02	0.42 (3,54)	.74	0.01
Pain intensity numeric rating scale	0.94 (1,18)	.35	0.02	1.34 (3,54)	.27	0.04	0.76 (3,54)	.52	0.02
Quality of life physical subscale	0.26 (1,18)	.62	0.01	5.55 (1.78,31.98)	.01[b]	0.04	1.08 (1.78,31.98)	.34[b]	0.01
Tampa Scale for Kinesiophobia, 11-item version	0.22 (1,18)	.64	0.01	1.05 (3,54)	.38	0.01	0.14 (3,54)	.93	0.00

[a]Generalized η^2.

[b]Greenhouse Geisser corrected.

Movement Tasks

There was no significant effect for group, assessment visit, and the interaction for lumbar spine or hip during the waiter bow and box lift tasks (Table 7).

Participant-Reported Outcomes

For the pain intensity NRS and fear of movement questionnaire, no significant effects for group, assessment visit, or their interaction were present (Table 7).

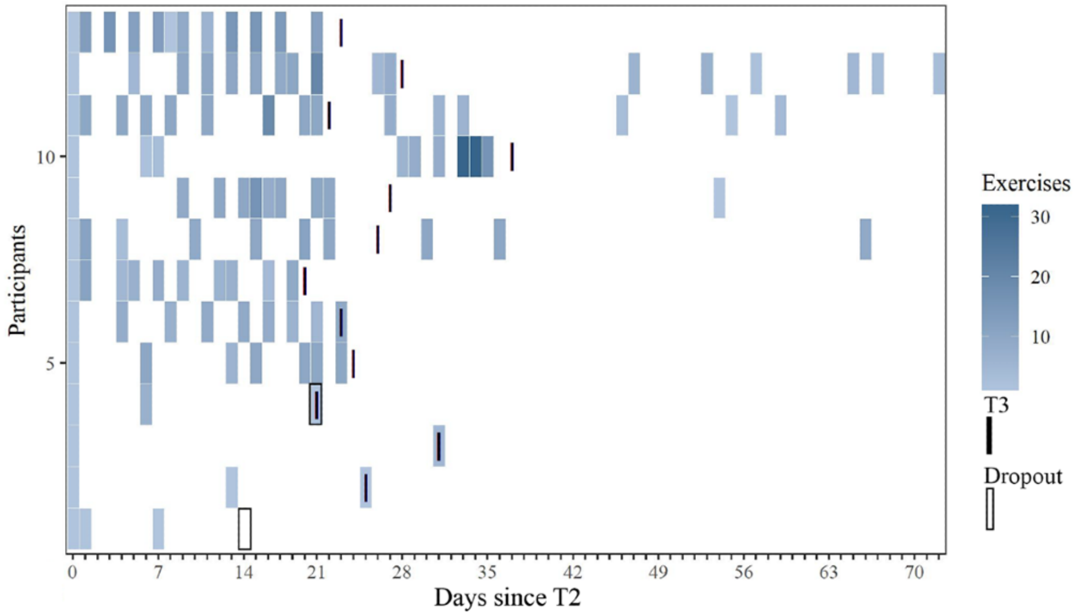
For the RMDQ scores, Friedman tests did not show significant differences across visits in the control group (χ^2_3=4.1; *P*=.25) or the intervention group (χ^2_3=6.0; *P*=.11). Bonferroni-corrected Wilcoxon rank-sum tests showed no difference among the groups at any assessment visit.

However, for the physical QOL, there was a significant main effect of assessment visit. Post hoc comparisons between assessment visits across both groups revealed that T3 scores were significantly higher than those at T2 (t_{19}=3.71; *P*=.009). Results for the social, psychological, and environmental QOLs are reported in Multimedia Appendix 6.

Adherence

Participants in the intervention group were instructed to complete a fixed set of 90 exercises between the assessments T2 and T3. Of these exercises, a median of 61% (55/90; range 2%-99%) were completed. As not all exercises were performed with the specified duration and frequency, and some participants performed the exercises that were provided from the device but were not intended as part of the program, effective time spent exercising differed from the completion of the program. In this period with a predefined schedule (T2-T3), participants exercised a median of 77.2% (139/180; range 3%-202%) of the targeted exercising duration of 180 minutes. The exercising time of 4 participants exceeded 180 minutes. During the intervention period with a schedule, a total of 7 participants performed a median of 9 exercises (minimum 1, maximum 41), equivalent to 17 minutes (minimum 2, maximum 109) that were not part of the program. In the intervention period without a schedule, of the 11 participants who had remained in the study, 4 (36%) participants performed a median of 27 exercises (minimum 1, maximum 29), equivalent to 82 minutes (minimum 2, maximum 101). An overview of the number of any exercise performed is provided in Figure 6.

Figure 6. Exercises that have been completed during the study per participant per day. Darker color indicates a larger number of exercises performed. All exercises, including exercises that have not been intended as a part of the schedule are displayed. Black bars show assessment visit T3 and black boxes show dropouts at or before T3.



Unintended Effects

There were no unintended effects that were related to the intervention. Although reasons for not adhering to the protocol were not assessed systematically and participants had been encouraged to contact the investigators with any difficulties, some participants in the intervention group reported problems with handling the devices. This included difficulties such as finding the right icon on the tablet and difficulties with the calibration of the IMUs and program failures of the tablet. These issues likely contributed to the low adherence of some participants.

Discussion

Principal Findings

Self-directed home exercising with feedback on trunk movements for a period of approximately 3 weeks did not enhance postural balance during quiet standing in study participants with LBP or significantly affect any other of the investigated outcomes. Comparisons of the groups with respect to the movement tasks showed, descriptively, a tendency toward slightly increased motion of the lumbar spine during both tasks in the intervention group, combined with a small reduction in the control group, which contradicted our predictions. Adherence to the scheduled exercising program was low. After the participants were no longer provided with a schedule to complete, only some participants kept using the training device repeatedly without specific instructions. Despite not showing intervention effects in this trial, it cannot be excluded that these interventions may still be beneficial when integrated into a therapy setting with patients. Furthermore, for other exercising interventions, it has been demonstrated that exercising could have more pronounced effects in patients than in other study participants with LBP [3]. A review showed that the results were positive for exercising with digital systems for LBP, when these exercises were delivered together with another intervention, but otherwise not [32].

Comparison With Previous Work

Postural Balance

In this study investigating an exercise intervention using mobile sensors under self-directed home conditions in people with moderate LBP, no improvement of postural balance during quiet standing was found. To our knowledge, no other studies using feedback on trunk movements and similar assessments of postural balance have been conducted with participants with LBP. In a study where exergaming with the Nintendo Wii was included into the treatment, participants were not able to maintain single-legged stance for longer than before the intervention [67]. In contrast, in a study with older participants with diverse chronic musculoskeletal complaints and an exergame that mainly focused on translations of the body weight, several postural balance parameters improved but not relative to participants who had performed similar exercises without gamification [68]. A meta-analysis on studies with older participants without complaints suggests that exergames affect different measures of postural balance positively, but an enhancement of postural balance assessed under stable, unperturbed conditions could not be confirmed either [69]. Consistent with these observations, differences observable at the level of the trunk may not necessarily translate to changes in COP-based assessments during quiet standing [70]. Postural balance regulation is the product of the complex interaction of different structures and systems, with the capacity to adapt to changing conditions [71]. Thus, adaptations in balance should be explored additionally under varying assessment conditions. For example, assessments of trunk balance during sitting may provide an isolated assessment of trunk control [7] and could possibly reveal more subtle changes. Unexpectedly, we observed changes in some postural balance parameters across assessment visits, but statistically significant differences between individual visits were not found, which may be owing to a lack of power. The descriptive pattern did not indicate a continuous trend that could have been interpreted as learning or other effects of repetition.

Movement Tasks

In this study, the average amount of lumbar spine movement observed was comparable with the values reported earlier by other researchers [46]. However, contrary to our expectations, descriptively, the participants in the intervention group showed small increases in movement in the lumbar region during the movement tasks compared with the control group, who showed comparable reductions. This was the case despite the instructions to not bend or extend the lumbar spine during the assessment. Nevertheless, if only the increase in lumbar spine motion in the intervention group independently of the decrease in the control group is considered, this increase was only during the box lift task in the PP analysis (6.03°) slightly larger than the minimally detectable change value of 5.3°, which was described in the study our assessments were adapted from, by Matheve et al [46]. These descriptive observations may indicate that the intervention might rather impact mobility than the precise control of the lumbar spine. This interpretation would be in line with the finding of other investigators who found an expansion in ROM after a similar intervention but did not clearly state whether there was a difference in comparison with the group without the exercises [9]. No impact on an intervention on ROM was found in another study [10]. A recent meta-analysis challenged the assumption that people with LBP tend to bend their spine more in lift tasks [72], and restrictions in ROM in the lumbar region of people with LBP have already been described [14]. Furthermore, it was found that during a box lift task, participants with chronic LBP moved less in the lumbar region than participants without LBP [73]. Hence, an increase in movement in the lumbar spine would not necessarily constitute an undesirable outcome. Future studies should clarify the role of lumbar spine posture during lifting and the influence of exercising interventions on lifting behavior.

Participant-Reported Outcomes

There were no statistically significant differences between the change scores of groups in participant-reported outcomes. In contrast, some other studies investigating similar interventions found positive effects on pain assessments [9,10]. Nevertheless, it should be considered that pain NRSs could be error prone to some degree [74], and the power in this study may have been insufficient to detect an effect. A reduction in pain intensity across both groups was observed within the first 3 weeks of the study, where no intervention was provided. This effect could possibly be caused by participants initiating study participation during periods in which their pain was perceived as slightly worse than usual. The amount of pain appeared to be comparable with the value of approximately 2.5 obtained from visual analog scales, which had been reported in a review that revealed postural balance differences between people with and without LBP [17]. As we have observed, a small study found that exercises with postural feedback in addition to standard care was not superior in reducing disability than the usual treatment alone [11]. This is in contrast with the results of a different study, which indicated that disability could be improved [10]. However, in that specific study the feedback from the wearable device was not only provided during exercises but also during everyday activities [10]. In our study, the RMDQ mean scores were generally low, which may have limited the range of possible improvement. We did not find an intervention effect on physical QOL. Contrary to this result, in another study, an intervention effect on the physical subscale of the QOL measure short form-36 was observed [9]. A small trial using comparisons of the scores, before and following a similar intervention, showed significant improvements in pain, disability, and QOL [22]. We did not find an effect of the intervention on fear of movement, and neither an intervention effect was found in another study [10].

Adherence

A particular strength of this work is the combination of an investigation of the exercising program at home in a period with a set training schedule and in a second interval, where participants could exercise as they wished. Comparison with studies on related interventions in home-based settings, which are considered similar, are difficult, as in a study a combined value including other exercises was investigated [11], and in another study, self-report methods had failed [6]. Furthermore, in a study that investigated exercising with the Nintendo Wii, completion of 71% of the advised time was achieved [36], which is comparable with the median of 77% obtained in this study. Nevertheless, the schedule provided was much more demanding and additional measures were used to improve adherence in the other study [36]. The comparatively short 3-week exercise period in combination with the low adherence may have limited the effects of the intervention. This assumption is in line with the findings of a review on virtual reality interventions, which indicated that interventions with more sessions may be more successful [33].

Results on adherence, considering time spent exercising, was more favorable than the number of exercises performed as requested by the investigators. Some participants exercised even more than required but did not follow the instructions precisely. In some cases, participants may have forgotten to reset the play time from the default 4 to 2 minutes, or the game may have motivated the participants to explore additional contents and may have provided stronger guidance than the instructions from the investigators. The 6-week period with flexible exercising opportunity resembled more closely to the conditions under which participants would be using such tools without the connection to a therapeutic setting. Only few participants kept exercising after T3. These results may imply that such interventions only get adopted by a small number of people or might rather be integrated within supervised programs on site. Within the setup of this study, it could not be determined whether the provided schedule or the participants' commitment to comply with the study protocol resulted in higher amounts of exercises between T2 and T3. Future studies should investigate if and how automated scheduling options can help improve adherence and how they should be integrated. Although the Valedo app offers the option to generate an exercising plan, such functions could be placed more prominently. In addition, the context and the kind of assistance required should receive more attention. Recently, for example, blended therapy setups for people with LBP have been explored [75].

Limitations

The low number of participants who could be recruited is an important limitation of this study. Although the individual components of the study protocol may not have been too time consuming, the overall effort associated with study participation, including diary methods and activity tracking not discussed in this manuscript, may have been a cause for low recruitment and retention rates. These assessments were included to answer additional research questions beyond the scope of a single manuscript but contributed to the effort by the study participants. In line with this presumption, reasons given for withdrawal were frequently related to time investment or perceived benefit and effort. This might also have contributed to the low adherence to the intervention. Time intervals between assessments were slightly stretched owing to frequent requests from participants to reschedule appointments, as the study participation was not part of an official treatment program and therefore often had to take place often outside of the working hours of the participants. The assessment of the movement tasks was preceded and followed by the participants shifting their weight to the sides and back, to time-synchronize the data from the IMUs with data collected simultaneously from the force platform. Although supporting analyses of change between T2 and T3, where data from trials that appeared to be performed from an unstable starting position were removed, appeared similar, this setup could have influenced the results. The study participants could not be blinded, and with most assessments conducted in the field, it could not be ruled out that the participants completed all exercises themselves. In a case with a particularly high number of exercises, it was suspected that other people may have completed some of those exercises. We did not record the use of pain medication during the study; therefore, confounding effects of pain medication could not be ruled out. Further, although we consider the availability of the questionnaires in different languages as a strength, this setup may have caused inconsistencies between the questionnaires that different participants received.

Conclusions

The results obtained in this study indicate that exercising with feedback on trunk movements alone may not influence postural balance during quiet standing in people with only moderate LBP intensity and disability. No significant intervention effects on lumbar spine and hip movement, pain intensity, disability, QOL domains, and fear of movement were observed. These results must be seen within the context of a small sample size and low adherence to the intervention, resulting in low doses of exercise. More work in this field is required; for example, to establish the effect of interventions using feedback on trunk movements in people more severely affected by LBP and clarify more proximal effects on trunk movement properties. As the amount of exercising dropped substantially in the intervention period without a schedule, future studies should investigate the impact of different scheduling options and explore such interventions in combination with other therapeutic settings or other strategies to improve adherence.

Acknowledgments

The authors thank Ramon Glättli, Kim Graf, Cinzia Maschio, Adrian Stutz, Tina Wunderlin, and Katharina Zahoranszky who assisted in the outcome assessments. The authors also thank Professor Oliver Distler for participating as a study physician and the participants who invested their time and provided the data. The authors would also like to thank Hocoma AG for providing the Valedo Pro used for the movement tasks and Lars Lünenburger (Hocoma AG) for providing support in exporting exercise adherence data from the Valedo app. The authors like to thank the Swiss National Science Foundation for funding this study (grant 167302) within the National Research Program 75 *Big Data*. Lars Lünenburger contributed to the writing of the grant proposal which funded this study and was administered by WK.

Authors' Contributions

WK (sponsor) and JS (principal investigator) initiated the study. RP, RK, and JS were consulted as physiotherapists, if eligibility could not be determined by simple questions. AM analyzed the data and drafted the first version of the manuscript. All authors have edited the manuscript and read and agreed to the publication of the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth; version 1.6.1) submission publication form.

[PDF File (Adobe PDF File), 1273 KB - games_v10i2e31685_app1.pdf]

Multimedia Appendix 2

Graphical representation of center of pressure parameters.

[PDF File (Adobe PDF File), 228 KB - games_v10i2e31685_app2.pdf]

Multimedia Appendix 3
Video showing the exercises.
[MOV File , 198831 KB - games_v10i2e31685_app3.mov]

Multimedia Appendix 4
Descriptive statistics and baseline comparisons.
[PDF File (Adobe PDF File), 238 KB - games_v10i2e31685_app4.pdf]

Multimedia Appendix 5
Data used for the intention-to-treat analysis of center of pressure data.
[PDF File (Adobe PDF File), 324 KB - games_v10i2e31685_app5.pdf]

Multimedia Appendix 6
Additional comparisons including all assessment visits.
[PDF File (Adobe PDF File), 103 KB - games_v10i2e31685_app6.pdf]

References

1. Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, et al. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann Transl Med 2020 Mar;8(6):299 [FREE Full text] [doi: 10.21037/atm.2020.02.175] [Medline: 32355743]
2. Lin I, Wiles L, Waller R, Goucke R, Nagree Y, Gibberd M, et al. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: systematic review. Br J Sports Med 2020 Jan;54(2):79-86. [doi: 10.1136/bjsports-2018-099878] [Medline: 30826805]
3. Hayden JA, van Tulder MW, Malmivaara A, Koes BW. Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev 2005 Jul 20(3):CD000335. [doi: 10.1002/14651858.CD000335.pub2] [Medline: 16034851]
4. van Middelkoop M, Rubinstein SM, Verhagen AP, Ostelo RW, Koes BW, van Tulder MW. Exercise therapy for chronic nonspecific low-back pain. Best Pract Res Clin Rheumatol 2010 Apr;24(2):193-204. [doi: 10.1016/j.berh.2010.01.002] [Medline: 20227641]
5. Saragiotto BT, Maher CG, Yamato TP, Costa LO, Menezes Costa LC, Ostelo RW, et al. Motor control exercise for chronic non-specific low-back pain. Cochrane Database Syst Rev 2016 Jan 08;2016(1):CD012004 [FREE Full text] [doi: 10.1002/14651858.CD012004] [Medline: 26742533]
6. Matheve T, Claes G, Olivieri E, Timmermans A. Serious gaming to support exercise therapy for patients with chronic nonspecific low back pain: a feasibility study. Games Health J 2018 Aug;7(4):262-270. [doi: 10.1089/g4h.2017.0173] [Medline: 29688818]
7. van Dieën JH, Reeves NP, Kawchuk G, van Dillen LR, Hodges PW. Motor control changes in low back pain: divergence in presentations and mechanisms. J Orthop Sports Phys Ther 2019 Jun;49(6):370-379 [FREE Full text] [doi: 10.2519/jospt.2019.7917] [Medline: 29895230]
8. Schmidt R, Lee T, Winsetein C, Wulf G, Zelaznik H. Augmented feedback. In: Motor Control and Learning: A Behavioral Emphasis (6th Edition). Champaign IL: Human Kinetics; 2018:341-373.
9. Magnusson ML, Chow DH, Diamandopoulos Z, Pope MH. Motor control learning in chronic low back pain. Spine (Phila Pa 1976) 2008 Jul 15;33(16):E532-E538. [doi: 10.1097/BRS.0b013e31817dfd9a] [Medline: 18628693]
10. Kent P, Laird R, Haines T. The effect of changing movement and posture using motion-sensor biofeedback, versus guidelines-based care, on the clinical outcomes of people with sub-acute or chronic low back pain-a multicentre, cluster-randomised, placebo-controlled, pilot trial. BMC Musculoskelet Disord 2015 May 29;16:131 [FREE Full text] [doi: 10.1186/s12891-015-0591-5] [Medline: 26022102]
11. Hügli AS, Ernst MJ, Kool J, Rast FM, Rausch-Osthoff A, Mannig A, et al. Adherence to home exercises in non-specific low back pain. A randomised controlled pilot trial. J Bodyw Mov Ther 2015 Jan;19(1):177-185. [doi: 10.1016/j.jbmt.2014.11.017] [Medline: 25603757]
12. Ferreira ML, Machado G, Latimer J, Maher C, Ferreira PH, Smeets RJ. Factors defining care-seeking in low back pain--a meta-analysis of population based surveys. Eur J Pain 2010 Aug;14(7):747.e1-747.e7. [doi: 10.1016/j.ejpain.2009.11.005] [Medline: 20036168]
13. Mazaheri M, Coenen P, Parnianpour M, Kiers H, van Dieën JH. Low back pain and postural sway during quiet standing with and without sensory manipulation: a systematic review. Gait Posture 2013 Jan;37(1):12-22. [doi: 10.1016/j.gaitpost.2012.06.013] [Medline: 22796243]
14. Laird RA, Gilbert J, Kent P, Keating JL. Comparing lumbo-pelvic kinematics in people with and without back pain: a systematic review and meta-analysis. BMC Musculoskelet Disord 2014 Jul 10;15:229 [FREE Full text] [doi: 10.1186/1471-2474-15-229] [Medline: 25012528]

15. Sadler SG, Spink MJ, Ho A, De Jonge XJ, Chuter VH. Restriction in lateral bending range of motion, lumbar lordosis, and hamstring flexibility predicts the development of low back pain: a systematic review of prospective cohort studies. BMC Musculoskelet Disord 2017 May 05;18(1):179 [FREE Full text] [doi: 10.1186/s12891-017-1534-0] [Medline: 28476110]
16. Berenshteyn Y, Gibson K, Hackett GC, Trem AB, Wilhelm M. Is standing balance altered in individuals with chronic low back pain? A systematic review. Disabil Rehabil 2019 Jun;41(13):1514-1523. [doi: 10.1080/09638288.2018.1433240] [Medline: 29382241]
17. Ruhe A, Fejer R, Walker B. Center of pressure excursion as a measure of balance performance in patients with non-specific low back pain compared to healthy controls: a systematic review of the literature. Eur Spine J 2011 Mar;20(3):358-368 [FREE Full text] [doi: 10.1007/s00586-010-1543-2] [Medline: 20721676]
18. Willigenburg NW, Kingma I, Hoozemans MJ, van Dieën JH. Precision control of trunk movement in low back pain patients. Hum Mov Sci 2013 Feb;32(1):228-239. [doi: 10.1016/j.humov.2012.12.007] [Medline: 23427936]
19. Alsubaie AM, Martinez-Valdes E, De Nunzio AM, Falla D. Trunk control during repetitive sagittal movements following a real-time tracking task in people with chronic low back pain. J Electromyogr Kinesiol 2021 Apr;57:102533. [doi: 10.1016/j.jelekin.2021.102533] [Medline: 33621756]
20. Matheve T, Brumagne S, Demoulin C, Timmermans A. Sensor-based postural feedback is more effective than conventional feedback to improve lumbopelvic movement control in patients with chronic low back pain: a randomised controlled trial. J Neuroeng Rehabil 2018 Sep 26;15(1):85 [FREE Full text] [doi: 10.1186/s12984-018-0423-6] [Medline: 30253807]
21. Korakakis V, O'Sullivan K, Kotsifaki A, Sotiralis Y, Giakas G. Lumbo-pelvic proprioception in sitting is impaired in subgroups of low back pain-But the clinical utility of the differences is unclear. A systematic review and meta-analysis. PLoS One 2021;16(4):e0250673 [FREE Full text] [doi: 10.1371/journal.pone.0250673] [Medline: 33901255]
22. Alemanno F, Houdayer E, Emedoli D, Locatelli M, Mortini P, Mandelli C, et al. Efficacy of virtual reality to reduce chronic low back pain: proof-of-concept of a non-pharmacological approach on pain, quality of life, neuropsychological and functional outcome. PLoS One 2019;14(5):e0216858 [FREE Full text] [doi: 10.1371/journal.pone.0216858] [Medline: 31120892]
23. Low DC, Walsh GS, Arkesteijn M. Effectiveness of exercise interventions to improve postural control in older adults: a systematic review and meta-analyses of centre of pressure measurements. Sports Med 2017 Jan;47(1):101-112 [FREE Full text] [doi: 10.1007/s40279-016-0559-0] [Medline: 27245061]
24. Casonatto J, Yamacita CM. Pilates exercise and postural balance in older adults: a systematic review and meta-analysis of randomized controlled trials. Complement Ther Med 2020 Jan;48:102232. [doi: 10.1016/j.ctim.2019.102232] [Medline: 31987246]
25. Manogharan S, Kongsted A, Ferreira ML, Hancock MJ. Do older adults with chronic low back pain differ from younger adults in regards to baseline characteristics and prognosis? Eur J Pain 2017 May;21(5):866-873. [doi: 10.1002/ejp.989] [Medline: 28295893]
26. Ferreira M, Ferreira P, Henschke N, Kamper S, Koes B, Hayden J, et al. Age does not modify the effects of treatment on pain in patients with low back pain: secondary analyses of randomized clinical trials. Eur J Pain 2014;18(7):932-938. [doi: 10.1002/j.1532-2149.2013.00438.x]
27. Stamenkovic A, Clark BC, Pidcoe PE, van der Veen SM, France CR, Russ DW, et al. Distinguishing chronic low back pain in young adults with mild to moderate pain and disability using trunk compliance. Sci Rep 2021 Apr 07;11(1):7592 [FREE Full text] [doi: 10.1038/s41598-021-87138-6] [Medline: 33828171]
28. Lomond KV, Henry SM, Hitt JR, DeSarno MJ, Bunn JY. Altered postural responses persist following physical therapy of general versus specific trunk exercises in people with low back pain. Man Ther 2014 Oct;19(5):425-432 [FREE Full text] [doi: 10.1016/j.math.2014.04.007] [Medline: 24853255]
29. Areeudomwong P, Buttagat V. Proprioceptive neuromuscular facilitation training improves pain-related and balance outcomes in working-age patients with chronic low back pain: a randomized controlled trial. Braz J Phys Ther 2019;23(5):428-436 [FREE Full text] [doi: 10.1016/j.bjpt.2018.10.005] [Medline: 30361077]
30. Gholami Borujeni B, Yalfani A. Reduction of postural sway in athletes with chronic low back pain through eight weeks of inspiratory muscle training: a randomized controlled trial. Clin Biomech (Bristol, Avon) 2019 Oct;69:215-220. [doi: 10.1016/j.clinbiomech.2019.09.006] [Medline: 31614295]
31. McCaskey MA, Wirth B, Schuster-Amft C, de Bruin ED. Postural sensorimotor training versus sham exercise in physiotherapy of patients with chronic non-specific low back pain: an exploratory randomised controlled trial. PLoS One 2018;13(3):e0193358 [FREE Full text] [doi: 10.1371/journal.pone.0193358] [Medline: 29522571]
32. Matheve T, Brumagne S, Timmermans AA. The effectiveness of technology-supported exercise therapy for low back pain: a systematic review. Am J Phys Med Rehabil 2017 May;96(5):347-356. [doi: 10.1097/PHM.0000000000000615] [Medline: 27584143]
33. Bordeleau M, Stamenkovic A, Tardif P, 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]
34. Peebles AT, van der Veen S, Stamenkovic A, France CR, Pidcoe PE, Thomas JS. A virtual reality game suite for graded rehabilitation in patients with low back pain and a high fear of movement: within-subject comparative study. JMIR Serious Games 2022 Mar 23;10(1):e32027 [FREE Full text] [doi: 10.2196/32027] [Medline: 35319471]

35. Kim S, Min W, Kim J, Lee B. The effects of VR-based Wii fit yoga on physical function in middle-aged female LBP patients. J Phys Ther Sci 2014 Apr;26(4):549-552 [FREE Full text] [doi: 10.1589/jpts.26.549] [Medline: 24764631]
36. Zadro JR, Shirley D, Simic M, Mousavi SJ, Ceprnja D, Maka K, et al. Video-game-based exercises for older people with chronic low back pain: a randomized controlledtable trial (GAMEBACK). Phys Ther 2019 Jan 01;99(1):14-27. [doi: 10.1093/ptj/pzy112] [Medline: 30247715]
37. Ribeiro DC, Milosavljevic S, Terry J, Abbott JH. Effectiveness of a lumbopelvic monitor and feedback device to change postural behaviour: the ELF cluster randomised controlled trial. Occup Environ Med 2020 Jul;77(7):462-469. [doi: 10.1136/oemed-2019-106293] [Medline: 32253227]
38. Ribeiro DC, Sole G, Abbott JH, Milosavljevic S. The effectiveness of a lumbopelvic monitor and feedback device to change postural behavior: a feasibility randomized controlled trial. J Orthop Sports Phys Ther 2014 Sep;44(9):702-711. [doi: 10.2519/jospt.2014.5009] [Medline: 25098195]
39. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. J Pain 2016 Dec;17(12):1302-1317 [FREE Full text] [doi: 10.1016/j.jpain.2016.08.011] [Medline: 27616607]
40. Ciabattoni L, Ferracuti F, Lazzaro G, Romeo L, Verdini F. Serious gaming approach for physical activity monitoring: a visual feedback based on quantitative evaluation. In: Proceedings of the 2016 IEEE 6th International Conference on Consumer Electronics - Berlin (ICCE-Berlin). 2016 Presented at: 2016 IEEE 6th International Conference on Consumer Electronics - Berlin (ICCE-Berlin); Sep 5-7, 2016; Berlin, Germany p. 209-213. [doi: 10.1109/icce-berlin.2016.7684757]
41. van der Veen SM, Bordeleau M, Pidcoe PE, France CR, Thomas JS. Agreement analysis between vive and vicon systems to monitor lumbar postural changes. Sensors (Basel) 2019 Aug 21;19(17):3632 [FREE Full text] [doi: 10.3390/s19173632] [Medline: 31438520]
42. Alazba A, Al-Khalifa H, AlSobayel H. RabbitRun: an immersive virtual reality game for promoting physical activities among people with low back pain. Technologies 2018 Dec 20;7(1):2. [doi: 10.3390/technologies7010002]
43. 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]
44. Meinke A, Peters R, Knols R, Karlen W, Swanenburg J. Exergaming using postural feedback from wearable sensors and exercise therapy to improve postural balance in people with nonspecific low back pain: protocol for a factorial pilot randomized controlled trial. JMIR Res Protoc 2021 Aug 26;10(8):e26982 [FREE Full text] [doi: 10.2196/26982] [Medline: 34435954]
45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
46. Matheve T, De Baets L, Rast F, Bauer C, Timmermans A. Within/between-session reliability and agreement of lumbopelvic kinematics in the sagittal plane during functional movement control tasks in healthy persons. Musculoskelet Sci Pract 2018 Feb;33:90-98. [doi: 10.1016/j.msksp.2017.08.005] [Medline: 28844565]
47. Valedo user manual (english). Hocoma. 2018. URL: https://hocoma.b-cdn.net/wp-content/uploads/2019/04/09081248/VH-UM1.4-en-1604.pdf [accessed 2020-12-18]
48. Bakeman R. Recommended effect size statistics for repeated measures designs. Behav Res Methods 2005 Aug;37(3):379-384. [doi: 10.3758/bf03192707] [Medline: 16405133]
49. Snow G. blockrand: Randomization for Block Random Clinical Trials. CRAN. URL: https://cran.r-project.org/web/packages/blockrand/index.html [accessed 2020-03-28]
50. Kassambara A. Pipe-friendly framework for basic statistical tests. rstatix. URL: https://rpkgs.datanovia.com/rstatix/ [accessed 2021-05-10]
51. 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. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
52. Ruhe A, Fejer R, Walker B. The test-retest reliability of centre of pressure measures in bipedal static task conditions--a systematic review of the literature. Gait Posture 2010 Oct;32(4):436-445. [doi: 10.1016/j.gaitpost.2010.09.012] [Medline: 20947353]
53. Winter D. Human balance and posture control during standing and walking. Gait Posture 1995 Dec;3(4):193-214. [doi: 10.1016/0966-6362(96)82849-9]
54. Prieto TE, Myklebust JB, Hoffmann RG, Lovett EG, Myklebust BM. Measures of postural steadiness: differences between healthy young and elderly adults. IEEE Trans Biomed Eng 1996 Sep;43(9):956-966. [doi: 10.1109/10.532130] [Medline: 9214811]
55. Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in non-specific low back pain. Eur Spine J 2015 Jun;24(6):1127-1142. [doi: 10.1007/s00586-015-3892-3] [Medline: 25841358]

56. Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LO, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. Pain 2018 Mar;159(3):481-495 [FREE Full text] [doi: 10.1097/j.pain.0000000000001117] [Medline: 29194127]
57. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976) 1983 Mar;8(2):141-144. [doi: 10.1097/00007632-198303000-00004] [Medline: 6222486]
58. Wiesinger GF, Nuhr M, Quittan M, Ebenbichler G, Wölfl G, Fialka-Moser V. Cross-cultural adaptation of the Roland-Morris questionnaire for German-speaking patients with low back pain. Spine (Phila Pa 1976) 1999 Jun 01;24(11):1099-1103. [doi: 10.1097/00007632-199906010-00009] [Medline: 10361659]
59. Chiarotto A, Roorda LD, Crins MH, Boers M, Ostelo RW, Terwee CB. PROMIS physical function short forms display item- and scale-level characteristics at least as good as the Roland Morris disability questionnaire in patients with chronic low back pain. Arch Phys Med Rehabil 2020 Feb;101(2):297-308. [doi: 10.1016/j.apmr.2019.09.018] [Medline: 31689418]
60. The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
61. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
62. Woby SR, Roach NK, Urmston M, Watson PJ. Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia. Pain 2005 Sep;117(1-2):137-144. [doi: 10.1016/j.pain.2005.05.029] [Medline: 16055269]
63. Rusu AC, Kreddig N, Hallner D, Hülsebusch J, Hasenbring MI. Fear of movement/(Re)injury in low back pain: confirmatory validation of a German version of the Tampa Scale for Kinesiophobia. BMC Musculoskelet Disord 2014 Aug 19;15:280 [FREE Full text] [doi: 10.1186/1471-2474-15-280] [Medline: 25138111]
64. Mair P, Wilcox R. Robust statistical methods in R using the WRS2 package. Behav Res Methods 2020 Apr;52(2):464-488. [doi: 10.3758/s13428-019-01246-w] [Medline: 31152384]
65. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
66. Lenhard W, Lenhard A. Berechnung von Effektstärken. Psychometrica. URL: https://www.psychometrica.de/effektstaerke.html [accessed 2021-12-07]
67. 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]
68. Ditchburn J, van Schaik P, Dixon J, MacSween A, Martin D. The effects of exergaming on pain, postural control, technology acceptance and flow experience in older people with chronic musculoskeletal pain: a randomised controlled trial. BMC Sports Sci Med Rehabil 2020;12:63 [FREE Full text] [doi: 10.1186/s13102-020-00211-x] [Medline: 33062284]
69. 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;9(1):11-23. [doi: 10.1089/g4h.2019.0016] [Medline: 31800322]
70. Schelldorfer S, Ernst MJ, Rast FM, Bauer CM, Meichtry A, Kool J. Low back pain and postural control, effects of task difficulty on centre of pressure and spinal kinematics. Gait Posture 2015 Jan;41(1):112-118. [doi: 10.1016/j.gaitpost.2014.09.004] [Medline: 25270326]
71. Shumway-Cook A, Woollacott MH. Normal postural control. In: Motor Control: Translating Research into Clinical Practice (5th ed). Philadelphia: Wolters Kluwer; 2017:153-182.
72. Saraceni N, Kent P, Ng L, Campbell A, Straker L, O'Sullivan P. To flex or not to flex? Is there a relationship between lumbar spine flexion during lifting and low back pain? A systematic review with meta-analysis. J Orthop Sports Phys Ther 2020 Mar;50(3):121-130. [doi: 10.2519/jospt.2020.9218] [Medline: 31775556]
73. 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]
74. Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement properties of visual analogue scale, numeric rating scale, and pain severity subscale of the brief pain inventory in patients with low back pain: a systematic review. J Pain 2019 Mar;20(3):245-263. [doi: 10.1016/j.jpain.2018.07.009] [Medline: 30099210]
75. Kloek CJ, van Tilburg ML, Staal JB, Veenhof C, Bossen D. Development and proof of concept of a blended physiotherapeutic intervention for patients with non-specific low back pain. Physiotherapy 2019 Dec;105(4):483-491 [FREE Full text] [doi: 10.1016/j.physio.2018.12.006] [Medline: 31031023]

Abbreviations

AP: anterior-posterior
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

COP: center of pressure
IMU: inertial measurement unit
ITT: intention-to-treat
LBP: low back pain
NRS: numeric rating scale
PP: per-protocol
QOL: quality of life
REDCap: Research Electronic Data Capture
RMDQ: Roland Morris disability questionnaire
ROM: range of motion
TIDieR: Template for Intervention Description and Replication
WHOQOL-Bref: World Health Organization Quality of Life Questionnaire-short version

Edited by N Zary; submitted 30.06.21; peer-reviewed by M Wittstein, A Stamenkovic, C Victor; comments to author 12.10.21; revised version received 17.01.22; accepted 14.04.22; published 10.06.22.

Please cite as:
Meinke A, Peters R, Knols RH, Swanenburg J, Karlen W
Feedback on Trunk Movements From an Electronic Game to Improve Postural Balance in People With Nonspecific Low Back Pain: Pilot Randomized Controlled Trial
JMIR Serious Games 2022;10(2):e31685
URL: https://games.jmir.org/2022/2/e31685
doi:10.2196/31685
PMID:35687390

Original Paper

Designing a Serious Game (Above Water) for Stigma Reduction Surrounding Mental Health: Semistructured Interview Study With Expert Participants

Rina R Wehbe[1,2,3], BSc, MSc, PhD; Colin Whaley[3,4,5], BSc, MSc; Yasaman Eskandari[2,3,5,6], BSc; Ally Suarez[2,3,7,8]; Lennart E Nacke[3,8], PhD; Jessica Hammer[9], PhD; Edward Lank[2,10†], PhD

[1]Human Computer Interaction for Social Good (HCI4GOOD), Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada

[2]Methods Group, Human Computer Interaction Lab, Cheriton School of Computer Science, Faculty of Mathematics, University of Waterloo, Waterloo, ON, Canada

[3]The Games Institute, University of Waterloo, Waterloo, ON, Canada

[4]Michael G DeGroote School of Medicine, McMaster University, Hamilton, ON, Canada

[5]School of Pharmacy, University of Waterloo, Waterloo, ON, Canada

[6]Faculty of Science, University of Waterloo, Waterloo, ON, Canada

[7]Faculty of Applied Health Sciences, University of Waterloo, Waterloo, ON, Canada

[8]Human Computer Interaction (HCI) Games Group, Stratford School of Interaction Design and Business, Faculty of Arts, University of Waterloo, Waterloo, ON, Canada

[9]OH! Lab, Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, United States

[10]Équipe LOKI, Inria Lille-Nord Europe, University of Lille, Lille, Villeneuve d'Ascq, France

[†]deceased

Corresponding Author:
Rina R Wehbe, BSc, MSc, PhD
Human Computer Interaction for Social Good (HCI4GOOD)
Faculty of Computer Science
Dalhousie University
6299 South Street
Halifax, NS, B3H4R2
Canada
Phone: 1 902 494 2211
Email: rina.wehbe@dal.ca

Abstract

Background: Although in many contexts unsuccessful games targeting learning, social interaction, or behavioral change have few downsides, when covering a sensitive domain such as mental health (MH), care must be taken to avoid harm and stigmatization of people who live with MH conditions. As a result, evaluation of the game to identify benefits and risks is crucial in understanding the game's success; however, assessment of these apps is often compared with the *nongame* control condition, resulting in findings specifically regarding entertainment value and user preferences. Research exploring the design process, integrating field experts, and guidelines for designing a successful serious game for sensitive topics is limited.

Objective: The aim of this study is to understand which elements of game design can guide a designer when designing a game for sensitive topics.

Methods: To carefully probe the design space of serious games for MH, we present Above Water (AbW), a game targeting the reduction of stigma surrounding MH, now in its second iteration. The game, AbW, serves as a consistent research probe to solicit expert feedback. Experts were recruited from a range of topic domains related to MH and wellness, game design, and user experience.

Results: By using this deployment as a research probe, this study demonstrates how to synthesize gained insights from multiple expert perspectives and create actionable guidelines for successful design of serious games targeting sensitive topics.

Conclusions: Our work contributes to a better understanding of how to design specialized games to address sensitive topics. We present a set of guidelines for designing games for sensitive subjects, and for each guideline, we present an example of how to apply the finding to the sample game (AbW). Furthermore, we demonstrate the generalizability to other sensitive topics by providing an additional example of a game that could be designed with the presented guidelines.

(JMIR Serious Games 2022;10(2):e21376) doi:10.2196/21376

KEYWORDS

human–computer interaction; games for change; games for mental health; sensitive topics; game design; empirical analysis; expert participants

Introduction

Background

Anyone discussing topics such as mental health (MH), poverty, socioeconomic status, homelessness, race, sexuality, identity as a person of color, and sexual health will likely agree that calling the conversation *difficult* is an understatement. Often, these topics may be considered extremely personal or *taboo* in our society; in other words, these are *sensitive* topics. Research on sensitive conversations explores the approaches, techniques, and outcomes of these situations and emphasizes the importance of having these discussions [1-3].

Above Water (AbW) is used as a probe to effectively elicit insights from experts. We then demonstrate how to apply these insights to the game, AbW, to exemplify how to design games that target sensitive topics. Contributions of this work demonstrates how to synthesize insights from multiple perspectives and create actionable items. We probed experts with different perspectives using a constant probe stimulus (AbW), which allowed us to capture the different perspectives from a range of experts and explore how these lessons may be generalized to other games targeting sensitive topics.

In this paper, we focus on the sensitive topic of MH. In our current sociocultural context, MH is often a difficult conversation. Researchers have demonstrated that parents need to have conversations early, to attempt to mitigate negative outcomes [4]. Untreated MH conditions are damaging to the individual and have cumulative effects on families, communities, business, and society. Treatment and support have significant tangible and intangible benefit to society. Considering only economic factors in the United States, the National Alliance on Mental Health reports US $193.2 billion in lost earnings per year owing to poor or unavailable treatment options. Moreover, this total does not speak to the additional nonmonetary benefits of increased mental wellness across society [5]. Although difficult to discuss, MH is an undermined aspect of overall health that is often overlooked in our culture, society, and sometimes even in the health care system itself.

Treatment for individuals is available; however, an individual may hesitate to seek treatment because of the stigma or the untrue, falsely held beliefs regarding MH [6,7]. It can be argued that the stigma and the associated feelings of shame, isolation, and misunderstandings perpetuated our culture and media. It can be argued that stigma causes more harm than the condition itself [8]. Stigma results in an avoidance of treatment and education about MH issues, thereby limiting an individual's willingness to get help for themselves [9] or offer effective help to others [10].

Research has shown that it is possible to reduce stigma through the introduction of resources, new information, or meeting with individuals [11]. In other words, increasing MH literacy (MHL) is inversely related to stigma. Increased MHL or knowledge of MH also improves one's ability to offer help [10]. In this paper, we explore how to increase MHL and reduce stigma in the context of a hybrid card and digital game. We leverage games and playful design to facilitate education and conversation.

The study by Juul [12] explains that games allow for the formation of a *magic circle*, a safe, playful space that encourages exploration. Within the magic circle, players can safely make mistakes and act in ways that are not in accordance with their everyday persona, thereby allowing the game to become an ideal environment for learning. Research on games and games for change demonstrates that learning through games can change behavior [13]. However, owing to the sensitive nature of MH and surrounding stigma, designing games for MH presents a particular challenge not seen with other health topics (eg, exergames) [14,15].

To understand what game mechanics and experiences would be most helpful, we invited experts in both MH and game design (GD) to participate in semistructured interviews aimed at eliciting information through the dismantling of a game designed to battle MH stigma. In this study, we use the game, AbW, as a research probe [16,17].

Through the dissection of the game, AbW, we see emergent themes of comfort, learning system, and technical design that led us to the presentation and discussion of successful mechanics for the design of sensitive topics. In this paper, we present design guidelines and demonstrate how to implement our guidelines using AbW as a case study. To further communicate the application of our guidelines and the applicability of these guidelines beyond MH, we also include a second example of a game that could be designed to tackle another sensitive topic.

Literature Review

MH is an important part of overall health. According to Mental Health America 2018 [18], 18% of adults or 43.3 million Americans are living with diagnosable MH concerns. Similarly, the Mental Health Commission of Canada has identified that 20% of Canadians will also experience an MH concern in their lifetime [19]. MH concerns are not limited to adults; Mental Health America 2018 reported that 11.93% of youth (aged 12-17 years) have had at least one episode of major depression [18]. These statistics are alarming given that untreated MH concerns can lead to further health and social challenges, including substance abuse disorder, loss of employment, reduced social functioning, thoughts of self-harm, and death by suicide. Furthermore, untreated MH concerns can also lead to community-level challenges, potentially resulting in greater rates of incarceration and increased strain on national health care systems [18]. In addition, companies may experience a loss of productivity or high turnover rate of employees owing to untreated MH concerns, resulting in higher personnel recruitment and training costs.

The current understanding of mental illness reflects the complex biological, psychological, and sociological factors that impact MH, in what is termed as the biopsychosocial model. Biological explanations include neurochemical imbalances, impairments of neural networks, and inherited genetic predispositions [20]. In addition, environmental factors interact with genetic aspects that ultimately result in complex networks of factors potentially predisposing an individual to mental illness [20].

A broad range of environmental factors are understood to potentially contribute to MH challenges. Some factors can be substance-mediated, such as substance abuse, whereas others are rooted in one's social environment and include adverse childhood events, sexual abuse, or trauma [21]. Literature on MH clearly shows that that both factors, internal and external to an individual, contribute to their MH [20,22].

Individuals living with an MH illness are less likely to receive treatment than those living with a physical health illness. It is estimated that only 20% of individuals with an MH concern or diagnosed condition sought medical treatment, and that only one-fifth of Canadian children who need MH care receive it [19,23]. In addition, approximately 60% of Canadians do not receive timely diagnosis and treatment (ie, <1 year from symptom onset) [24]. MH conditions represent serious challenges to those who live with them and our society at large.

The median time of diagnosis from the onset of symptoms is 4.4 years for depression and 6.2 years for anxiety disorders [24]. The time gap is even higher for illnesses such as schizophrenia and bipolar disorder [24].

Considering that MH illnesses are treatable and manageable, the previously mentioned statistics are disheartening. The effects of stigma or untrue beliefs perpetuated about a particular condition provide a possible explanation for this phenomenon.

Effects of Stigma on MH Issues

Stigma around MH has been shown to result in individuals not seeking support or treatment, even when they are readily available [6,7,9,25]. For example, youth may struggle to seek help from a parent, teacher, or caregiver concerning thoughts of self-harm and suicide. After seeking help, stigma may cause individuals to feel that their own problems with anxiety are very embarrassing to share with a support group, thereby continuing the perpetuation of MH stigma. Therefore, reducing stigma can lead people to seek treatment or get help from their community.

Stigma is perpetuated in society through multiple methods, not limited to media, including television and video games [8]. The root causes of MH stigma are complex and actively perpetuated in western media [8]. Even promoting an understanding of the causes of mental illness can lead to stigma and have adverse, unexpected effects as people may view MH illness as something out of the control of people with MH illnesses, promoting the stereotype of unpredictability [26].

Although a better understanding of the biology alleviates negative feelings of blame for those experiencing the mental illness, resulting feelings of pessimism for the prognosis have also been noted [26]. Improving MHL is a key factor for reducing stigma; however, it must be done comprehensively to ensure efforts to reduce stigma are successful [23]. Studies have shown that this is possible through the introduction of resources, new information, or meeting with individuals [11]. Therefore, a discussion group with new information and connecting individuals with different lived experiences with mental illness can potentially reduce stigma. Sensitive topics, such as MH, require a safe environment for discussion. Consequently, the ability to access information and ask questions while preserving privacy of individuals is imperative for success.

If intervention is delayed, for example owing to stigma, one disorder is likely to progress to multiple comorbid disorders that are more difficult to treat, with higher chances of recurrence [27]. In particular, anxiety disorders tend to have a longer delay, from years to decades, owing to early onset and lower perceived need for treatment [27]. The solution is changing peoples' beliefs about MH. Therefore, we need to explore platforms that provide a safe place to allow for learning, increase motivation, and change behavior.

Platforms for Battling MH Stigma

Social media has the potential to be used as an MH stigma–fighting platform. The most notable form of this strategy has been the sharing of narratives, especially testimonials from those experiencing mental illnesses [28]. Johnson et al [28] showed that presenting content in the form of a narrative combined with homophily in social media is more effective in education about MH than simply stating facts and statistics. Social media interactions foster a peer-to-peer support system that can be helpful in allowing individuals to share similar experiences and can even be beneficial for users who merely view the content anonymously [29]. However, information regarding MH illnesses tends to contain more inaccuracies and stereotypes than those regarding physical illnesses [30].

Direct MHL education campaigns have the potential to provide accurate information and have been shown to improve attitudes toward MH [30]. However, the effects were not lasting and did not improve the confidence of participants in helping and supporting those experiencing MH issues [30]. Moreover, the emphasis of the biological causes of mental illnesses seemed to encourage a helpless attitude and fear of unpredictability and dangerousness [30].

Campaigns and programs that have been successful rely on multiple platforms and approaches. For instance, Time to Change, an extensive outreach program to improve public attitude toward MH issues and reduce stigma improved the attitudes of 5.4 million people in the United Kingdom since 2008 [31,32]. Their approach involved national social marketing and grassroots-level programs in communities involving social contact with those affected by MH-related stigma [31]. Beyondblue is another example of an extensive national initiative founded in Australia with the aim to improve MHL [31]. One of their notable approaches was providing a platform for well-known actors to share their experiences with MH illnesses [31].

Methods of fighting stigma involving both contact and education may be the most successful [30]. Serious games involving discussion of MH topics combine the benefits of both forms of approaching MH stigma. Games can provide players with a safe environment [12] and change their attitudes and behaviors [13]. They can also provide players with a sense of belonging or community, which may allow for the discussion of more sensitive topics [33,34].

The Use of Games Related to MH

Playful apps, including games, gamified apps, and simulations, have been used adjunct to therapy, often by gamifying a well-known method of clinical treatment such as cognitive behavioral therapy. The game SuperBetter [35] uses a point system and principles of cognitive behavioral therapy to reduce depressive symptoms [35]. Similarly, virtual reality (VR) has gained some traction for exposure therapy of phobias. Although playful, VR apps are not necessarily considered to be games owing to the missing cooperative or competitive aspect [36]. The application of games in therapy has gained support because games are known to be a vehicle for motivating behavior owing to the complex and responsive rewards systems built into their design [37-39].

Learning complex ideas is part of many game systems. For example, the ability to catch a Pokémon in a game is an abstraction of a larger probability equation. Therefore, games have often been leveraged for learning [40]. There is an abundance of support for educational games, serious games, and gamification [36,41,42]; these games and apps aim to computerize aspects of existing evidence-based therapies. More research needs to be done on their contribution to the effectiveness of therapy; however, they can be potentially useful for learning about MH.

Another category of games targeting the lay person or general public aim to improve MHL. For example, Stigma-Stop is a nonimmersive VR game in which each participant plays the role of a person who is facing adversity from a mental illness [43]. Information about the illness and symptoms is presented to the player, who receives feedback based on their decisions [43]. The game succeeded in helping players gain a better understanding of each of the conditions and is demonstrated to effectively reduce stigma [43].

Researchers have shown that when players enter the game world, they suspend their disbelief to become more open to concepts [12]. As a result, researchers have demonstrated that games can provide a chance for reevaluation of ideas [44]. Therefore, games can provide an opportunity to reevaluate beliefs held as inherently true regarding MH. Stigma-Stop [43] was found to be successful in debunking common misconceptions about certain MH illnesses; however, its long-term effects are yet to be evaluated to confirm a long-term reduction in stigma [43].

Cangas et al [43] suggest that contact with those experiencing a mental illness may be the missing element. Researchers tested a serious game to provide training for responding to MH concerns in the workplace [45]. Hanisch et al [45] reported that management staff who played the game had improved knowledge about MH. In both games tested by researchers, accurate information was provided, and the players could interact with other characters experiencing a mental illness. Moreover, testimonials of real people experiencing mental illness were provided as a supplemental video. Therefore, both the education and contact strategy were used to reduce stigma and improve knowledge of MH issues.

In addition, in the context of games, the actions of players including failures are relatively inconsequential. Unlike other commonly encountered scenarios, it is acceptable to fail or not know the correct answer in a game [46,47].

Overall, games and playful apps that provide training were proven effective [35,43,45,48]; however, these apps are not always accessible to the public. Moreover, it is possible that these games are not appealing to the public owing to the formality surrounding sensitive topics, which again contributes to perpetuating stigma by discouraging comfortable discussion.

The Potential of Serious Discussion Games in Mitigating MH Stigma

Although games have been used alongside therapy or as a part of formal training to improve MHL, there is an unmet need for games that facilitate discussion of this serious topic aimed at the general public. Serious games for MH are continuing to expand in the field as more game designers are beginning to explore the use of gaming strategies that can increase the positive outcomes for MH issues [49]. However, most of these games are either not publicly available (ie, only available to registered health professionals [HPs]) or not clinically tested [49]. In response, this study aims to explore GD that can provide a safe environment for thoughtful discussion of MH that has the potential to reduce stigma.

More traditional board games are also being used to facilitate these discussions. For instance, the Learning Life Game [50] is a simple, noncompetitive board game, which randomizes questions for discussion. The goal of the game is to help players learn about themselves and others. The game hopes to teach problem-solving strategies by leveraging role-playing game (RPG) strategies, imagined situations, and hypothetical questions [50]. The purpose of this game is to help adolescent therapy groups better convey their feelings and communicate. Answering the questions in a group was found to elicit cooperative behavior from other group members, particularly when challenging questions or vocabulary were encountered [50]. Overall, this game was a helpful tool in helping adolescents develop verbal skills and provided structure for group therapy sessions [50]. However, as mentioned before, games designed for a formal therapy setting are not necessarily suitable or appealing to the public for whom the reduction of stigma is a more immediate issue.

If serious topics are to be discussed in an unsupervised and informal setting, game mechanics must work to provide a safe space to improve the quality of the discussion. Privacy is a crucial element that can be afforded through the mechanics of gameplay such as anonymity, randomization, or hidden information. For example, anonymity is a contributing factor to the execution of cheating behaviors; therefore, GD can choose to include anonymity as a mechanism and leverage it to direct the atmosphere of the game space [51].

GD has shown mixed results for anonymity as a mechanism to shape the play space. Cheating behaviors and competitive actions can be encouraged through the use of deception or anonymity [51]. In contrast, the permissive environment encouraged by anonymity can be used positively. For example, AbW leverages anonymity positively to foster communication about MH and to reduce fears of experiencing stigma while playing the game. Similarly, other studies have also indicated that anonymity can be used to create more welcoming environments. A study [52] on a web-based focus group, which discussed dating and intimacy for survivors of childhood cancer, highlights key design elements that can aid the discussion of serious or sensitive topics. It was found that participants preferred the anonymity that the chat provided. They also preferred groups with >2 people and a moderator to stimulate the conversation [52]. AbW is a novel game that combines these design features to construct the ideal environment for the discussion of stigma surrounding MH.

Finally, games can be used as vehicles for storytelling, which has been found to increase empathy [53]. Playing games cooperatively can also increase feelings of empathy [54,55] and reduce feelings of antifriendship or pleasure from the misery of others [56]. Emotions can be leveraged in the design of educational games [57], which AbW plans to do through the emotional journey that the players experience. Although simulated, real-emotion experiences can be felt through games and stories [21,53,58].

GD can use the created empathy to design the mood and atmosphere in their games. For example, consider *Brothers: A Tale of Two Sons* (Starbreeze Studios) [16], which uses the relationship between the 2 main characters to perpetuate the story and elicit empathy from its players. Following this strategy, AbW attempts to use personal connection and empathy as part of the GD. AbW attempts to bring together members of a group to share stories and experiences that would otherwise not be discussed in everyday conversations, in the hope of effectively reducing stigma surrounding MH challenges.

Design

The game, AbW [17,59], was designed to help educate individuals about anxiety, specifically generalized anxiety disorder (GAD) and panic disorder (PD). Upon further iteration and in-house testing, the concept of the game was widened to include general information about MH. The current version of AbW intends to educate individuals on general MH factors within a wider breadth and focuses on anxiety in depth.

In summary, AbW uses both physical cards and an integrated digital component in the form of a website, which is intended to be used similar to a phone app. The game was designed to be an educational tool and not a treatment in and of itself. Potential use cases include support groups, caregiver–child dyadic pairs, and MH training.

AbW is designed to be an experience in which participants experience a range of emotions; for example, anticipation of getting the right card, relief for being able to ask desired questions, joy when discussing victories related in their own or their bespoke character's journeys, and silliness when playing a minigame. Although the game is not designed to be inherently stressful, participants may converse in depth about situations that are stressful and choose to disclose personal information. The variance of emotional experience will depend on the group, setting, and individuals' decisions to disclose information. Owing to the variability in conversation, the game has inherent replayability, because it is nearly impossible to replicate a human interaction as no 2 conversations are the same.

Design History

The design of AbW has been detailed in previous work [16,60]. AbW began as a seminar project focused on exploring the designing of games for health. The GD was presented at 2 international conferences in the form of a student design competition submission [59] and an interactive demonstration [16].

Upon moving to evaluate the efficacy of the game as an educational game and an intervention for MH stigma, there were concerns about the safety of presenting information about MH to naïve participants owing to the sensitive nature of MH conversations. Researchers worked closely with the office of research ethics to determine the safest approach to evaluate the game. To ensure the safety of players, it was decided that an evaluation with expert participants should precede the evaluation of the game with a large group of naïve player-participants.

Digital and Physical Card Game

The game is played with a deck of cards and a personal computing device (ie, mobile phone or tablet; Figures 1 and 2 [17,59]). Most gameplay is conducted through physical game cards, and the game is moderated by the app.

The goal of the game is to collect the player's chosen *life goal cards*. Among other functionalities, the digital component serves to track the *life goal* the players have selected.

Added challenge stems from the appearance of *anxiety cards*, randomly drawn from the deck. Management of *anxiety cards* is made possible with *treatment cards*, which convey information and remove the negatively-scored *anxiety* cards the player may accumulate as the group traverses the card deck (Figure 3 [17,59]). In-game effectiveness of *treatment cards* is rated based on the amount of help that each treatment requires from an MH professional. The more involved an MH professional is in the execution of the treatment, the more anxiety cards the treatment card can counteract. For example, a healthy eating card would only counteract 1 anxiety card, whereas the card representing a psychiatrist's prescription would counteract significantly more.

The game is played open-face, with accumulated *life goal* and *anxiety* cards displayed openly (face-up) on the table (Figure 4 [17,59]). The collected cards in front of each player denote the player's current progress.

Figure 1. An image of the app, which comprises part of the gameplay space [17,59].

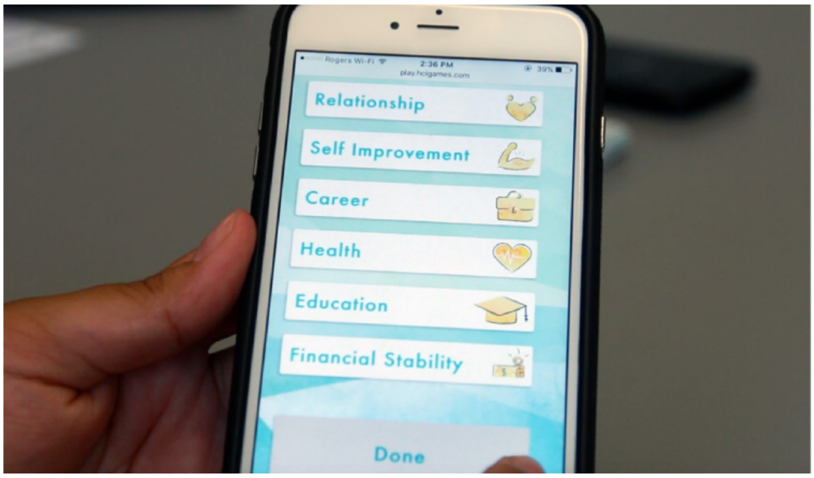
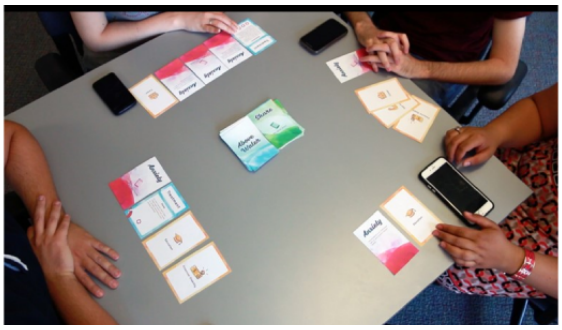


Figure 2. The card types in the game [17,59].

Figure 3. The treatment cards come together to symbolize the complexity of treatment plans, which usually comprise multiple interventions or habit-changing efforts [17,59].

Figure 4. A picture of the game in play using the bring-your-own-device policy. Note that the game is played open-handed [17,59].



Finally, the last category of cards is the *share card* which asks users to share experiences, tips, or ask questions. The game facilitates dialogue though *share cards*, which are meant to stimulate discussion, self-reflection, and discourse with other players. The *share card* activity always points to the digital app.

The app serves to augment the card game. During the gameplay, players may draw a card with a phone-icon; this game event instructs players to use the mobile app. Treatment cards, share cards, and anxiety cards point to minigames in the digital component such as myth-busting, deep breathing, and yoga exercises.

Activities associated with these cards use the players' phones for facilitation. These include engaging players in minigames to allow them to try some of the basic at-home methods of managing anxiety, such as guiding players through yoga poses and deep breathing exercises. The share cards ask players to type a response to a prompt in their phone, which are randomized and shared anonymously with other players.

By augmenting a traditional in-person card-based game with phone-based features, AbW seeks to provide players a way to fully engage in the game with other players to express feelings and talk comfortably about sensitive topics, while maintaining anonymity and privacy.

Designing for Space and Privacy

The game features a distributed platform design, which spans the mobile device and a tabletop card game. We took a novel approach to the tabletop game paradigm by orienting and guiding the card game using a mobile app. We follow a bring-your-own-device (BYOD) protocol (Figure 4 [16,59]). The BYOD protocol allows players to have control over their own privacy. Participants can choose to access the web portal through the browser of their choice with a device calibrated to their preferred settings; for example, the participant can choose to use a privacy-aware browser or their browser's *private browsing* mode.

The game is designed to foster discussion and protect anonymity of the users, while allowing for a shared group learning experience. The game leverages anonymity [51] through augmentation of the card game through a mobile app. The app allows users to input answers, and it randomly selects 1 answer to be presented for group discussion. The app balances anonymity through face-to-face group discussion. Both these design decisions, anonymity and in-person interaction, have demonstrable effects on the game's atmosphere and were selected to reduce stigma associated with diagnoses of MH illnesses or MH concerns [11,61]. For a full description of the game mechanics, please refer to previous work [16,17].

Accessing Trusted Resources

The GD is the contribution of the project, the content itself is not. The content is sourced from credible trusted resources such as medical organizations, government organizations, and outreach programs. Information presented in the game is designed to spark curiosity and encourage players to read about MH and pursue their own active education after the game ends. With the BYOD paradigm, players leave the game with the links to resources on their phone and can access the website again anytime.

Methods

Overview

In this study, we asked MH professionals and game designers to provide feedback on the game, AbW, through semistructured interviews.

The expert evaluation of our game allows the researchers to ensure that our intervention, AbW, is safe and has a likelihood of being effective before proceeding to conduct a playtest with a group of nonexpert participants.

AbW was used as a probe to effectively elicit insights from experts. We then demonstrate how to apply these insights to the game, AbW, to exemplify how to design games that target sensitive topics. Contributions of this work demonstrate how to synthesize insights from multiple perspectives and create actionable items. We probed experts with different perspectives using a constant probe stimulus (AbW), which allowed us to capture the different perspectives from a range of experts and explore how these lessons may be generalized to other games targeting sensitive topics.

Participants

In total, 14 participants were interviewed for the study, 7 (50%) HPs and 7 (50%) game designers, between April 2018 and November 2019. Demographic information of the participants is shown in Table 1. We sought to speak with individuals with a variety of roles in MH care provision and in the games industry. Within the scope of MH professionals, we included social workers, counselors, and community leaders, who contribute to the biopsychosocial definition of health in the community; for example, MH3 is a hospital chaplain who services patients, families, and caregivers to support coping and provide counseling. In addition, expert participants from game, user experience (UX), and related disciplines were recruited internationally (North America and Australia).

Expert participants in MH and wellness were recruited by contacting health institutions and practices in the local community. Further calls for participation were sent to HPs nationally via professional network organizations, networking, message boards, and listservs. Recommendations from participants were also followed.

Table 1. Participants' occupation.

Demographic type and ID	Demographic information
Mental health professionals	
MH1	Clinical and counseling psychologist
MH2	Registered practical nurse
MH3	Chaplain (hospital)
MH4	School guidance counselor
MH5	Psychology professor
MH6	Clinical psychology lecturer
MH7	Occupational therapist
Game designers	
GD1	Game designer and experience designer
GD2	Game designer
GD3	Game designer
GD4	Experience designer
GD5	Game producer
GD6	Game designer and programmer
GD7	Senior user experience professional

Interview Protocol

Interviews were made accessible by allowing participants to choose how and where the interview was conducted, as web-based or in person, and at a time of their choosing. Consent was obtained before participating.

Participants were sent a digital copy of the game and a link to the game website to review before the interview. During in-person interviews, a physical copy of the card deck was also presented. Participants were informed that the game was still in the prototype phase, so that participants would feel comfortable to be critical of the game.

Interview questions focused on the benefits and risks of playing the game socially, in a clinical or professional setting; the feasibility of the game mechanics, especially the digital components; and the ability of the game to achieve its goal of reducing stigma. Expert participants were also asked to convey their opinions on the risks and benefits of AbW and games for change as interventions.

The interviews were semistructured. Prompts were included for the research team to pursue based on their own judgment. A copy of the provided prompts is included in Multimedia Appendix 1. Typical to semistructured interviews, information arising from past interviews provided direction for continued discussion. The interview was designed to take approximately 1 hour; however, in clinical and business settings, this time was often shortened. When interviews were shortened, we concentrated on larger themes (Multimedia Appendix 2).

Interview Analysis

The interview protocol itself was semistructured, with prompts for continued discussion. Interviews were audio recorded and then transcribed. Interviews were coded using NVivo (version 12.2.0; QSR International) for macOS using thematic coding in a grounded theory methodology [62].

As data were collected, the researchers began to build theories and look for conversion within the emerging results, beginning with a line-by-line assessment for our initial 2 participants, which grounded the further thematic analysis with subsequent participants. Coding conducted on a line-by-line basis abstracted key themes from participant responses. Relevant codes were then organized into a mind map, allowing codes to be organized into emergent themes for inductive analysis. A positivist approach was taken during this analysis, as the data analyzer believed that evaluating participants' thoughts about the game in an objective manner would result in conclusions about the game that would increase the game's utility to stakeholders, including clinicians [63]. That is, viewing participant responses as impartially discussing the game allowed more generalizations to be made about it.

Upon completion of the data collection phase and transcription, thematic analysis based on the initial grounded theory results was used to understand larger themes. Three members of the research team (RRW, CW, and YE) reviewed the transcripts independently and decided the potential codes individually. Then, they merged their codebooks through discussion and reached consensus to form a preliminary codebook. These preliminary codes were then organized into 12 major codes inductively.

Then, to remove possible biases, a full code analysis was performed by YE and naïve coder AS, who coded all the interviews using NVivo (version 12; QSR International) for macOS, paragraph by paragraph using the major codes. The overall unweighted agreement among coders was κ of 0.67, indicating moderate agreement among coders.

Following the completion of this iteration of coding, RRW and CW rejoined the discussion, and then, all the 4 researchers discussed the results of the coding while providing their respective takeaways for each of the 12 major codes. Following the conclusion of this protocol, all the authors were invited to review the findings. The complete code reference is shown in Table 2.

Table 2. Qualitative coding reference table.

Theme and major code	Description of major code	Clinical or medical	GD[a] or UX[b]	Minor codes
Comfort				
Security and privacy	Owing to sensitivity of the data, precautions on how it is stored and accessed and how participants feel about sharing their information (Do they feel safe and trusting of the system and other players?)	58	54	Anonymity, privacy, risks, safety and vulnerability, sensitive information, checkpoints, and sharing
Social dimensions	The social environment surrounding the game, including how participants feel (awkwardness and tension), talk to each other, room atmosphere, and the resulting consequences (ie, quality of conversation, fear of being incorrect, or challenging or correcting others).	72	49	Social environment, wrongness, conversation facilitation, and awkwardness
Facilitator	Facilitator role or game leader role as it is needed or as the game is designed.	21	40	Authority, facilitator, and facilitator skill
Clinical value or outcomes	The outcome of the game as it applies to players leaving the game (What did they learn? Did it affect their thoughts or behaviors?)	83	92	Game outcome and clinical effects
People				
Community	The ability of the game to create a connected environment within a community. In this category, we refer to the community as a group with a clinical facilitator.	41	18	Clinical and community value of game
Target	Who is the target population (age and familiarity with content) and what are their relation to each other? How does that affect the design of the game?	66	40	Population and age
Replay or customization	Features adding to replay value or customization of the game.	11	24	Customization for player, clinician, or facilitator and replay value of game
Content	The content of the game, specifically the density, depth, and breadth of the information.	90	87	Content of resources, game content, data dense, and data sparse
Learning system	How learning is approached, handled, and reinforced. In this category we are specific to the learner (not the facilitator role).	60	42	Learning, scaffolding, pedagogy, question structure, presentation of content, and reinforcing
Design				
GD	How the game is designed and the game's attitude (ie, fun or cooperative).	103	163	Pacing, simplicity, fictional person and narrative, cooperative vs competitive goals, impact chain, strategy, analogy, and random
Technical design	The design of the app itself (void of all learning and game elements). Here we refer to the technical elements of the interface (navigation, buttons, etc).	33	43	Mechanic of typing, keyboard accessibility, and input methods; autonomy; navigation of app; coordination; syncing; and hard to reference
Onboarding	Onboarding procedures and clarity of instructions for both the facilitator and the learner.	22	28	Onboarding and instruction clarity

[a]GD: game design.

[b]UX: user experience.

Ethics

This study received ethics approval from the University of Waterloo, Canada, registered under office of research ethics (number: 32195). The data set for this study is not available for access.

Results

Overview

The iterative process that resulted in the final codebook can be represented in 3 stages. In the first stage, the resultant codes of individual exploration of the data were compiled. After additional reflection and exploration of the data set, the authors proceeded to combine the codes into larger themes. Then, the themes resulting from this stage were used to code the data set. Next, line-by-line coding of the interviews with both medical and UX experts and the findings from each domain were discussed in detail. Finally, an overall look at the data set revealed the presented themes. In this stage, the authors discussed the resultant categories and developed 3 main theories. Each theory is described in detail in this section. Quotes are presented for each theme to illustrate the findings from the data.

Comfort

Overview

Owing to the sensitive nature of the discussion topic, the research team worked closely with the office of research ethics to ensure that the content presented in the game was not triggering. Here, we use *triggering* to describe the negative effects that can occur with the presentation of a stimulus to a venerable party. As a result of discussions with the ethics board, it emerged that we would need to probe our expert participants for their assessment of safety and possible risks of AbW.

The first emergent theme from the data points to the importance of participant comfort. Upon analysis of the results, we see the emergent theme of participant comfort, which goes beyond our initial research question pertaining solely to the presentation of triggering information. As it became clear that our probe was very narrow, we began to collect data on factors that contribute to a concussive environment for discussion. The results indicated that comfort goes beyond the content provided on the cards. From the data, we see a clear need for an overall consideration of security; privacy of data; social environment, clinical and community values; and a facilitator.

Am I Comfortable With the System?

Does the player trust the game? We know from the literature that privacy and security on the web is a challenging topic for the nonspecialized, lay user to understand [64]. Therefore, part of the work in designing a game that connects to the internet is to ensure that users feel that data entered into the app will not be used maliciously. Designed with this in mind, the game does not store information or participant data. In addition, the game was designed to obtain user input and randomly select an answer without identifying the user themselves:

> *The anonymous component of it is very helpful, because then a group of people can get through a game session and not necessarily put them in a position where they're - they're vulnerable about something that might be difficult to discuss, but still they're able to get the information or get the - educational resources that you're trying to convey.* [MH2]

Am I Comfortable With the Game Content?

We want to ensure that the content of the game is comfortable to interact with. This includes the content being worded and presented in a way that demonstrates sensitivity. For instance, the game should not betray the user's trust by suddenly presenting triggering content. Instead, it should allow users to ease into the material:

> *[Life Goal Cards] could actually cause some people to be anxious, when they're picking it. And like I could imagine somebody sitting there going `you know, I really wanted to get that degree in computer science, 15 years ago. I never got that degree in computer science.* [GD5]

Am I Comfortable With This Group?

Creating a comfortable social game environment for discussion of sensitive topics is an important and challenging aspect of the game's design. Initially, we sought information particularly about triggers for individual experiences; however, expert feedback created a better understanding of how the design of the game should be inoculated against situations arising from the social dimensions of a group.

The goal is clear and well-articulated by GD1 in the following quote:

> *If you can get the group to take it seriously and offer serious conversations, then you can open up things that a group of friends even or family would never talk about normally.* [GD1]

A challenge could be that people sometimes react in a way that belies our true feelings (eg, laughing because you feel awkward). Moreover, it is possible that a player may react strongly to a general statement (eg, assuming that there is personal commentary, when none exists):

> *Some people talking about, or sharing, or disclosing—even if it's in an anonymous format—can still have some strong emotional reactions, either to their stuff being selected, or to discovering that that's not the social norm, or to putting themselves out there and not getting the positive feedback that they would hope.* [MH3]

Moreover, if participants are playing the game in a clinical or formal setting (eg, support group, group therapy, or workshop), where there may be a facilitator or exercise leader, the idea of them not being involved as a player in the game may also be intimidating:

> *I'd like to have them actually play the game with the participants. That would always be my inclination, rather than kind of standing off to the side observing. Because that obviously leaves a kind of a cold clinical feel to it, like everybody's - well if they weren't anxious before, they are now because the psychotherapist is watching. If there was an expert, I'd rather have them play.* [MH1]

Am I Comfortable in This Setting?

After considering the social environment that the game has built for each individual, we moved on to understanding the individual as a group, creating an environment of group comfort.

As MH2 explains, it may be a matter of starting the discussion and allowing it to flourish, suggesting that the game needs only to *spark* the discussion:

> *It - as far as like stigma, yeah because people get to talk about how they deal with it together. Raise awareness, to have it, normalize it into a game that...yeah. It's an educational tool. It's - it - and that, it's very helpful. It - it makes it so people are talking about it more instead of not dealing with it, so I mean it's very helpful - it looks like a very helpful tool to use.* [MH2]

Does the Game Need a Facilitator?

A large portion of the discussion arising from the results was the question of whether the game needs a dedicated facilitator.

A facilitator in GD is not an unusual principle. For example, the classic Dungeons and Dragons role-playing tabletop game, has a *dungeon master* role in which one of the players facilitates the gameplay [65]. Commercial games available on the market, such as the Ultimate Werewolf (Ted Alspach, Bézier Games) follows a similar format.

Expert participants provided careful feedback regarding this concept. Points of consideration for the role of the facilitator were related to the maintenance of the comfortable environment:

> *The only danger might be in - it's common to any group, where somebody is really insensitive in responding to someone else's sharings, and they're like `I can't believe whoever wrote this was worried about going to the grocery store, like that's the dumbest thing I ever heard'. That I do wonder - that's where having a trained group moderator might be helpful.* [MH1]

In addition to a question of whether the game should be designed with a facilitator role, there was a question about who can fill the facilitator role. Particularly, expert participants cautioned that the situation in which the game is played may require a knowledgeable individual (eg, a clinician or a teacher) to lead the group:

> *I would also suggest that if this game was being used as anything other than background information for like a health class, then probably it should probably be targeted for use with someone who has competency in delivering it.* [MH6]

The feedback from the participants about having a facilitator feed into a larger conversation about the overall learning system design. Data demonstrate that there is reasonable concern that a nonfacilitated session depends on the player's ability to read and interpret information; especially in groups, it is possible that the *loudest voice* may be accepted as correct or misinformation. The game is a learning environment, as such participants should explore the concept and think through the new information; however, participants may not always be trusted to completely read the information available or interpret it correctly.

Moreover, the game presents ambiguous situations with no completely correct answer. For example, the cards may ask about how one would deal with stigma from a family member or in the workplace. In the following quote, GD1 discusses the mechanics of the share card. Each participant provides an answer to the situation, and the game chooses an answer and presents it anonymously:

> *Sometimes people give advice that comes from a perspective of not really understanding the issues. So I'd be curious if that comes up or how - how the game deals with situations like that. Because there's no authority of voice here so the game - the game is gonna just pick one at random, right?* [GD1]

The concern with a game, such as AbW, is that most of the learning occurs within the context of the group conversation and that is encouraged, but discussions cannot be controlled by a card game, even one augmented with an app.

Learning System

Overview

Given that this is an educational game, learning is at the forefront of the game's goals. We began simply by analyzing the data sets for information regarding the level of content. We then looked at the specificity of the game mechanics to the target group, and finally, we wanted to understand whether the game's learning system met the learning objectives and if it was replayable.

Content and Replayability

The overall goal of the game was to provide information about GAD and PD. Most participants indicated that the content's depth and breadth were sufficient, with few participants stating that more content should be added and fewer participants indicating that the content was very expansive.

To decide the level of content needed, expert participants pointed to how the game was targeted to a particular subset of individuals:

> *Having gone through PTSD, someone saying it's an anxiety disorder doesn't teach me anything or feel useful, it just feels a bit dismissive, which seems opposite to your goal.* [GD6]

The game designer points out in the quote above that there is a difference in how the material feels to someone who is just beginning to understand the problem as opposed to someone who has had lived experience. To a person beginning to learn about clinical anxiety, information about the many subcategories of the large disorder (eg, PD, posttraumatic stress disorder, and GAD) may be surprising.

Participants also mentioned the idea of a facilitator as a way to control the content and flow of the game. That is, if the facilitator is there to control the environment, can they not also control the consumption of information?

Overall, replayability was viewed differently for games designed for general play and nonsensitive topics. By our participants, we are advised that multiple sessions of repeated play can help in effective learning and messaging:

> *Single run is probably not sufficient or not as good as, but once a week is perhaps too frequent. I would that using the game as part of a process. So, if I meet with a student weekly, that we talk we play the game for some bit. Maybe next week we talk about something else. Maybe we'll explore one of the ideas in greater depth. And then maybe later on, it might be also be group, right. So, it might be individually, and then later it might be some group session and so-on.* [MH6]

The value of replaying the game for the participants may be understated. Revisiting or replaying the game may help with guidance over a period of growth:

> *And again with a moderator, a therapist, they might be able to help - actually help that person draw in the positive attributes of their treatment, and how they're feeling about the treatment they've received over a period of time. So whether they can - at that point they would be saying - probably sharing their own experiences, but if - even if they don't want to share their experiences they would share things that they've been told by their therapist, right? So they've kind of got to that point where they're just talking about it, and it might be even helping them destigmatize their own health issues.* [GD5]

It is possible that as game designers, we see the onus of making the game replayable on the game mechanics, and clinicians see the replayability of the game as an assistant to one's journey of growth.

Target Market, Clinical, and Community Value

From our data, we see that the game, AbW, sits at a branching point in design. Expert participants made us aware that the game could be designed in multiple and highly specific situations. In our results, we captured a large range of use cases. As participants were provided a copy of the game before the interview, many participants came to the interview with ideas for whom the game may be valuable. As our research questions included understanding possible target groups, we probed participants to understand the situations and use cases they had envisioned.

Suggestions included, but were not limited to, therapy groups, support groups, GD workshops for non-MH professionals, in-home family or parent–child dyad communication, and educational institutions such as schools. The diversity of responses was in juxtaposition to the only point of agreement—the game would not be a *pick-up-and-play* game at a friend's game night or at a game cafe:

> *As far as the mobile component, like if you can make - maybe if you made the whole game able to be done through an app, people can just meet up for the community - the community focus ones. Like there's - there's treatment centres and there's community focus therapies. So people go to - go to their social work and such and if you have those groups there, you can just pull up the app on their phone and they can all play together.* [MH2]

When dealing with sensitive topics, our data suggest that the customization of the game should be in accordance with the delicate nuances of the population. These nuances are only revealed when we have gained significant insight into a demographics of participants. These findings are consistent with literature, which indicates that personalization of the learning system is more effective [66].

Technical and Practical Design

Finally, the last emergent theme that we identified was the technical and practical design. Although, generally, all apps will need some degree of design improvements, here we focus on improvements particular to designing a game on sensitive topics.

Onboarding and Instruction

There is a need for intuitive navigation, user interface design, and GD. Here, the focus is on being able to support navigation and research during the game with an emphasis on discretion. This includes being autonomous in leaving and entering the game and not being negatively impacted by accessing resources during gameplay.

Our results indicated that the navigation may be unintuitive for a mixed audience of people with different levels of computer literacy. In the following quote, our expert participant highlights the lack of flexibility in navigation owing to the currently implemented networking structure of the app:

> *I'm not as technologically sophisticated as the two of you, it takes me forever on my phone to type a message. So if I were to be playing, there would be all these twelve year olds who are done, and I would still be hunting around looking for capitals and punctuation. So that part would be a barrier for somebody like me.* [MH3]

The goal is to create a pick-up-and-play game to allow participants to concentrate on learning instead of strategy. Again, the contrast between game designers and HPs provided an interesting space for design exploration. Slower and simpler mechanics and pacing were mainly advised by HPs; this was in contrast to game designers, who felt the need to add more mechanisms and allow for divergent strategy and player choice. The divergence among groups of expert professionals converged on the discussions of pacing, emphasizing icebreakers before sharing cards. Allowing to develop a sense of comfort before encouraging more personal discussion was emphasized:

> *Like once you get started on telling a story - the first time that you tell a personal story should not be to a group of strangers, let's just - so that's the risk, right? And so what if the person gets into it and then they get overwhelmed and so they're getting more and more upset. They've said something that's really upsetting, possibly upsetting other people who have had similar experiences, but even if that doesn't*

> *happen, who's the person that can handle this and contain it?* [MH3]

Mediation of these concepts led to the discussion of onboarding or learning to play the game first. Instructions can remove the discomfort of ambiguity:

> *For the share functionality, the uncertainty here seems like it would be uncomfortable. There aren't clear instructions for who is going to share, or how that should take place. My friends with anxiety all like things to be clear and understandable - uncertainty is a key thing that makes them anxious.* [GD7]

Designing Options

The data revealed that the design of the app should allow more flexibility for players to traverse the content freely during gameplay with other players. The game is designed to link back to resources for further information after the game, but the design should be flexible enough to allow back and forth movement or in-line definitions.

To improve the original design, the game should leverage the moment of curiosity. Acting within the moment to deliver information may be more effective because participants have just uncovered the content and are actively interested:

> *It's one thing to have a page of resources but if you can't access them when you're not playing the game, then that's less useful. But if you - if you're kind of curious is you could have - you know email this to myself or download a PDF onto my phone, then it's there.* [MH3]

In the quoted content above, the expert participant makes an important point about note-taking. Participants may want to feel such that they have collected or saved notes during the game.

Discussion

Principal Findings

Our results provide us with an insight into the problems that are currently present in the game, AbW; however, the implications of our results are beyond the redesign of a single game. We use AbW as a research probe to understand how to design games for MH and sensitive topics. For each theme, we provide design guidelines for games about sensitive topics. We also provide an example of how to redesign the example game, AbW. Using AbW, we can demonstrate the application of our findings in a redesigning of the game.

In this section, we include a design idea for a possible game for another sensitive topic as an alternative example to demonstrate the applicability of these guidelines outside the game, AbW, showcasing the generalizability of our findings.

Comfort

By presenting this design guideline, we emphasize that the onus of responsibility to provide this comfortable environment is on the game designers and not the players. Similarly, the designers are also seen as responsible for providing the scaffolding for methods and procedures to maintain a comfortable environment.

Focus on the Atmosphere and the Meta-game

Design for comfortable environments that encourage discussion of sensitive topics is a critical challenge to the design of games for sensitive topics. Owing to the sensitive nature of the information, designers need to consider mechanisms for dealing with awkward interactions, icebreakers, and oversharing.

Designing the atmosphere of the game may include considering the design of the system and content of the game and allowing for variables including the group of people, the setting in which the game is being played, and the possibility of a game facilitator.

If we were designing a game that attempts to tackle the topic of racism, it is likely that conversations surrounding slavery, historical violence, and systematic bias in governing bodies would be discussed. For example, given the recent political climate following the brutality of police and murder of George Floyd Jr [67] captured on a cell phone, conversations related to the found footage can be triggering to individuals. Therefore, it is necessary to allow participants to decide the comfortable limitations to the game, possibly by allowing players to create the atmosphere of the game through emergent gameplay.

Consider we propose a design for a board game in which participants are given a large quantity of playable pieces of 8 possible colors. The goal is to fill the board to pursue two possible winning conditions: (1) a single player or alliance of players controls the board and floods the game with a single color, thereby dominating as the majority color until the other players can longer make legal moves or (2) all players win when negotiation leads to an equal ratio of representation of all colors on the board. As participants play token pieces on the board, we may decide to emphasize the injustices faced by minority groups by allowing players represented by the color in majority on the board and allow that player to impose arbitrary overpowered and unfair penalties. As games can become heated, it may be necessary to impose a *clear the board card* or *equal ratios card* to allow the possibility of the second winning condition representing a peaceful ending. Perhaps, a player who is dedicated to pursuing the first winning condition may be *voted out* of the game by a majority vote when the game board is closer to an equal ratio.

A game idea such as this allows players to choose the atmosphere of the game. If the group as a majority prefers collaboration, they may choose to pursue a peaceful ending together as a metaphor for a world with perfect equity.

Assign a Facilitator Role to Monitor the Game Environment and Keep a Positive Atmosphere (Optional)

A larger point of discussion within the research was the need for a facilitator. A facilitator is a group leader who would be a helpful factor in maintaining a comfortable environment because responsibility has been assigned.

Expert participants suggested any varieties of facilitation ideas from providing resources to accommodate a player-facilitator to having the game played within a supervised clinical practice. However, convergence on the central need for a facilitator was clear—maintenance and arbitration to maintain a comfortable working environment or the creation of an environment for participants, in which they may feel safe to express their feelings, knowledge, and lived experience. Recommendations from experts in this area suggest the need for a responsible party to act as a group leader—the facilitator.

In the case of AbW, the facilitator role can be added as a nonplayer. The facilitator role allows for the enforcement of a comfortable environment as dictated by the game designer. For example, one option is to limit the game to a clinical setting, played in the presence of a clinician. In this iteration, the information in the rule set on what the facilitator should know, what types of conversations or language to steer away from, or the stipulation that this game should be attended to by a clinician. The clinician, acting as the facilitator, can then use the game as a tool for directed discussion.

Let us propose another GD that may benefit from a facilitator. Perhaps we are hoping to teach players about the diversity that exists in the autism spectrum disorder (ASD). We might decide to extend the known metaphor of a jigsaw puzzle often used by ASD awareness groups by designing a *mystery puzzle*. In the game, participants may be tasked with collaboratively putting together the puzzle without a visual guide. To solve the mystery and piece the puzzle together, participants might be given information from a facilitator. For example, we might have a facilitator who is given special knowledge about the story's outcome and is tasked with leading other players to the solution by providing hints through telling stories and anecdotes from characters who are identified as being on the ASD spectrum. By promoting empathy through storytelling, the facilitator can complement the collaborative environment needed to piece together a puzzle of an unknown visual outcome.

Learning System

The core of any educational game is the learning system. In this section, we present the findings that we have concluded from the data to design learning systems within games for sensitive topics.

Target the Game to a Player Group in a Specific Setting

In the evaluation of the learning system, we find that there is a need for careful consideration and targeting of the player group because of the sensitive nature of the conversation. Context plays a large role in players' experience of the game. Playing the game at work, home, school, or a workshop could result in unique experiences with other players and game content. The needs of a single player at these instances change owing to the environment, which encapsulates them. Therefore, player 1 may have a specific set of needs that are divided into player 1.1 in a professional setting, player 1.2 in a social setting, and so on. Hence, there is a need for specific targeting.

Human-centered design, the paradigm of work that is described in the paper by Norman [68] is nearly synonymous with UX, human–computer interaction, and subsequently games user research. However, the targeting recommended by experts in our data set demonstrates that the level at which you specify your user-focus depth and breadth needs to be carefully considered. Comments from experts demonstrate that *a person who is interested in learning* with age-group specifier is not enough. For example, the results suggest that there is a difference between a person approaching this with lived experience and a naïve player with only passing knowledge:

> *When you get a bunch of people together, and one of those people is a serious anxiety disorder sufferer and everybody else isn't, then when they talk about their experiences and anxiety, there's a whole pile of commonality, and one person is thinking `I'm not like these other people, I'm gonna probably keep my mouth shut'.* [Participant 3]

From the above quote, we see a demonstration of how a group with commonalities in identified target demographics (eg, age) can approach the game and unsuccessfully find common ground.

As game designers, we need to design not only for our target users but also to create an environment in which we can capture a subset of the user's personality and engage them in a conversation about sensitive topics. Therefore, customization should be based on the common traits of individuals in the target market and not specifically on GD heuristics for that age group.

To apply this design guideline to AbW, we would need to select a targeted setting for the game. For example, redesigning the game to be played with other people in a support group, which meets anonymously at a university space. The game would need to support a larger player group and specifically focus on challenges and emotions, as support groups are designed around discussing the impact of a condition or label and the lived experience of or feelings experienced by the users and design mechanics for player–player interaction. These player–player interactions can be adding a button in the app, which communicates that the feelings or experiences expressed by a player have resonated with others.

As an example of how to apply this principle to other game ideas, let us imagine a game that tackles the topic of children who are terminally ill. We may decide to design the game to be played with healthy siblings or other children visiting the hospital. To do so, we would need to understand the setting of the children's hospital and the disparity in ability among the groups of players. We may choose to design a game that requires less cognitive effort or physical dexterity, such as games that focus on creative expression and imagination. For example, imagine a game in which participants are using tangram pieces to tell a story. In the game, younger children may have a chance to express their feelings indirectly by telling stories using these large, easy-to-grip, color pieces. Here, this playful activity can be played on small surfaces such as a bedside table or larger surfaces such as the floor.

Design Collaborative Challenges

A large body of literature has shown that challenge in games can be important to the motivation to play, the game atmosphere, and player–player interactions [69-71].

A subset of participants felt that the game needed more strategy, in the traditional sense that board games allow participants to occupy their thoughts between their own turns by thinking about their next move. However, AbW's challenge was the discussion of the sensitive topic of MH:

> *There's - this - this to me right now, unless I'm missing something as a naive player, doesn't feel like there's any dominant strategy, except for doing exactly the choice that you need to do each round to be able to move forward. So I think from a gameplay perspective, having to make some - some uncertain or unclear choices at some point would enhance the play value, but as often - I've seen like with games like this before, I don't know - I don't know if that's requesting too much cognitive lay from people who are already managing a lot of other information, you know. And it also depends on the context and audience. You know, if this is for a discussion group, it is possible that you don't - you don't want to put too much game play at the expense of people really just being able to have an experience shared conversation together. So - so that's gonna shift with your audience, but if it was - if the charge was just increase game play interest level and this, I would have - I would have more choices and more obscured information about what's the best choice to make was.* [GD2]

To apply this guideline to AbW, the goal of the game would need to change from *collect all your life goal cards* to *help everyone in the group to collect their life goal cards*. Changing the goal would orient players to a more collaborative environment and complement existing mechanisms for sharing and discussion, such as the share cards or open-hand card game protocol.

We present an alternative example of how this could be used is designing a game to discuss privilege, which can be considered as personal factors, which convey advantage in a situation. As a thought experiment, consider a massive multiplayer web-based game world (eg, Elder Scrolls Online, Bethesda 2014), and imagine instead of the game asking players to choose their race, they were randomly assigned one. Furthermore, let race X be extremely disadvantaged, unable to fight against other races in hand-to-hand combat, and reduce playability of the game. Instead of making the goal *triumph over others in hand-to-hand combat*, consider making the goal *design an arena system with hand-to-hand combat that balances the scales for players who randomly were born into X race*. Changing the goal of the game from individualist goals to collective player goals creates an entirely different game, and this new game creates a better environment for the discussion of privilege by abstracting the conversation from personal traits to in-game traits.

Technical and Practical Design (Autonomy and Divergent Paths)

The following section refers to the design of interface, navigation, and meta-game environment (eg, how the game is physically situated based on how we expect players to set up the game). In this section, we also discuss practical aspects of researching and designing the game itself for transparency and replicability.

Navigation design is especially important because participants will want to discretely access information during gameplay. We also need to allow for an *escape route* or a method for participants to exit if they are uncomfortable.

For AbW, we may apply this principle in the redesign by allowing participants to change the topic of a share card by anonymously triggering a redraw. To disguise player identity, we may also use the app to randomly call a redraw at times. This would allow participants to avoid conversations based on their own preferences and potentially pass the blame for abrupt change of subject.

We may apply this principle to other game ideas as well. Imagine a game in which web-based players can explore an open world such as a massive multiplayer RPG. Storylines written in the game can tackle sensitive issues through web-based role-play; however, common to massive multiplayer RPGs, there is freedom to choose what areas of the map to explore, form a guild or team with other players, or follow a narrative story path.

Guidelines for Researching Pregame Design

Overview

The review of the literature presented above provides credence to the importance of games user research. In turn, GD should also be based on research. Simply, this may be researching the persona of the target audiences or gathering information for the story based on real-world objects. Researching before designing games only becomes more important as we add purpose outside the entertainment value. As our purposes shift to more serious and sensitive topics, our results demonstrate that games for sensitive topics especially need to be researched.

When researching a sensitive topic with the intention of designing a game, consider the diversity of the topic as it spans multiple points of further inquiry. Results from the 2 expert participant groups allowed for triangulation and insight into the technical and practical design of the game. Therefore, we suggest guidelines for research.

Research Diversely

To research diversely means to collect data beyond the scope of information presently captured in one field (eg, GD or health). To implement this principle, the scope of the research must be open to all relevant fields and experts within the scope of the project. Doing so allows one to design a game based on a wide variety of expertise and diversify research beyond general field knowledge and GD.

In other words, game designers should not claim expertise beyond game designing without formal training, hence the need for expert reviews.

To apply this design guideline, we must consider which professionals engage in the topic materials. In the case of AbW, this meant including data from both game designers and MH professionals. It is evident from our results that these 2 groups of experts provided different information that could be contrasting or complimentary at times. To apply this guideline to improve *AbW*, future work may consider interviewing another set of expert professionals: social workers, personal support workers, registered practical nurses, and registered nurses.

Unlike other MH professionals interviewed in this study, these individuals have a role in extending care, beyond the individual, to the family as a whole. These experts would be able to provide insight on not only the individual's health journey but also the individual's effect on their own community [65].

As an alternate example, consider a game about *sexual orientation and LGBTQ2++ lifestyle*. As usual, game designers will look at the literature, existing state of the art, and other GD professionals; however, we might also ask physicians, psychologists, and neurologists. Moreover, a game designer may also find that activists, women's and feminist studies scholars, community leaders, and club owners may provide insight into the complex and wide-reaching aspects that encompass the life of an individual belonging to the lesbian, gay, bisexual, transgender, and queer 2++ community.

Beware of Assumed Specialization and Misconceptions That Surround the Game When Conducting Research

During the development of AbW and the presented research project, we spoke to many individuals. We consider the assumptions of individuals to whom we communicated our research ideas. During the development and research of the game, AbW, it was apparent that a game for MH was assumed to be for individuals with MH concerns. Ironically, the game was targeted at the exact people who assumed this to battle stigma. Ideally, the game would communicate that it is the community that needs to be informed, work together, and battle misinformation. The target is to impart an understanding that MH is a part of overall health and to restructure the idea that MH is a condition that affects only a subset of people and instead think of MH similar to the way we think of any physical aliment. For example, consider the common cold. At any time, a subset of people might be battling a cold; however, any individual in our community could catch a cold. Therefore, all community members should recognize the basic signs and symptoms of a cold and know basic information about treatment options and where to seek help. Similar to the common cold, a community should also treat individuals with MH concerns as working toward recovery.

During the course of the study, we found that both HPs and game designers got the impression that the game was designed to be a treatment or clinical tool despite our emphasis on it being a game for education. Expert participants were not immune to the assumption, and some expert participants expected the game to be played with a clinical target, even when specified otherwise. To apply this design guideline to AbW, we would see this misclassification as a need for better clarity in the presentation of the game.

As explained in the results, the game, AbW, is at the crossroads of development and therefore one way to counter the assumptions that follow this game (ie, AbW is a game only for people with MH concerns) and apply this design guideline would be to rebrand the game. For example, we may consider a larger group size, a younger audience [13-17], and a school setting. Then, the game would be for general education of high school children during a health or gym course session.

As an alternative example, consider a game being designed for education about a health condition such as type 2 diabetes. The assumption may be that the game would educate individuals newly diagnosed with this condition. However, this disease is likely preventable with lifestyle modifications such as a healthy diet and exercise [72]. Therefore, the game being designed could be for a larger audience. Designers would need to decide, in the early design stages, to determine whether the game will be marketed to all individuals or only those who were recently diagnosed.

Limitations

Our study combined expertise from multiple disciplines to yield multiple perspectives on the same set of information. As explained in the aforementioned discussion section *(Research Diversely)*, increasing the participant sample to a new domain of expert participants will yield a new viewpoint on the same problem. For example, interviewing social workers [73] may yield new information. Future research should aim to gain these additional perspectives.

Moreover, a counter methodology to presenting a cultural probe (as we have done in this study) may reveal additional insights that are not considered by the game designers. For example, the use of participatory design methodology [74] may reveal interesting insights and approaches.

Conclusions

On the basis of our presented study, we discuss the design of games for sensitive topics. We used the bespoke game, AbW, as a research probe. Our work resulted in the contribution of guidelines that focus on the comfort of participants, optimization of the learning system, and technical and practical guidelines. We made our findings accessible by providing examples for their applications, as applied to AbW, and suggested a possible game idea to germinate creative thinking in our readers.

Comparison With Previous Work

Previous work on this project was presented as part of a Association Computing Machinery (ACM) special Interest Group (sig) Computer Human Interactions in Play (CHI PLAY) Student Game Design Competition (SGDC) [59] and a conference demonstration with accompanying short paper [17]. These articles can be referred to as earlier versions of the game presented with accompanying documentation and conceptualizations. In this paper, we see the work in its most recent iteration evaluated. Through the process of analysis, we contribute a better understanding of how to research and design games for sensitive topics and MH. In addition, we solidify our conclusions and make them accessible with examples, to iterate on both the presented project and alternative game examples.

Acknowledgments

The authors would like to thank the alumni of this project, laboratory assistants, and volunteers for their support. The authors warmly thank their expert participants. In addition, RRW would like to extend gratitude to Chrysanne DiMarco, who instructed at the University of Waterloo. The authors would also like to thank the supporting institutions and departments who have supported this study, including the Games Institute and Cheriton School of Computer Science, University of Waterloo, Canada. We also thank the OH! Lab and Human Computer Interaction Institute, Carnegie Mellon University, United States. This project would not have been possible without the support of the funding bodies: National Sciences and Engineering Research Council of Canada that awarded RRW with Postgraduate Scholarships D and further supported LEN and EL, each with a discovery grant. LEN also would like to thank the National Sciences and Engineering Research Council for the Collaborative Research Experience Training Program Saskatchewan Waterloo Games User Research grant and Canadian Institutes of Health Research Active and Assisted Living Program grant. The authors would like to thank the Games Institute, University of Waterloo, for their support for this project through the Interactive and Multi-Modal Experience Research Syndicate and Social Science and Humanities Research Council. This paper was published in memory of author EL, whose impact and contribution to science will not be forgotten.

Authors' Contributions

RRW is the lead author, lead game designer, and researcher. RRW worked on all components of the project including project conceptualization, research design, ethics, data collection, and data analysis. RRW also led the writing and worked on all sections: literature review, methods, results, discussion, and editing, and the submission process. RRW also handled administration and advised less experienced colleagues in direction. CW is the second author and lead volunteer coordinator; advised less experienced colleagues; handled ethics, data collection, and data analysis; wrote methods and results; and worked on editing, and the submission process. YE is the third author and former volunteer and undergraduate research assistant funded by the National Sciences and Engineering Research Council award. YE worked on data collection, data analysis, literature review, and editing, and the submission process. AS is the fourth author, undergraduate research assistant, and volunteer who worked on data analysis, literature review, and editing, and the submission process. LEN was the undergraduate research advisor to CW and AS. LEN also participated in the ethics application and submission process. JH is the external collaborator, advised and mentored less experienced colleagues, handled ethics, worked on editing, and was involved in the submission process. EL is the PhD supervisor of RRW; guided ethical research protocol; and worked on research design, writing, and editing of the paper, and the submission process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Related publications and media.
[DOCX File , 19 KB - games_v10i2e21376_app1.docx]

Multimedia Appendix 2
Interview questions.
[DOCX File , 22 KB - games_v10i2e21376_app2.docx]

References

1. Polito JM. Effective communication during difficult conversations. Neurodiagn J 2013;53(2):142-152. [Medline: 23833841]
2. Meyer EC, Sellers DE, Browning DM, McGuffie K, Solomon MZ, Truog RD. Difficult conversations: improving communication skills and relational abilities in health care. Pediatr Crit Care Med 2009;10(3):352-359. [doi: 10.1097/PCC.0b013e3181a3183a] [Medline: 19325506]
3. Staller KM. Difficult conversations: talking with rather than talking at. Qual Soc Work 2014;13(2):167-175. [doi: 10.1177/1473325014521335]
4. Bögels SM, Hellemans J, van Deursen S, Römer M, van der Meulen R. Mindful parenting in mental health care: effects on parental and child psychopathology, parental stress, parenting, coparenting, and marital functioning. Mindfulness 2014;5(5):536-551. [doi: 10.1007/s12671-013-0209-7]
5. Insel TR. Assessing the economic costs of serious mental illness. Am J Psychiatry 2008;165(6):663-665. [doi: 10.1176/appi.ajp.2008.08030366] [Medline: 18519528]
6. Parcesepe AM, Cabassa LJ. Public stigma of mental illness in the United States: a systematic literature review. Adm Policy Ment Health 2013;40(5):384-399 [FREE Full text] [doi: 10.1007/s10488-012-0430-z] [Medline: 22833051]
7. Dingfelder SF. Stigma: alive and well. American Psychological Association. 2009. URL: https://www.apa.org/monitor/2009/06/stigma [accessed 2022-02-15]
8. Shapiro S, Rotter M. Graphic depictions: portrayals of mental illness in video games. J Forensic Sci 2016;61(6):1592-1595. [doi: 10.1111/1556-4029.13214] [Medline: 27783392]

9. Martin JM. Stigma and student mental health in higher education. High Educ Res Dev 2010;29(3):259-274. [doi: 10.1080/07294360903470969]
10. Morgan AJ, Ross A, Reavley NJ. Systematic review and meta-analysis of Mental Health First Aid training: effects on knowledge, stigma, and helping behaviour. PLoS One 2018;13(5):e0197102 [FREE Full text] [doi: 10.1371/journal.pone.0197102] [Medline: 29851974]
11. Corrigan PW, Morris SB, Michaels PJ, Rafacz JD, Rüsch N. Challenging the public stigma of mental illness: a meta-analysis of outcome studies. Psychiatr Serv 2012;63(10):963-973. [doi: 10.1176/appi.ps.201100529] [Medline: 23032675]
12. Juul J. The magic circle and the puzzle piece. In: Proceedings of the 2008 Conference The Philosophy of Computer Games. 2008 Presented at: The Philosophy of Computer Games '08; May 8-10, 2008; Postdam, Germany.
13. Granic I, Lobel A, Engels RC. The benefits of playing video games. Am Psychol 2014;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
14. Mueller F, Gibbs MR, Vetere F. Design influence on social play in distributed exertion games. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery; 2009 Presented at: CHI '09; April 4-9, 2009; Boston, MA p. 1539-1548. [doi: 10.1145/1518701.1518938]
15. Mueller F, Richard B, Andres J, Patibanda R. Experiencing the body as play. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery; 2018 Presented at: CHI '18; April 21-26, 2018; Montréal, Canada p. 210. [doi: 10.1145/3173574.3173784]
16. Fares J. Brothers: a tale of two sons. Starbreeze Studios. 2013. URL: https://store.steampowered.com/app/225080/Brothers__A_Tale_of_Two_Sons/ [accessed 2022-02-15]
17. Wehbe RR, Watson DK, Fortes Tondello G, Nacke L. ABOVE WATER: extending the play space for health. In: Proceedings of the 2016 ACM on Interactive Surfaces and Spaces. New York: Association for Computing Machinery; 2016 Presented at: ISS '16; November 6-9, 2016; Ontario, Canada p. 497-499. [doi: 10.1145/2992154.2996882]
18. The state of mental health in America 2018. Mental Health America. 2017. URL: http://www.mentalhealthamerica.net/issues/state-mental-health-america [accessed 2022-02-15]
19. Making the case for investing in mental health in Canada. Mental Health Commision of Canada. 2013. URL: https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf [accessed 2022-02-15]
20. Schmidt CW. Environmental connections: a deeper look into mental illness. Environ Health Perspect 2007;115(8):A404-A410 [FREE Full text] [doi: 10.1289/ehp.115-a404] [Medline: 17687431]
21. Schneider EF, Lang A, Shin M, Bradley SD. Death with a story: how story impacts emotional, motivational, and physiological responses to first-person shooter video games. Hum Commun Res 2004;30(3):361-375. [doi: 10.1093/hcr/30.3.361]
22. Felitti VJ, Anda RF, Nordenberg D, Williamson DF, Spitz AM, Edwards V, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. Am J Prev Med 1998;14(4):245-258. [doi: 10.1016/s0749-3797(98)00017-8] [Medline: 9635069]
23. Centers for Disease Control and Prevention, Substance Abuse and Mental Health Services Administration, National Association of County Behavioral Health and Developmental Disability Directors, National Institute of Mental Health, The Carter Center Mental Health Program. Attitudes toward mental illness: results from the behavioral risk factor surveillance system. Centers for Disease Control and Prevention. 2012. URL: https://www.cdc.gov/hrqol/Mental_Health_Reports/pdf/BRFSS_Full%20Report.pdf [accessed 2022-02-15]
24. Cheung R, O'Donnell S, Madi N, Goldner E. Factors associated with delayed diagnosis of mood and/or anxiety disorders. Health Promot Chronic Dis Prev Can 2017;37(5):137-148 [FREE Full text] [doi: 10.24095/hpcdp.37.5.02] [Medline: 28493658]
25. Son C, Hegde S, Smith A, Wang X, Sasangohar F. Effects of COVID-19 on college students' mental health in the United States: interview survey study. J Med Internet Res 2020;22(9):e21279 [FREE Full text] [doi: 10.2196/21279] [Medline: 32805704]
26. Lebowitz MS. Biological conceptualizations of mental disorders among affected individuals: a review of correlates and consequences. Clin Psychol Sci Prac 2014;21(1):67-83. [doi: 10.1111/cpsp.12056]
27. Wang PS, Angermeyer M, Borges G, Bruffaerts R, Tat Chiu W, De Girolamo G, et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007;6(3):177-185 [FREE Full text] [Medline: 18188443]
28. Quintero Johnson JM, Yilmaz G, Najarian K. Optimizing the presentation of mental health information in social media: the effects of health testimonials and platform on source perceptions, message processing, and health outcomes. Health Commun 2017;32(9):1121-1132. [doi: 10.1080/10410236.2016.1214218] [Medline: 27573748]
29. Ridout B, Campbell A. The use of social networking sites in mental health interventions for young people: systematic review. J Med Internet Res 2018;20(12):e12244 [FREE Full text] [doi: 10.2196/12244] [Medline: 30563811]
30. Committee on the Science of Changing Behavioral Health Social Norms, Board on Behavioral, Cognitive, and Sensory Sciences, Division of Behavioral and Social Sciences and Education, National Academies of Sciences, Engineering, and Medicine. Ending discrimination against people with mental and substance use disorders: the evidence for stigma change. Washington, D.C: The National Academies Press; 2016.

31. Committee on the Science of Changing Behavioral Health Social Norms, Board on Behavioral, Cognitive, and Sensory Sciences, Division of Behavioral and Social Sciences and Education, National Academies of Sciences, Engineering, and Medicine. Approaches to Reducing Stigma. In: Ending discrimination against people with mental and substance use disorders: the evidence for stigma change. Washington, D.C: The National Academies Press; 2016:69-92.
32. Our impact. Time to change. 2022. URL: https://www.time-to-change.org.uk/about-us/our-impact [accessed 2022-02-15]
33. Bowman SL. The functions of role-playing games: how participants create community, solve problems and explore identity. Jefferson, NC: McFarland & Company; 2010.
34. O'Connor EL, Longman H, White KM, Obst PL. Sense of community, social identity and social support among players of massively multiplayer online games (MMOGs): a qualitative analysis. J Community Appl Soc Psychol 2015;25(6):459-473. [doi: 10.1002/casp.2224]
35. 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]
36. Lau HM, Smit JH, Fleming TM, Riper H. Serious games for mental health: are they accessible, feasible, and effective? A systematic review and meta-analysis. Front Psychiatry 2017;7:209 [FREE Full text] [doi: 10.3389/fpsyt.2016.00209] [Medline: 28149281]
37. Wang H, Sun CT. Game reward systems: gaming experiences and social meanings. In: Proceedings of the 2011 DiGRA International Conference: Think Design Play. 2011 Presented at: DiGRA '11; September 14-17, 2011; Hilversum, The Netherlands p. 1-15.
38. Jakobsson M, Sotamaa O. Special issue: game reward systems. Game Stud 2011;11(1).
39. Vugts MA, Zedlitz AM, Joosen MC, Vrijhoef HJ. Serious gaming during multidisciplinary rehabilitation for patients with chronic pain or fatigue symptoms: mixed methods design of a realist process evaluation. J Med Internet Res 2020;22(3):e14766 [FREE Full text] [doi: 10.2196/14766] [Medline: 32149720]
40. Gudoniene D, Bartkute R, Rutkauskiene D, Blazauskas T. Technological aspects of the gamification model for e-learning participant's engagement. Balt J Mod Comput 2016;4(4):1008-1015. [doi: 10.22364/bjmc.2016.4.4.25]
41. Apostol S, Zaharescu L, Alexe I. Gamification of learning and educational games. In: 9th International Scientific Conference eLearning and software for Education. 2013 Presented at: eLSE '13; April 25-26, 2013; Bucharest, Romania p. 67-72.
42. van der Meulen H, McCashin D, O'Reilly G, Coyle D. Using computer games to support mental health interventions: naturalistic deployment study. JMIR Ment Health 2019;6(5):e12430 [FREE Full text] [doi: 10.2196/12430] [Medline: 31094346]
43. Cangas AJ, Navarro N, Parra JM, Ojeda JJ, Cangas D, Piedra JA, et al. Stigma-stop: a serious game against the stigma toward mental health in educational settings. Front Psychol 2017;8:1385 [FREE Full text] [doi: 10.3389/fpsyg.2017.01385] [Medline: 28878702]
44. Frasca G. Rethinking agency and immersion: video games as a means of consciousness-raising. Digit Creat 2010;12(3):167-174. [doi: 10.1076/digc.12.3.167.3225]
45. Hanisch SE, Birner UW, Oberhauser C, Nowak D, Sabariego C. Development and evaluation of digital game-based training for managers to promote employee mental health and reduce mental illness stigma at work: quasi-experimental study of program effectiveness. JMIR Ment Health 2017;4(3):e31 [FREE Full text] [doi: 10.2196/mental.7600] [Medline: 28778839]
46. Krotoski A. Serious fun with computer games. Nature 2010;466(7307):695. [doi: 10.1038/466695a]
47. Juul J. The art of failure: an essay on the pain of playing video games. Cambridge, MA: MIT press; 2013.
48. 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;8(1):e13459 [FREE Full text] [doi: 10.2196/13459] [Medline: 32134391]
49. 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;2(2):e11 [FREE Full text] [doi: 10.2196/mental.4133] [Medline: 26543916]
50. Corder BF, Whiteside R, Vogel M. A therapeutic game for structuring and facilitating group psychotherapy with adolescents. Adolescence 1977;12(46):261-268. [Medline: 883537]
51. Chen VH, Wu Y. Group identification as a mediator of the effect of players' anonymity on cheating in online games. Behav Inf Technol 2015;34(7):658-667. [doi: 10.1080/0144929x.2013.843721]
52. Wettergren L, Eriksson LE, Nilsson J, Jervaeus A, Lampic C. Online focus group discussion is a valid and feasible mode when investigating sensitive topics among young persons with a cancer experience. JMIR Res Protoc 2016;5(2):e86 [FREE Full text] [doi: 10.2196/resprot.5616] [Medline: 27161146]
53. Mar RA, Oatley K, Peterson JB. Exploring the link between reading fiction and empathy: ruling out individual differences and examining outcomes. Communications 2009;34(4):407-428. [doi: 10.1515/comm.2009.025]
54. Greitemeyer T. Playing video games cooperatively increases empathic concern. Soc Psychol 2013;44(6):408-413. [doi: 10.1027/1864-9335/a000154]
55. Orlick TD. Positive socialization via cooperative games. Dev Psychol 1981;17(4):426-429. [doi: 10.1037/0012-1649.17.4.426]

56. Greitemeyer T, Osswald S, Brauer M. Playing prosocial video games increases empathy and decreases schadenfreude. Emotion 2010;10(6):796-802. [doi: 10.1037/a0020194] [Medline: 21171755]
57. Conati C. Probabilistic assessment of user's emotions in educational games. Appl Artif Intell 2002;16(7-8):555-575. [doi: 10.1080/08839510290030390]
58. Coulson M, Barnett J, Ferguson CJ, Gould RL. Real feelings for virtual people: emotional attachments and interpersonal attraction in video games. Psychol Pop Media Cult 2012;1(3):176-184. [doi: 10.1037/a0028192]
59. Wehbe R, Watson D, Tondello GF, Ganaba M, Stocco M, Lee A, et al. ABOVE WATER: an educational game for anxiety. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2016 Presented at: CHI PLAY Companion '16; October 16-19, 2016; Austin, TX p. 79-84. [doi: 10.1145/2968120.2971804]
60. Herrera AK, Wehbe RR, Mathews FZ. Educational games and health sciences. Network of the National Library of Medicine New England Region (NNLM NER) Repository. 2017. URL: https://escholarship.umassmed.edu/ner/44/ [accessed 2022-02-15]
61. Corrigan P. How stigma interferes with mental health care. Am Psychol 2004;59(7):614-625. [doi: 10.1037/0003-066X.59.7.614] [Medline: 15491256]
62. Charmaz K. Constructing grounded theory. Thousand Oaks, CA: Sage Publications; 2014.
63. Charmaz K. Grounded theory. In: Smith JA, Harré R, Van Langenhove L, editors. Rethinking methods in psychology. London, UK: Sage Publications; 1996:27-49.
64. Swanson C, Urner R, Lank E. Naïve security in a Wi-Fi world. In: 4th IFIP WG 11.11 International Conference Proceedings. 2010 Presented at: IFIPTM '10; June 16-18, 2010; Morioka, Japan p. 32-47.
65. Chappell A. Of dice and men: the story of dungeons and dragons and the people who play it. Am J Play 2014:290-292.
66. Essalmi F, Ben Ayed LJ, Jemni M, Kinshuk, Graf S. A fully personalization strategy of E-learning scenarios. Comput Hum Behav 2010;26(4):581-591. [doi: 10.1016/j.chb.2009.12.010]
67. Barbot O. George Floyd and our collective moral injury. Am J Public Health 2020;110(9):1253. [doi: 10.2105/AJPH.2020.305850] [Medline: 32614647]
68. Norman DA. Human-centered design considered harmful. interactions 2005;12(4):14-19. [doi: 10.1145/1070960.1070976]
69. Hung CY, Sun JC, Yu PT. The benefits of a challenge: student motivation and flow experience in tablet-PC-game-based learning. Interact Learn Environ 2015;23(2):172-190. [doi: 10.1080/10494820.2014.997248]
70. Tüzün H, Yılmaz-Soylu M, Karakuş T, İnal Y, Kızılkaya G. The effects of computer games on primary school students' achievement and motivation in geography learning. Comput Educ 2009;52(1):68-77 [FREE Full text] [doi: 10.1016/j.compedu.2008.06.008]
71. Moon H, Baek Y. Exploring variables affecting player's intrinsic motivation in educational games. In: Proceedings of the 17th International Conference on Computers in Education. 2009 Presented at: ICCE '09; November 30-December 4, 2009; Hong Kong p. 718-722.
72. Jermendy G. Can type 2 diabetes mellitus be considered preventable? Diabetes Res Clin Pract 2005;68(Suppl1):S73-S81. [doi: 10.1016/j.diabres.2005.03.010] [Medline: 15955380]
73. Videka-Sherman L. Metaanalysis of research on social work practice in mental health. Soc Work 1988;33(4):325-338 [FREE Full text] [doi: 10.1093/sw/33.4.325]
74. Douglas S, Namioka A. Participatory design: principles and practices. Mahwah, NJ: L. Erlbaum Associates; 1993.

Abbreviations

AbW: Above Water
ASD: autism spectrum disorder
BYOD: bring-your-own-device
GAD: generalized anxiety disorder
GD: game design
HP: health professional
MH: mental health
MHL: mental health literacy
PD: panic disorder
RPG: role-playing game
UX: user experience
VR: virtual reality

Edited by N Zary; submitted 25.08.20; peer-reviewed by M Lucassen, K Rodrigues; comments to author 14.11.20; revised version received 29.03.21; accepted 12.10.21; published 19.05.22.

Please cite as:
Wehbe RR, Whaley C, Eskandari Y, Suarez A, Nacke LE, Hammer J, Lank E
Designing a Serious Game (Above Water) for Stigma Reduction Surrounding Mental Health: Semistructured Interview Study With Expert Participants
JMIR Serious Games 2022;10(2):e21376
URL: https://games.jmir.org/2022/2/e21376
doi:10.2196/21376
PMID:35588056

Original Paper

Associations Between Addictive Behaviors, Individual Characteristics, and the Use of Gambling Services Within the World of Gaming: Cross-sectional Survey Study

Mark Kisch[1,2], MD; Anders Håkansson[1,3], MD, PhD

[1]Division of Psychiatry, Department of Clinical Sciences, Lund University, Lund, Sweden

[2]Malmö Addiction Center, Region Skåne, Malmö, Sweden

[3]Clinical Research Unit/Clinical Gambling Disorder Unit, Malmö Addiction Center, Region Skåne, Malmö, Sweden

Corresponding Author:
Anders Håkansson, MD, PhD
Division of Psychiatry
Department of Clinical Sciences
Lund University
Baravägen 1
Lund, 22100
Sweden
Phone: 46 070 3135677
Email: anders_c.hakansson@med.lu.se

Abstract

Background: Gambling within the world of gaming is an emerging phenomenon that may share common conceptual characteristics with traditional forms of gambling. The current literature suggests a higher degree of problematic behaviors in this gambling pattern, but studies are few, prompting for further research regarding individual characteristics and comorbid conditions associated with this activity.

Objective: The aim of the study is to investigate correlations between the use of gambling services within the world of gaming and individual characteristics and addictive behaviors including problem gambling.

Methods: A cross-sectional web survey was distributed to an existing panel of online respondents in Sweden. A total of 2001 respondents were included. Chi-square and Mann-Whitney *U* tests, followed by a logistic regression, were used in order to determine independent variables associated with gambling in the context of gaming.

Results: A total of 2.9% (58/1984) of respondents reported past-year gambling within gaming. Significant associations were found with male sex, younger age, history of treatment-seeking for alcohol problems, and higher Gaming Addiction Scale scores.

Conclusions: The demonstrated findings strengthen previously found associations between gambling in gaming and younger age, male sex, and problematic gaming behaviors. Additionally, the association with a history of treatment needs for alcohol problems adds to the previous impression of increased problem severity and comorbidity in within-gaming gamblers.

(JMIR Serious Games 2022;10(2):e29077) doi:10.2196/29077

KEYWORDS

gambling disorder; gaming disorder; behavioral addiction; mental health; gambling; gaming; addiction; behavior; cross-sectional; online survey; age; gender

Introduction

Gambling disorder is a psychiatric diagnosis describing a problematic gambling pattern leading to substantial impairment or mental distress [1]. The prevalence of either problem gambling or gambling disorder has been estimated to be between 0.1% and 5.8%, [2], and risk factors for problem gambling include male sex, young age, substance use, poor mental health, low level of formal education, and low socioeconomic status [3,4]. Gambling disorder can be treated in psychological therapies in the form of cognitive behavioral therapy [5]. Since the 1980s, the world has seen an incomparable growth in global commercial gambling, which is suggested to be due to changes in attitudes toward legal gambling or the increasing presence of internet and mobile devices and other nontraditional platforms that provide gambling services [6].

A number of harms have been described to be associated with a history of gambling disorder [7]. Financial harm is a dominant theme that may include debts and personal bankruptcy, affecting the gambler, relatives, and society as a whole. Furthermore, there are documented high rates of comorbid psychiatric disorders among gamblers. A meta-analysis performed on 36 studies documented comorbid, current psychiatric disorders in 74.8% (95% CI 36.5-93.9) of treatment-seeking gamblers [8]. The most prominent types are nicotine dependence, substance use disorders, mood disorders, and anxiety disorders [9]; suicidal ideation, suicide attempts [10], and completed suicides [11] are also seen.

An increasing proportion of gamblers engage in online platform activities [12], characterized by high availability, a short time between betting and outcome, light and sound effects, uninterrupted gaming sessions, variable stake sizes, near miss features, anonymity, and illusion of control [13]. As found in an annual report from the Gambling Commission in Great Britain, 18% of the respondents had gambled online in the past 4 weeks [14]. In addition, problem gambling is more common in internet gambling compared to noninternet gambling [15]. When compared to noninteractive gamblers, interactive gamblers are more likely to be younger, male, educated, and part of a group household. Interactive gamblers also tend to gamble on more activities and bet higher amounts [16]. In Sweden, a large majority of treatment-seeking gamblers are involved in online gambling activities [17]. Younger adults may represent a vulnerable subgroup with an increased risk to experience gambling problems, and this appears to be independent of gambling modality or degree of gambling participation [18,19]. The comparatively increased risk-taking and poor consequence thinking in this group has been suggested as a possible explanation [20].

The phenomenon of digital- and video-gaming is increasingly highlighted with respect to its addictive properties and problematic behaviors with negative outcomes [11,21]. Gaming disorder has appeared in the World Health Organization's *International Classification of Diseases* since 2018. The diagnosis is characterized by an at least 12-month period of impaired control of gaming, reprioritization of daily activities and interests that allows more time to spend on gaming, and continued gaming despite negative consequences, along with a clinically significant loss in function in important life areas [22]. In the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), internet gaming disorder is recognized as a condition warranting more clinical research before it might be listed as a formal disorder [1].

Computer gaming is a common phenomenon, especially in the lower age categories where the majority is engaged in this activity in some way [23]. Concerning the prevalence rate of internet gaming disorder among adolescents, data indicate that this is on the rise and has been estimated to be 1.6% in European adolescents and significantly more common in boys [24,25]. Due to the limited amount of longitudinal studies, causes and consequences in the field of pathological gaming are not fully investigated [26,27]. However, it seems that, for example, lower psychosocial well-being, including lower social competence and self-esteem, is an antecedent rather than a consequence of pathological gaming. Loneliness, on the other hand, turns out to be both a cause and consequence in this manner [28]. Other risk factors for developing internet gaming disorder include greater amounts of gaming participation and impulsivity traits, while depression, anxiety, and lower school grades should be seen as consequences [29]. In addition, general population data show that daily playing of computer games may be associated with a marked increase in the risk of problem gambling, regardless of gender [23].

In the context of gaming, specific features appear to overlap structurally and psychologically with conventional types of gambling for money [30]. Built-in gambling elements in games can take the form of loot boxes purchasable for real money, randomly generating virtual items that can be either cosmetic assets (ie, skins) or performatively advantageous in the game in question [31,32]. Virtual currencies can be sold or cashed out for real money, which creates a situation similar to traditional forms of gambling [30]. These convergences between gambling and gaming are often facilitated by digital microtransactions [33-35]. Notably, the authors label monetization technologies in games such as loot boxes as predatory due to their properties to encourage escalating player spending [36].

According to a recent Danish national survey, more than half of the gaming adolescents in the sample, with a male predominance, had engaged in loot box activity at some time [37]. Additionally, a positive correlation was identified between loot box engagement and problem gambling. In a Swedish population study, at-risk gambling was considerably more common in those who had bet on loot boxes [23]. Similar correlations have been recognized in the context of skin betting (gambling with items within the game that have value to the player's character) among adolescents [38]. Loot box involvement, giving rise to more problematic gaming behaviors, also has been indirectly linked to mental distress [39], and links between the amount spent on loot boxes and the degree of gambling problems have been acknowledged [32]. Loot box users are more likely to be young, employed, low-educated, and have an average household income [40].

Given these concerns, an important challenge worth emphasizing is the lack of international consensus for where the boundaries should be set for gambling-in-gaming activities to meet legal definitions on gambling [41-43]. These blurred boundaries between gambling and gaming call for more research addressing individual characteristics associated with the use of gambling services in the context of gaming.

Overall, the phenomenon of gambling within computer games has been increasingly highlighted, but studies of its correlates in the population are still few. Additionally, uncertainty remains about how this habit is associated with other health hazards, such as poor mental health and more traditional addictive disorders to tobacco, alcohol, and drugs, over and above the need to study it in the context of problem gambling and problem gaming. Therefore, the aim of this study is to, on the basis of responses from an electronic self-response questionnaire, investigate correlations between the use of gambling within gaming services while controlling for a number of sociodemographic, substance use, and mental health variables, as well as established measures of problem gaming and problem gambling.

Methods

Study Sample

A targeted Swedish segment of Userneeds web panel, consisting of a total of 115,000 individuals aged 16 years and older, was derived with the aim of accurately reflecting Sweden's general population in terms of sex and age distribution. From the way that data collection took place, it follows that it is not possible to determine how many individuals in fact were contacted in order to reach the intended number of respondents.

Out of a total of 2124 respondents, 116 individuals did not meet the primary inclusionary criteria which were formulated as having completed all the questions of the survey. Additionally, 7 sets of answers could be paired with possible doublets, originating from the same IP addresses and indicating duplicate answers. Consequently, the second round of these responses were excluded from the material, resulting in a final sample of 2001 completed surveys. Of these, only those responses that were consistent with the defined outcome measures for each question were used for further analysis in this study. This means that respondents who reported answers "don't know and/or don't want to answer" on questions where these answer options were available were designated as missing cases and hence not included in numerical calculations. The number of missing cases was different for each question but amounted at most to 48, so the total number of respondents with full data included in the final statistical analysis was n=1953.

Procedure

Recruitment of participants was carried out using the electronic questionnaire services of the company Userneeds, resulting in a cross-sectional study design. More specifically, a quantitative web survey was distributed to the Swedish segment of the company's multinational web panel consisting of preexisting survey respondents who had voluntarily signed up as members and provided Userneeds with personal information in order to regularly receive email invitations to participate in online surveys. The invitations sent out to this study contained written information regarding the anonymity of participation as well as on the topic of the survey referred to as "gambling, addiction to gambling, and mental behavioral and substance use disorders." In order to start responding to the questionnaire, participants were required to give their written electronic consent. Data collection took place over 17 days in September 2019 and was scheduled to cease when 2000 individual responses were achieved. As a general reward for completing surveys distributed from Userneeds, responders receive credits equivalent to approximately one euro in the company's bonus system.

As mentioned above, the survey contained a number of questions focusing on gambling behaviors but beyond that also intended to capture the sociodemographic background of respondents as well as comorbidities in the form of mental illness and other addictions. Based on the replies to whether or not the respondent in the past year had engaged in gambling services within computer games, the material was divided into 2 main groups: those who had engaged in this activity (n=58) and those who had not (n=1926). A total of 17 respondents answered "don't know and/or don't want to answer" on this question and were consequently not included in the study. The 2 groups with valid answers, however, were studied separately and compared in respect of correlates by means of variables obtained from the survey responses.

Ethics Approval

From the same overall research project, one study has been published and another is under review addressing other aspects of gambling and behavioral addictions [44]. The Ethics Review Authority, Sweden, reviewed the ethics application (number 2019-04176) and found the study did not formally require ethical permission as no personal data were used and also expressed that it had no ethical concerns with the study. Respondents enrolled in the study participated on a completely voluntary basis and gave their written consent in order to be able to start answering the questionnaire.

Measures

Sociodemographic Variables

Gender was categorized into the binary groups of male and female; those who preferred not to answer the question of gender identity were excluded from the study. Age was stratified into 6 brackets of age ranges (16-19 years, 20-24 years, 25-29 years, 30-39 years, 40-49 years, and 50 years and older). Monthly income was stratified into 10 brackets measured in Swedish krona (SEK) from <10,000 SEK (US $1000) to >50,000 SEK (US $5000). Level of education was categorized on the basis of ever having attended university or not, regardless of obtained degree. Finally, the living status was spaced into 5 categories: single with children living at home, single without children living at home, living with a husband/wife or partner with children at home, living with a husband/wife or partner without children, or living with parents.

Comorbidity Variables

In order to identify serious mental illness among the respondents, the Kessler Psychological Distress Scale (K6) was integrated into the survey. K6 is a 6-item self-reported tool designed to screen for mood and anxiety disorders over the past 6-month period. Each question in the inventory investigates to what extent distinct feelings of nervousness, hopelessness, restlessness, depressiveness, lethargy, and worthlessness have been experienced. The responses are consequently coded with scores depending on the stated frequency of each question where 0=none of the time, 1=little of the time, 2=some of the time, 3=most of the time, and 4=all of the time. This system results in a total score range extending from 0 to 24 [45]. A categorical variable was defined by a summated cutoff score of ≥13, having a sensitivity of 0.36, and a specificity of 0.96 for detecting serious mental illness [45]. In the collected material for this study, 24 individuals had missing answers on one or two K6 items. In these cases, the individuals' median values of the available items were used to calculate their complete K6 scores.

In addition to K6, the survey included questions about having a history of seeking treatment for either psychological distress or problems with alcohol, illicit drugs, or prescription drugs. A question about daily use of nicotine, as in smoking or snuff, was also included.

Problem Gambling Severity Index

The 9-item scale [46,47] of the Problem Gambling Severity Index (PGSI) consists of questions asking about how frequently the responder engage in behaviors related to problem gambling, thus aiming to determine problem gambling severity. Each answer is scored from 0 to 4 based on a Likert scale consisting of the following response options: 0 (never), 1 (sometimes), 2 (most of the time), and 3 (almost always). Hence, depending on the total score, ranging on a continuum from 0 to 27, responders can traditionally be classified as being a nongambler or nonproblem gambler (0 points), low-risk gambler (1-2 points), moderate-risk gambler (3-7 points), or problem gambler (≥8 points). In this study, however, PGSI scores were used in a continuous manner rather than with the intention of categorizing participants into certain gambling profiles.

In order to assess gambling behaviors more generally, the survey contained questions on whether respondents have engaged in other forms of past-year gambling activities (online casino, physical casino, online horse betting, physical horse betting, online sports betting, physical sports betting, online poker, physical poker, online bingo, and physical gambling machines).

Gaming Addiction Scale

The Gaming Addiction Scale (GAS) is a 7-item questionnaire documented to have high reliability and generate consistent results across various samples [48]. The questions included in the GAS are based on DSM-based criteria for internet gaming addiction, and response options were subsumed according to a Likert scale ranging from 1=never to 5=very often. Assessing the past 6-month period, the questions asked intend to measure gaming-related preoccupation, tolerance, escapism, difficulties of quitting, withdrawal signs, social conflicts, and problems [48]. In the same way as for the PGSI, GAS scores were used as continuous measures in this study.

Statistical Analysis

The statistical analyses were conducted using SPSS (version 26, IBM Corp). First, the group of responders that in the past year had engaged in gambling services within computer games was compared with the group who had not engaged in this activity with regard to 20 categorical variables. These comparisons were made using chi-squared tests and included gender; age; monthly income; living status; level of education; serious mental illness according to K6; history of treatment-seeking due to psychological distress, alcohol problems, or problems with narcotics/drugs; and daily use of nicotine as well as past-year engagement in other traditional gambling activities. The total scores of the variables PGSI and GAS were kept as continuous scales rather than classifying tools and compared between the groups in terms of median values using nonparametric Mann-Whitney *U* tests. In order to investigate adjusted correlates of the use of gambling services within computer games, variables that according to the above-mentioned statistical models demonstrated statistically significant differences between the groups (P<.05) could enter a binary logistic regression. In order to stay within the recommended range of minimum events per predictor variable [49], the variables measuring other forms of gambling activity in the past year were not included in the logistic regression analysis. That way, a total of 10 variables entered the final analysis, which meant an events per predictor variable of 5.8. Results of the logistic regression were reported as odds ratios with 95% confidence intervals.

Results

Among the 2001 responders who completed the survey, 17 individuals answered "don't know and/or don't want to answer" on whether they had engaged in gambling services within computer games in the past year. Among the remaining 1984 responders with valid answers on this question, 2.9% (58/1984) reported engagement in this activity during the given period.

The unadjusted comparison between gamblers in computer games and the remaining respondents of the sample showed that those who participated in gambling services within computer games in the past year were significantly (P<.05) more likely to be male, younger, and single or living with their parents, as well as to have a lower degree of education. In terms of comorbidity variables, significant associations were found between gambling in gaming and serious mental illness according to K6; treatment-seeking for problems with alcohol, illicit drugs, and prescription drugs; and daily use of nicotine. Involvement in gambling services within the world of gaming was significantly associated with higher past-year engagement in other forms of gambling activities (online casino, physical casino, online horse betting, physical horse betting, online sports betting, physical sports betting, online poker, physical poker, online bingo, and physical gambling machines) as well as higher median scores on the PGSI (8 vs 0, P<.001) and GAS (21.5 vs 7, P<.001; Table 1).

Table 1. Comparison between gamblers in computer games and the rest of the sample using chi-square and Mann-Whitney analyses (n=1986).

	Past-year gambling within computer games (n=58), n (%)	No past-year gambling within computer games (n=1926), n (%)	*P* value	Missing ("don't know or don't want to answer")
Male gender	45 (77.6)	930 (48.3)	<.001	2
Age (years)	—[a]	—	<.001	0
16-19	9 (15.5)	49 (2.5)	—	—
20-24	13 (22.4)	98 (5.1)	—	—
25-29	12 (20.7)	178 (9.2)	—	—
30-39	8 (13.8)	352 (18.3)	—	—
40-49	13 (22.4)	477 (24.8)	—	—
>50	3 (5.2)	772 (40.1)	—	—
Monthly income (SEK)	—	—	.19[b]	0
<10,000	12 (20.7)	175 (9.1)	—	—
10,000-15,000	6 (10.3)	182 (9.4)	—	—
15,000-20,000	5 (8.6)	179 (9.3)	—	—
20,000-25,000	4 (6.9)	199 (10.3)	—	—
25,000-30,000	8 (13.8)	308 (16.0)	—	—
30,000-35,000	8 (13.8)	302 (15.7)	—	—
35,000-40,000	4 (6.9)	225 (11.7)	—	—
40,000-45,000	2 (3.4)	127 (6.6)	—	—
45,000-50,000	1 (1.7)	66 (3.4)	—	—
>50,000	8 (13.8)	163 (8.5)	—	—
Living status	—	—	<.001	0
Single with children at home	5 (8.6)	103 (5.3)	—	—
Single without children at home	17 (29.3)	496 (25.8)	—	—
With husband/wife or partner and children	14 (24.1)	574 (29.8)	—	—
With husband/wife or partner and no children	11 (19.0)	659 (34.2)	—	—
Living with parents	11 (19.0)	94 (4.9)	—	—
Attended university	19 (32.8)	1084 (56.3)	<.001	0
Serious mental illness on the K6[c] (≥13 points)	18 (31.0)	170 (8.9)	<.001	10
Need of treatment for psychological distress	27 (47.4)	695 (36.6)	.10	27
Need of treatment for alcohol problems	13 (23.6)	66 (3.4)	<.001	14
Need of treatment for illicit or prescription drug problems	8 (14.3)	30 (1.6)	<.001	8
Daily use of nicotine (smoking or snuffing)	18 (32.1)	301 (15.7)	.001	8
Past-year gambling activities	—	—	—	—
Casino, online	26 (47.3)	133 (6.9)	<.001	9
Casino, physical	16 (28.1)	57 (3.0)	<.001	6
Horse betting, online	25 (43.1)	248 (12.9)	<.001	8
Horse betting, physical	16 (29.1)	196 (10.2)	<.001	13
Sports betting, online	26 (47.3)	303 (15.8)	<.001	8
Sports betting, physical	21 (36.8)	217 (11.3)	<.001	6
Poker, online	19 (33.3)	49 (2.5)	<.001	4

	Past-year gambling within computer games (n=58), n (%)	No past-year gambling within computer games (n=1926), n (%)	*P* value	Missing ("don't know or don't want to answer")
Poker, physical	12 (21.8)	54 (2.8)	<.001	5
Bingo, online	25 (44.6)	84 (4.4)	<.001	6
Gambling machines, physical	22 (37.9)	81 (4.2)	<.001	1

[a]Not applicable.

[b]Linear by linear association.

[c]K6: Kessler Psychological Distress Scale.

In logistic regression of the statistically significant variables (except the variables on past-year engagement in other forms of gambling activities), gambling in gaming remained significantly associated with male sex, younger age groups, treatment-seeking for alcohol problems, and higher median scores on the GAS (Table 2).

Table 2. Potential correlates of gambling within the world of gaming (logistic regression including variables with significance in bivariate analyses [n=1953]).

	Odds ratio (95% CI)
Male gender	3.30 (1.53-7.12)
Age	0.76 (0.60-0.97)
Living status	0.87 (0.65-1.15)
Attended university	0.65 (0.33-1.31)
Serious mental illness on the K6[a] (≥13 points)	0.96 (0.42-2.18)
Need of treatment for alcohol problems	3.08 (1.09-8.69)
Need of treatment for narcotics or addictive drug problems	0.59 (0.16-2-20)
Daily use of nicotine (smoking or snuffing)	1.39 (0.63-3.03)
PGSI[b] median score	0.98 (0.91-1.05)
GAS[c] median score	1.27 (1.19-1.37)

[a]K6: Kessler Psychological Distress Scale.

[b]PGSI: Problem Gambling Severity Index.

[c]GAS: Gaming Addiction Scale.

Discussion

Principal Findings

The aim of this study was to determine potential correlates of the use of gambling services within the world of gaming, focusing on sociodemographic variables, co-occurring mental health problems and substance use, and gaming addiction and problem gambling levels. In the study sample, which had a cross-sectional design and was intended to reflect the general population of Sweden, 2.9% (58/1986) of respondents (aged 16 years and older) reported they had engaged in gambling services within computer games in the past year. After controlling the variables for one another, however, significant associations were found between gambling in gaming and male gender, younger age, history of treatment-seeking due to alcohol problems, and higher scores on the GAS. In the unadjusted analyses, individuals involved in gambling within gaming had a markedly higher problem gambling score. However, no independently significant association was found between gambling in gaming and problem gambling when controlling for other relevant variables.

In terms of the prevalence of gambling in gaming, previous research conducted in the area has quite unanimously targeted subgroups of the population such as adolescents or excessive gamers, making our findings one of the first measures of prevalence in the general population. The only comparable figure found was 1.3%, calculated from a recent German sample of internet users [40].

This study examined a series of sociodemographic variables as potential correlates of the involvement in gambling in gaming. Of these, gambling in gaming could be significantly associated with male sex and younger age in the final logistic regression model. More specifically, among the respondents who reported they had engaged in gambling-like activities within computer games the past year, 78% were male and 59% were aged 29 years or younger. Corresponding figures for the compared group were 48% and 17%, respectively. These results were consistent with findings from the limited studies conducted in the area [39,40]. Considering the young male predominance in the context of excessive gambling and gaming in general [23,49], it was not surprising that the distribution of age and sex in the field of convergence between these activities seems to follow the same pattern.

Since poor mental health and substance use problems are commonly co-occurring with problem gambling [3,4], this study intended to investigate whether these associations exist in the case of individuals who engage in gambling in gaming. Li and colleagues [39] had previously indirectly, through the adverse effects of excessive gaming, linked loot box involvement to higher levels of mental distress, yet no significant association was found in our study between gambling in gaming and serious mental illness in the adjusted analysis. Here, one should bear in mind the low sensitivity (0.36) of the K6 scale's ability to detect serious mental illness [45], but on the other hand, a history of treatment-seeking for psychological distress showed no significant association with gambling in gaming in the unadjusted or adjusted analysis, indicating that such a direct association may not exist. Instead, a history of treatment-seeking for alcohol problems was 8 times as common in the group that had engaged in gambling in gaming. To the best of the our knowledge, previous data for comparison is lacking in this regard. Moreover, drinking alcohol while gambling increases the risk of developing gambling problems and vice versa [50], and comorbidity between problem gambling and alcohol use disorders is high [51]. Thus, as traditional problem gamblers share symptom characteristics and to some extent neurobiological and psychological pathophysiology with substance use disorders [52,53], it is worth investigating if this is also the case for the gambling in gaming phenomenon.

Comparing median scores on the GAS between the group that reported they had engaged in gambling in gaming the past year versus the group that had not, significantly higher scores were found in the former group, both in the unadjusted and adjusted analyses. More specifically, respondents who had engaged in gambling in gaming displayed a median score of 21.5 while the group for comparison scored 7, which according to the adopted scale model is the lowest possible total points, indicating absence of gaming-related problems in a very large portion of the general population. This should be read as such criteria largely based on what have been suggested to define internet gaming disorder according to the DSM-5 being found to be fulfilled to a greater extent in the group that had engaged in gambling in gaming. Given that the objective of the study, in this aspect, was to look for associations between gambling in gaming and gaming-related problems in general, continuous scoring was an appropriate approach that would presumptively detect such a link if it existed. At the same time, it follows that the use of the GAS scale as a continuous variable did not claim to detect and address gaming disorder as a psychiatric diagnosis among the respondents, given that this would require demarcations on how many criteria of the 7 items, and to what extent, each respondent met [48]. The positive association found between gambling in gaming and gaming-related problems, however, is concurrent with the discoveries of the few studies conducted in the area, although different screening tools had been used and the monetization type of the games had in these cases been limited to loot boxes [39,40].

In terms of the association between gambling in gaming and problem gambling, previous research has been fairly unanimous that such a link seems to exist [37,38]. Here, however, while the uncontrolled analyses showed a statistically significant and large difference in the PGSI median scores (8.0 vs 0.0) between the respondents with and without engagement in gambling in gaming, these figures did not retain their significance in the final linear regression analysis model. A likely explanation may be that problematic gambling behavior was more strongly explained by elevated GAS scores than PGSI scores, with reference to the previously cited, indirect relationship between daily gaming and problem gambling [23]. In the context of multidimensional gambling behaviors, however, all other forms of gambling participation in the past year tested for in the unadjusted analysis were more common among respondents who had also engaged in gambling in gaming. Wardle [38] concluded in a recent study that while skin betting as an isolated phenomenon was not significantly associated with at-risk gambling, in combination with other gambling activities it was highlighted as a considerable risk factor for developing problematic gambling behaviors by underpinning the same behaviors. Thus, it is theoretically possible that such multifactorial explanation models exist, which prompts for more research in the area.

Implications

This study strengthens the limited scientific basis for the associations previously found between participation in gambling-like activities within computer games and male gender, younger age, and problematic gaming behaviors. However, the demonstrated association between gambling in gaming and a history of treatment-seeking for alcohol problems can be considered a new supplement in the emerging description of the profile characteristics of individuals who engage in this overlapping activity between gambling and gaming. Additionally, this newfound link may provide hints about underlying dependence mechanisms or risk factors in this field of fusion between gambling and gaming, prompting for closer investigation of the associations with substance use disorders as well as possible bidirectional relationships.

As no significant associations could be found between gambling in gaming and serious mental illness or problem gambling, which were findings that to varying degrees differed from existing data, these relationships must be explored further. Since excessive gaming and in-game gambling services are predominantly a problem in the younger population, longitudinal study designs are needed in order to be able to monitor whether involvement in and normalization of such activities at a young age could predispose to gambling problems once these individuals reach legal age. Likewise, given that gambling in gaming is most prevalent in younger segments, this is an urgent aspect that needs to be further investigated for the basis of prevention work.

Strengths and Limitations

The main strength of this study, given its exploratory intentions, is that it is based on a broad study sample with participants recruited from the general population rather than from psychiatric clinical settings where this kind of screening for correlates could be deceptive.

The limitations of this study largely depend on data acquisition–related aspects, given that the collected material consisted of self-estimates and self-responses, and it is a likely assumption that web panel members with a greater interest in or experience of with gambling, mental behaviors, and substance use may have been more prone to participate in the study after the topic was presented in the invitation mailings. Whether online behaviors of individuals who are part of a web panel and recurrently enroll in online market surveys differ from the general population cannot be ruled out. These settings may have contributed to a selection bias, skewing the studied population toward a higher degree of problematic behaviors in several respects that the study intended to investigate. In the light of the findings, however, these circumstances are most important to bear in mind when drawing conclusions about associations between gambling and alcohol issues. On the other hand, the self-administered online questionnaires have the advantage of decreasing the impact of social desirability bias, as respondents were anonymized toward the authors and final scores and potential study implications were withheld.

Another possible limitation of the study that may have impacted the results is that the study did not include individuals younger than age 16 years, a group where, considering the current state of knowledge, involvement in gambling services within computer games can be assumed to be high. This means that the estimated 3% prevalence of past-year engagement in gambling in gaming could be on the low side and should consequently not be taken for granted or applied to other populations, especially because online habits may differ considerably between different countries. However, it is also possible that the age bracket comprising those aged 16-19 years may have encompassed a minor number of participants aged 16 or 17 years, whose hypothetical impact on the results of the study is somewhat difficult to eradicate since the items of the PGSI instrument in many ways undertake legal age and responsibility for one's own finances.

Despite the relatively high rates of engagement in gambling in gaming, the absolute number of respondents belonging to this subgroup was low (n=58) and all of the intended variables could not be included in the overall logistic regression analysis model. Although this can be considered a limiting factor of this study, our findings may serve as guidance for future studies when sample sizes are to be determined in order to allow for more elaborate and multifactorial statistical procedures.

Finally, although it was not the aim of the study, it is noteworthy that this study design does not allow for the determination of temporality aspects between gambling in gaming and the associations found. In order to establish such time relationships and ultimately understand underlying pathological pathways, further research of longitudinal and prospective approaches is required.

Conclusions

The demonstrated findings of this study have strengthened previously found associations between gambling in gaming and younger age, male sex, and problematic gaming behaviors. Additionally, an association was found between engagement in gambling services within computer games and a history of treatment-seeking due to alcohol problems, which should be considered as an insight that expands the current scientific knowledge about how individuals involved with this activity differ from their counterparts in the general population in terms of individual characteristics and comorbidities.

Acknowledgments

The authors would like to thank the companies Patient Information Broker AB (Landskrona, Sweden) and I-Mind Consulting AB (Lund, Sweden) for their help with the data collection and MD Vincent Henzel for help with study preparations and data collection.

Authors' Contributions

MK and AH planned the study together. AH was mainly responsible for ethics permission and data collection. MK performed the statistical analyses and wrote the draft of the paper. Both authors edited the paper and approved the final version.

Conflicts of Interest

AH has a position at Lund University, Lund, Sweden, which is supported by the state-owned gambling operator, AB Svenska Spel. AH has research funding from the research council of Svenska Spel for specific research projects and from the research council of the Swedish alcohol monopoly, Systembolaget. AH is also involved in preparations for a clinical research project in collaboration with Kontigo Care that is planned to provide clinical follow-up devices in clinical treatment of gambling disorder as part of the study but without any other study funding or personal fees. None of the entities mentioned above were involved in or had any influence on this research project, which also was conducted without any specific funding.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington: American Psychiatric Publishing; 2013.
2. Calado F, Griffiths MD. Problem gambling worldwide: an update and systematic review of empirical research (2000-2015). J Behav Addict 2016 Dec;5(4):592-613 [FREE Full text] [doi: 10.1556/2006.5.2016.073] [Medline: 27784180]

3. Buth S, Wurst FM, Thon N, Lahusen H, Kalke J. Comparative analysis of potential risk factors for at-risk gambling, problem gambling and gambling disorder among current gamblers-results of the Austrian Representative Survey 2015. Front Psychol 2017 Dec 14;8:2188 [FREE Full text] [doi: 10.3389/fpsyg.2017.02188] [Medline: 29312056]
4. Dowling N, Merkouris S, Greenwood C, Oldenhof E, Toumbourou J, Youssef G. Early risk and protective factors for problem gambling: a systematic review and meta-analysis of longitudinal studies. Clin Psychol Rev 2017 Feb;51:109-124 [FREE Full text] [doi: 10.1016/j.cpr.2016.10.008] [Medline: 27855334]
5. Goslar M, Leibetseder M, Muench HM, Hofmann SG, Laireiter A. Pharmacological treatments for disordered gambling: a meta-analysis. J Gambl Stud 2019 Jun 20;35(2):415-445 [FREE Full text] [doi: 10.1007/s10899-018-09815-y] [Medline: 30570700]
6. Abbott M. The epidemiology and impact of gambling disorder and other gambling-related harm. Geneva: World Health Organization; 2017. URL: https://www.who.int/docs/default-source/substance-use/the-epidemiology-and-impact-of-gambling-disorder-and-other-gambling-relate-harm.pdf [accessed 2022-03-15]
7. Langham E, Thorne H, Browne M, Donaldson P, Rose J, Rockloff M. Understanding gambling related harm: a proposed definition, conceptual framework, and taxonomy of harms. BMC Public Health 2016 Jan 27;16(1):80 [FREE Full text] [doi: 10.1186/s12889-016-2747-0] [Medline: 26818137]
8. Dowling NA, Cowlishaw S, Jackson AC, Merkouris SS, Francis KL, Christensen DR. The prevalence of comorbid personality disorders in treatment-seeking problem gamblers: a systematic review and meta-analysis. J Pers Disord 2015 Dec;29(6):735-754. [doi: 10.1521/pedi_2014_28_168] [Medline: 25248010]
9. Lorains F, Cowlishaw S, Thomas S. Prevalence of comorbid disorders in problem and pathological gambling: systematic review and meta-analysis of population surveys. Addiction 2011 Mar;106(3):490-498. [doi: 10.1111/j.1360-0443.2010.03300.x] [Medline: 21210880]
10. Moghaddam JF, Yoon G, Dickerson DL, Kim SW, Westermeyer J. Suicidal ideation and suicide attempts in five groups with different severities of gambling: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. Am J Addict 2015 Mar 24;24(4):292-298. [doi: 10.1111/ajad.12197]
11. Karlsson A, Håkansson A. Gambling disorder, increased mortality, suicidality, and associated comorbidity: a longitudinal nationwide register study. J Behav Addict 2018 Dec 01;7(4):1091-1099 [FREE Full text] [doi: 10.1556/2006.7.2018.112] [Medline: 30427214]
12. Wardle H, Moody J, Spence S, Orford A, Volberg R. British Gambling Prevalence Survey 2010. National Centre for Social Research. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/243515/9780108509636.pdf [accessed 2020-02-16]
13. Meyer G, Fiebig M, Häfeli J, Mörsen C. Development of an assessment tool to evaluate the risk potential of different gambling types. Int Gambl Stud 2011 Aug;11(2):221-236. [doi: 10.1080/14459795.2011.584890]
14. Gambling participation in 2018: behaviour, awareness and attitudes: annual report. Gambling Commission. 2019. URL: https://www.gamblingcommission.gov.uk/PDF/survey-data/Gambling-participation-in-2018-behaviour-awareness-and-attitudes.pdf [accessed 2022-02-16]
15. Gainsbury SM. Online gambling addiction: the relationship between internet gambling and disordered gambling. Curr Addict Rep 2015 Jun 01;2(2):185-193. [doi: 10.1007/s40429-015-0057-8]
16. Gainsbury SM, Russell A, Hing N, Wood R, Lubman D, Blaszczynski A. How the Internet is changing gambling: findings from an Australian Prevalence Survey. J Gambl Stud 2015 Mar 11;31(1):1-15 [FREE Full text] [doi: 10.1007/s10899-013-9404-7] [Medline: 23934369]
17. Håkansson A, Mårdhed E, Zaar M. Who seeks treatment when medicine opens the door to pathological gambling patients-psychiatric comorbidity and heavy predominance of online gambling. Front Psychiatry 2017 Nov 29;8:255 [FREE Full text] [doi: 10.3389/fpsyt.2017.00255] [Medline: 29238309]
18. Welte JW, Barnes GM, Wieczorek WF, Tidwell MO, Hoffman JH. Type of gambling and availability as risk factors for problem gambling: a tobit regression analysis by age and gender. Int Gambl Stud 2007 Aug;7(2):183-198. [doi: 10.1080/14459790701387543]
19. Gainsbury SM, Angus DJ, Blaszczynski A. Isolating the impact of specific gambling activities and modes on problem gambling and psychological distress in internet gamblers. BMC Public Health 2019 Oct 25;19(1):1372 [FREE Full text] [doi: 10.1186/s12889-019-7738-5] [Medline: 31653242]
20. France A. Towards a sociological understanding of youth and their risk-taking. J Youth Stud 2010 Aug 19;3(3):317-331. [doi: 10.1080/713684380]
21. Brunborg GS, Mentzoni RA, Frøyland LR. Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking, or conduct problems? J Behav Addict 2014 Mar;3(1):27-32. [doi: 10.1556/jba.3.2014.002]
22. Gaming disorder. Geneva: World Health Organization; 2018. URL: https://www.who.int/features/qa/gaming-disorder/en/ [accessed 2020-02-13]
23. Public Health Agency of Sweden (Folkhälsomyndigheten). Samband mellan dataspel och spel om pengar. 2018. URL: https://www.folkhalsomyndigheten.se/spelprevention/om-spelproblem/dataspel/ [accessed 2020-02-18]

24. Kaess M, Parzer P, Brunner R, Koenig J, Durkee T, Carli V, et al. Pathological internet use is on the rise among European adolescents. J Adolesc Health 2016 Aug;59(2):236-239. [doi: 10.1016/j.jadohealth.2016.04.009] [Medline: 27267140]
25. Müller KW, Janikian M, Dreier M, Wölfling K, Beutel ME, Tzavara C, et al. Regular gaming behavior and internet gaming disorder in European adolescents: results from a cross-national representative survey of prevalence, predictors, and psychopathological correlates. Eur Child Adolesc Psychiatry 2015 May 5;24(5):565-574. [doi: 10.1007/s00787-014-0611-2] [Medline: 25189795]
26. Mihara S, Higuchi S. Cross-sectional and longitudinal epidemiological studies of Internet gaming disorder: a systematic review of the literature. Psychiatry Clin Neurosci 2017 Jul 31;71(7):425-444 [FREE Full text] [doi: 10.1111/pcn.12532] [Medline: 28436212]
27. Paulus FW, Ohmann S, von Gontard A, Popow C. Internet gaming disorder in children and adolescents: a systematic review. Dev Med Child Neurol 2018 Jul 06;60(7):645-659 [FREE Full text] [doi: 10.1111/dmcn.13754] [Medline: 29633243]
28. Lemmens JS, Valkenburg PM, Peter J. Psychosocial causes and consequences of pathological gaming. Comput Human Behav 2011 Jan;27(1):144-152. [doi: 10.1016/j.chb.2010.07.015]
29. 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 Feb;127(2):e319-e329. [doi: 10.1542/peds.2010-1353] [Medline: 21242221]
30. Drummond A, Sauer JD. Video game loot boxes are psychologically akin to gambling. Nat Hum Behav 2018 Aug 18;2(8):530-532. [doi: 10.1038/s41562-018-0360-1] [Medline: 31209310]
31. Skin gambling: teenage Britain's secret habit. ParentZone. 2018. URL: https://parentzone.org.uk/system/files/attachments/Skin_Gambling_Report_June_2018.pdf [accessed 2020-02-19]
32. Zendle D, Cairns P. Loot boxes are again linked to problem gambling: results of a replication study. PLoS ONE 2019 Mar 7;14(3):e0213194 [FREE Full text] [doi: 10.1371/journal.pone.0213194] [Medline: 30845155]
33. King D, Delfabbro P, Griffiths M. The convergence of gambling and digital media: implications for gambling in young people. J Gambl Stud 2010 Jun 16;26(2):175-187. [doi: 10.1007/s10899-009-9153-9] [Medline: 19756979]
34. Gainsbury SM. Gaming-gambling convergence: research, regulation, and reactions. Gaming Law Rev 2019 Mar;23(2):80-83. [doi: 10.1089/glr2.2019.2323]
35. Activision Blizzard announces fourth-quarter and 2017 financial results. Activision Blizzard, Inc. 2018 Feb 08. URL: http://investor.activision.com/releasedetail.cfm?releaseid=1056935 [accessed 2020-02-19]
36. King DL, Delfabbro PH. Predatory monetization schemes in video games (e.g. 'loot boxes') and internet gaming disorder. Addiction 2018 Nov 28;113(11):1967-1969. [doi: 10.1111/add.14286] [Medline: 29952052]
37. 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]
38. Wardle H. The same or different? Convergence of skin gambling and other gambling among children. J Gambl Stud 2019 Dec 9;35(4):1109-1125 [FREE Full text] [doi: 10.1007/s10899-019-09840-5] [Medline: 30852730]
39. 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]
40. von Meduna M, Steinmetz F, Ante L, Reynolds J, Fiedler I. Loot boxes: a game changer?. Hamburg: Gambling Research Division, University of Hamburg; 2019. URL: https://www.researchgate.net/publication/331935977_Loot_Boxes_-_A_Game_Changer [accessed 2022-03-15]
41. Schwiddessen S, Karius P. Watch your loot boxes! Recent developments and legal assessment in selected key jurisdictions from a gambling law perspective. Interact Entertain Law Rev 2018 Jun 20;1(1):17-43. [doi: 10.4337/ielr.2018.01.02]
42. Lee D. Video game gambling banned in Belgium. BBC News. 2018. URL: https://www.bbc.co.uk/news/technology-43906306 [accessed 2018-06-29]
43. Postrado L. Legal definition saves loot boxes from gambling classification in France. CalvinAyre.com. 2018. URL: https://calvinayre.com/2018/07/05/business/legal-definition-saves-loot-boxes-gambling-classification-france/ [accessed 2018-08-11]
44. Håkansson A, Henzel V. Who chooses to enroll in a new national gambling self-exclusion system? A general population survey in Sweden. Harm Reduct J 2020 Oct 21;17(1):82 [FREE Full text] [doi: 10.1186/s12954-020-00423-x] [Medline: 33087113]
45. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. Arch Gen Psychiatry 2003 Feb 01;60(2):184-189. [doi: 10.1001/archpsyc.60.2.184] [Medline: 12578436]
46. Ferris J, Wynne H. The Canadian Problem Gambling Index: final report. Canadian Consortium for Gambling Research. 2001 Feb 19. URL: https://www.greo.ca/Modules/EvidenceCentre/files/Ferris%20et%20al(2001)The_Canadian_Problem_Gambling_Index.pdf [accessed 2022-03-25]
47. Currie SR, Hodgins DC, Casey DM. Validity of the Problem Gambling Severity Index interpretive categories. J Gambl Stud 2013 Jun 17;29(2):311-327. [doi: 10.1007/s10899-012-9300-6] [Medline: 22426971]
48. 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]
49. Public Health Agency of Sweden (Folkhälsomyndigheten). Results from Swelogs 2018. 2018. URL: https://www.folkhalsomyndigheten.se/globalassets/livsvillkor-levnadsvanor/andts/spel/swelogs/resultat-swelogs-2018-2019.pdf [accessed 2020-04-14]

50. Public Health Agency of Sweden (Folkhälsomyndigheten). Alkohol and gambling. 2020. URL: https://www.folkhalsomyndigheten.se/spelprevention/om-spelproblem/alkohol-och-spel/ [accessed 2020-04-15]
51. Dowling NA, Cowlishaw S, Jackson AC, Merkouris SS, Francis KL, Christensen DR. Prevalence of psychiatric co-morbidity in treatment-seeking problem gamblers: a systematic review and meta-analysis. Aust N Z J Psychiatry 2015 Jun 03;49(6):519-539 [FREE Full text] [doi: 10.1177/0004867415575774] [Medline: 25735959]
52. Potenza MN. The neural bases of cognitive processes in gambling disorder. Trends Cogn Sci 2014 Aug;18(8):429-438 [FREE Full text] [doi: 10.1016/j.tics.2014.03.007] [Medline: 24961632]
53. Conversano C, Marazziti D, Carmassi C, Baldini S, Barnabei G, Dell'Osso L. Pathological gambling: a systematic review of biochemical, neuroimaging, and neuropsychological findings. Harv Rev Psychiatry 2012;20(3):130-148. [doi: 10.3109/10673229.2012.694318] [Medline: 22716504]

Abbreviations

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GAS: Gaming Addiction Scale
K6: Kessler Psychological Distress Scale
PGSI: Problem Gambling Severity Index

Edited by N Zary; submitted 25.03.21; peer-reviewed by C Entwistle; comments to author 16.04.21; revised version received 12.09.21; accepted 24.09.21; published 22.04.22.

Please cite as:
Kisch M, Håkansson A
Associations Between Addictive Behaviors, Individual Characteristics, and the Use of Gambling Services Within the World of Gaming: Cross-sectional Survey Study
JMIR Serious Games 2022;10(2):e29077
URL: https://games.jmir.org/2022/2/e29077
doi:10.2196/29077
PMID:35451974

Original Paper

Video Game Addictive Symptom Level, Use Intensity, and Hedonic Experience: Cross-sectional Questionnaire Study

Bhavneet Walia[1], PhD; Jeeyoon Kim[2*], PhD; Ignatius Ijere[1*], PhD; Shane Sanders[2*], PhD

[1]Department of Public Health, Falk College, Syracuse University, Syracuse, NY, United States

[2]Department of Sport Management, Falk College, Syracuse University, Syracuse, NY, United States

[*]these authors contributed equally

Corresponding Author:
Bhavneet Walia, PhD
Department of Public Health
Falk College
Syracuse University
430B Barclay, White Hall
Syracuse, NY, 13244
United States
Phone: 1 3154433930
Email: bwalia@syr.edu

Abstract

Background: The effects of behavioral addiction to video games has received increasing attention in the literature, given increased use intensity among subgroups of video game players.

Objective: This study seeks to empirically determine the relationship between intensity of video gaming and hedonic experience of the player.

Methods: We conducted a survey of 835 individuals who regularly play video games to determine the relationship between intensity of use and hedonic experience. We divided the sample into quartiles by self-reported video game addictive symptom level (from the Internet Gaming Disorder Scale) and conducted polynomial regressions separately for each quartile.

Results: We found that the higher video game addictive symptom level groups experienced a U-shaped (curvilinear) relationship between hedonic experience and intensity of play, whereas groups with lower video game addictive symptom levels exhibited no such relationship. The coefficients for the highest addictive symptom level group (quartile 4) for marginal effects for hours played per week and hours played per week squared were significantly negative (P=.005) and significantly positive (P=.004), respectively.

Conclusions: Our results are consistent with sensitization and tolerance theories, which suggest that high-symptom groups experience frustration and disappointment until they achieve excessive dopamine release, at which point their hedonic experience is expected to improve with additional play. Conversely, low-symptom groups experience no such fall-and-rise pattern. This result is consistent with the outcome that members of the latter group play the game for the direct experience, such that their hedonic experience is more directly related to events occurring in the game than to the increasingly elusive pursuit of excessive dopamine release. We also find that high-symptom groups spend substantially more time and money to support video game use and are much more likely to engage in video game use at the expense of other important activities, such as working, sleeping, and eating.

(JMIR Serious Games 2022;10(2):e33661) doi:10.2196/33661

KEYWORDS

video game use; addictive behaviors; mental health; video game addiction; videogames; addiction; video games

Introduction

In our current technological era, video games have become pervasive in the United States and many other developed countries. This form of gaming has high prevalence among people of different ages, socioeconomic levels, and cultural environments. The Entertainment Software Association (ESA) reports that "we are living in the golden age of video games, and video game players are thriving" [1]. In 2020, the ESA found that 75% of Americans had a gamer in their household and 64% of American adults played video games. A primary appeal of video games is their ability to provide pleasure or mental stimulation to players [2]. These attributes, along with technological improvements allowing for remote competitive play, have helped to drive the observed high rate of growth in video game participation. In general, video game players can vary widely in terms of intensity and competitive level [3,4]. In 2020, the ESA reported that 80% of video game players said the activity provided mental stimulation, and 79% said they experienced high levels of relaxation [1].

Video games are not only pervasive in many areas but also potent enough to provide hedonic experiences to some players. The pursuit of this hedonic experience has caused many players to make video games an essential routine in their lives. Video games have become a booming component of the entertainment industry; however, some health issues have been associated with them [5]. Studies have identified negative consequences of chronic play, which include social, professional, and educational impairment [6]. Hence, very much like substance use disorders, video game disorder is a "persistent or recurrent behavior pattern of sufficient severity to result in significant impairment in personal, family, social, educational, occupational or other important areas of functioning" [3]. Consequently, in 2018, the World Health Organization included video game disorder as an essential type of mental health disorder [3]. This is in accord with the fifth revision of the Diagnostic and Statistical Manual of Mental Disorders.

Many studies have been conducted on the psychological components of video games; others have concentrated on the "emotional states of the game player" [6] and the "playful-consumption experience of videogame[s]" [7]. From the hedonic perspective, the Greek philosopher Aristippus wrote that a good life is "to experience the maximum amount of pleasure, and that happiness is the totality of one's hedonic moments." From this perspective, the pursuit of hedonic enjoyment though the attainment of pleasure and avoidance of pain improves one's well-being as subjectively assessed based on cognitive and affective evaluation of one's perceived happiness in life [8]. In this sense, the term "subjective well-being" (or alternatively, "hedonic well-being") is widely used in attempts to understand how hedonic enjoyment contributes to well-being. Hedonic enjoyment increases positive affect, decreases negative affect, and heightens life satisfaction [9]. Although construing well-being based on pleasure or happiness provides a useful understanding of self-perceived psychological states, the subjective well-being approach has limitations in that pleasurable experiences are not necessarily optimal in promoting wellness (eg, drug use or alcohol consumption). Experiencing pleasure (or hedonic value) is a key motive for video gaming [10]. When pleasure is the motivation, the pursuit of hedonic value becomes the end goal; the pursuit of pleasure can lead to enhanced happiness and hedonic well-being (ie, the aspect of well-being related to pleasure) but can also lead to perverse elements of learned addictive behavior over time [11]. Griffiths [12] suggests mood modification (ie, "the subjective experience that people report as a consequence of engaging in the particular activity") and tolerance (ie, "the process whereby increasing amounts of the particular activity are required to achieve the former effects") are common components of addiction; such components have been confirmed to be present in the video gaming context [13]. Tolerance indicates that the more one is exposed to a stimulus, the greater becomes the threshold for a rewarding experience [14].

The present study examines tolerance in the context of video games. Previous work has shown that excessive players of video games tend to be more reward dependent (for dopamine) [15]. Previous research has also found that when reward expectations for dopamine are not met, disappointment can be experienced (eg, in those with internet addiction disorder) [16,17]. Sensitization (ie, hypersensitive reaction to a stimulus) is another common response among addicted individuals, according to the literature [18,19]. That is, addicted individuals are more responsive to stimuli that trigger their need to engage in the behavior of interest. Hypersensitive response and stimulation make self-regulatory behaviors more challenging for these individuals, even when the consequences of failed self-regulation are not consistent with long-term well-being; time-consuming addictive behaviors can contribute to loss of priorities, sleep deprivation, or job loss. Further, research finds that individuals with behavioral addictions act as they do because the activity in question releases excessive levels of dopamine in the brain [20]. The behavior or activity increasingly becomes a vehicle for dopamine release rather than chiefly an activity in and of itself. The combination of tolerance and sensitization can create complementary difficulties for addicted individuals. Sensitization makes individuals more responsive to stimuli that trigger a behavior, whereas tolerance increases the intensity of the activity needed to achieve the desired release of dopamine [12,13]. The result is something of a trap: addicted individuals are easily triggered to begin the activity but must engage in the activity with increasing intensity to achieve the desired outcome. As the desired outcome relates less to the activity and more to its effects, moreover, individuals may not enjoy the activity until the desired cognitive effect is achieved.

Within the context of these prior research findings, we hypothesized that video game players who show addictive symptoms would experience a U-shaped relationship between hedonic experience in video games and video game hours played per week. If this relationship was present, then we would observe a significant negative association with video games and a significant positive association with video games squared, as in a standard polynomial regression analysis. Sensitization would increase the likelihood that such video game players would initiate play even when it does not improve quality of life (eg, even when the person has other pressing life priorities). As these types of video game players play additional hours on average, we expect that an individual's hedonic response to the game will decrease with play intensity up to the point of excessive release of target levels of dopamine. Once video game play is initiated, we expect such users to be increasingly frustrated until they attain excessive dopamine release, due primarily to tolerance. Upon attainment of this target, however, we expect hedonic response will increase in play intensity. For video game users not experiencing addictive symptoms, we hypothesized that there would be no such U-shaped relationship. Rather, we expected such users to derive hedonic experiences directly from features of the game rather than from chasing a dopamine response. As such, we hypothesized that there would be a comparatively flat relationship between hedonic experience and video game hours played per week for this latter group of video game players.

To test this hypothesis, we divided the study subjects into quartiles based on self-reported addictive video game symptom levels using the Internet Gaming Disorder Scale (IGDS). Quartiles are commonly used in applied statistics to avoid arbitrary or convenient partition points. This method breaks the data into balanced subsamples by definition, because cutoffs are not arbitrarily determined by the researcher. Bennette and Vickers [21] state that “in contemporary epidemiologic practice, continuous variables are typically categorized into tertiles, quartiles and quintiles as a means to illustrate the relationship between a continuous exposure and outcome.” This approach was further validated by Maggiore et al [22], Kjos et al [23], and Sousa et al [24]. Good alternative methods of subsampling exist, such as the method employed by Zhu et al [25] to define categories for leisure gamers, excessive gamers, and pathological gamers based on cutoff values.

In our study, quartile 1 reported the lowest level of symptoms (mean symptom score 1.38) and quartile 4 reported the highest (mean symptom score 3.69). We used separate polynomial regression analyses for each quartile to test for significant nonlinear relationships between hedonic experience and video game hours played. Our use of polynomial regression to test for curvilinear relationships is a key contribution of this study, given that previous regression analyses of this topic have used linear regression. We studied this U-shaped relationship as a potential indicator of sensitization and tolerance in the addicted individuals. We specified a statistical model to explore this and, more specifically, estimated the “tipping point” at which the typical video game player exhibits addictive symptoms, such as attaining a target dopamine response and subsequently facing an increasing hedonic experience given additional video game play.

Methods

Overall Aims

This study aimed to examine how intensity of video game play (ie, casual or heavy usage) moderates the relationship between time spent on playing video games and hedonic experience. First, we hypothesized that subjects reporting higher behavioral addiction symptoms would show a U-shaped relationship between time spent on video games and hedonic experience. Second, we hypothesized that subjects reporting lower behavioral addiction symptoms would not show a U-shaped relationship between time spent on video game and hedonic experience, but instead would show a comparatively flat relationship.

Recruitment

To test the hypotheses, an online survey-based study was conducted after acquiring approval from the institutional review board at Syracuse University (IRB number: 19-186). Based on several screening questions (eg, to determine use and knowledge of video games), subjects who played video games on a regular basis were identified and allowed to participate in the survey study. The target population were United States adults who played video games on a regular basis. Respondents were recruited and the survey was distributed via Amazon Mechanical Turk. The initial request asked them to participate in a survey about well-being. The survey was developed with Qualtrics software (SAP) and was a 6-screen survey. Interested subjects answered questions about their daily lives, including 2 screening questions: (1) Do you play video games on a regular basis? (responses were either yes or no) and (2) How often do you play video games? (multiple responses were available). Subjects who answered “no” to the first question and “not at all” to the second question were automatically excluded. Additionally, respondents who chose video game genres that they played in a multiple-choice question were also asked to list the game titles they had played in that genre. Respondents who listed titles that did not match the genre were excluded.

Data Summary

Of 1072 attempts, 835 (77.8%) participants passed the screening and completed the survey. The average age of the participants was 32 years. More details on the respondents are presented in Table 1. To examine different types of players more specifically, we grouped respondents into quartiles based on game addiction symptoms.

In Lemmens et al’s study [26], video games players who were categorized as “addicted/disordered” constituted less than 5% of video game players. Quartile 4, representing the highest quartile, corresponded closely to the optimal cutoff point for disordered gamers defined by Qin et al [27]. Their optimal cutoff for defining problem gamers was a total 9-item IGDS score of 32, whereas the lower bound in our quartile 4 was 28. Therefore, all video game players that would have been classified as problem gamers in Qin’s definition were included in our quartile 4, along with some additional gamers with scores in the range of 28 to 31, which is near Qin’s cutoff. Given this close correspondence, as well as the empirical precedent of generating quartile categorical variables discussed previously, we maintained our quartile dummies in the regression analysis to follow. We adopted 3 items from the popular self-report scale described by Keyes [28], which measures hedonic experiences, to estimate respondents’ hedonic experience in video games (these items measure interest in life, satisfaction in life, and happiness experienced in video game play on a 6-point Likert scale ranging from “not at all” to “very extremely”); these items have been applied to previous video gaming studies [29]. The Cronbach α of our data was .9. Time spent on video games was measured with a single item: “On average, how many hours per week do you play video games?” (answers were open ended, ranging from 1 to 70 hours per week). The IGDS, defined by Lemmens et al [26], was also employed. This scale assesses 9 criteria for internet gaming disorder: preoccupation, tolerance, withdrawal, persistence, escape, problems, deception, displacement, and conflict. This represents one of the most commonly used gaming addiction symptom scales (responses were on a 5-point Likert scale, ranging from “never” to “very often”); the mean score was employed to group respondents based on behavioral addiction symptoms. The Cronbach α was .91. Additional items to determine the level of neglect of other activities due to video games were also constructed. The first item was “Have you neglected school or work so that you could play games?” Additional items were phrased similarly, asking about neglect of sleeping, eating, socializing with others, and physical activity (responses were on a 5-point Likert scale, ranging from “never” to “very often”). Other items were used to examine money spent on video games (1 item, “How much money do you spend on video games per year?” Answers were open ended); health condition (1 item, “Do you have any health conditions that limit the kind of physical activities you could do?” Answers were yes or no; 84.3% answered “no”), physical condition (1 item, “How physically competent are you in sports and outdoor games?” Answers were on a 4-point Likert scale, ranging from “not at all competent” to “very competent”; the average score was 2.7) and alcohol consumption (1 item, “How many drinks of alcohol do you have per day?” Answer were open ended; the average number was 0.84, ranging from 0 to 15). Demographics were also assessed in the survey, including age, gender, race, education, marital status, and employment. These items are reported in Table 2.

On average, the highest addiction quartile individuals had markedly higher use intensity than any other group. However, some of these individuals exhibited low usage and low hedonic well-being from game play. Though less representative of the highest addiction quartile group, the presence of these individuals represents a potential paradox. Their presence in the sample might be attributable to access issues or the application of self-control, whereby individuals might accept a temporary loss of hedonic well-being in an attempt to break behavioral addictive symptoms. Indeed, at any point in time, a fraction of addicted individuals are motivated to quit [30]. Further, withdrawal symptoms from self-regulation would be consistent with the observations of low use and low hedonic well-being in high behavioral addiction individuals.

Conversely, the lowest addiction quartile individuals, while exhibiting the lowest average use profile, did have some individuals who were high-intensity video game players. This also represents an apparent paradox. However, behavioral addiction involves dependence. That some low addiction quartile individuals reported high-intensity use does not necessarily indicate dependence. For example, it may be that these individuals enjoy gaming and have ample opportunity to game (eg, due to a dearth of life responsibilities) but that they could self-regulate usage if life responsibilities dictated the need to do so. While correlated with each other, intense and problematic use of addictive activities are distinct processes [31].

Table 1. Summary table of mean survey results by behavioral addiction quartile.

Quartile	Addictive symptoms score[a], mean (SD)	Time spent on video games per week (hours)	Activities neglected due to video games[b]					Money spent on video games per year (US$)	Hedonic experience in video games (mean score)[c]
			School or work	Sleeping	Eating	Social activities	Physical activities		
1	1.38 (0.22)	13.96	1.25	1.82	1.27	1.25	1.63	200.52	4.36
2	2.04 (0.19)	17.32	1.72	2.36	1.76	1.74	2.20	236.95	4.27
3	2.74 (0.20)	21.10	2.24	2.84	2.36	2.49	2.74	290.83	4.29
4	3.69 (0.48)	28.13	3.05	3.54	3.27	3.39	3.55	817.74	4.39

[a]Five-point Likert scale (Internet Gaming Disorder Scale).

[b]Five-point Likert scale.

[c]Six-point Likert scale.

Table 2. Survey respondent demographics (N=835).

Characteristics	Value
Gender, n (%)	
Male	525 (62.8)
Female	310 (37.2)
Age, average years	32.4
Race, n (%)	
White	608 (72.8)
Black	111 (13.3)
American Indian/Alaska native	8 (1)
Asian or Pacific islander	70 (8.4)
Other	37 (4.4)
Not applicable	1 (0.1)
Education, n (%)	
Less than high school	4 (0.5)
Highschool	97 (11.6)
Some college	249 (29.8)
College	352 (42.2)
Some graduate school	45 (5.4)
Graduate school	88 (10.5)
Employment, n (%)	
Working	635 (76.1)
Looking for a job	70 (8.4)
Retired	17 (2)
Housewife	34 (4.1)
Student	57 (6.8)
Other	22 (2.6)
Marital status, n (%)	
Married	338 (40.5)
Single	341 (51.6)
Widowed	4 (0.5)
Divorced	35 (4.2)
Separated	12 (1.4)
Married, spouse absent	1 (0.1)
Not applicable	14 (1.7)

Results

The key variables from the survey are summarized in Table 1. We divided the subjects into 4 balanced groups by self-reported video game addictive symptom (quartiles 1 to 4, with quartile 1 exhibiting the lowest addictive symptoms according to the IGDS). We observed that quartile 4, the group self-reporting the highest addictive symptoms, was the most distinct from its neighboring groups in terms of mean IGDS score, video game time expenditure, video game money expenditure, video game hedonic experience, and neglect of school or work, sleeping, eating, social activities, and physical activities. In other words, quartile-4 players were less bounded to the video game play patterns and costs exhibited by their neighbors in the sorted data. At the individual video game player level, we regressed hedonic experience in video games on the average hours of video game play per week, average hours of video game play per week squared, and a set of variables that served to control for any individual heterogeneity among survey respondents when estimating the relationship of interest. We conducted this regression separately for each addictive symptom quartile (quartiles 1 to 4, in ascending order of addictive symptoms).

Video game play per week squared was included to establish the possibility of a nonlinear, quadratic relationship between hedonic experience and video game play intensity (by addictive symptom quartile). Controls included race, education level, primary video game genre, marital status, alcohol consumption, whether the individual played a physical sport, physical condition, and health condition. All interpretations from the model are based on inferential statistical significance testing at standard significance (α) levels, as reported in Table 3. The main (nonparenthetical) values in the table represent regression coefficients or marginal effects. They refer to the change in the dependent variable with respect to a unit change in the explanatory variable. The values in parentheses are 2-sided *t* statistics for the regression coefficients.

The video game players exhibited a very different hedonic experience in weekly hours played across different addiction quartiles. As addictive quartile rose from quartile 1 to quartile 4, the hours played variables moved from an overall insignificant (not significantly different from flat) relationship with hedonic experience in the quartile 1 and quartile 2 regression results to the predicted U-shaped relationship for quartile 3 and quartile 4. In terms of coefficient value, hours played became smaller for quartiles 1 to 4 and more significant for quartiles 2 to 4 (for quartiles 1, 2, 3, and 4, respectively, P=.15, P=.17, P=.06, and P=.01; t_{216}=1.46, t_{146}=–1.365, t_{176}=–1.872, and t_{155}=–2.589) and hours played squared became larger and more significant (for quartiles 1, 2, 3, and 4, respectively, P=.31, P=.11, P=.03, and P=.008; t_{216}=–1.020, t_{146}=1.603, t_{176}=2.131, and t_{155}=2.676) as we moved from quartile 1 to quartile 4. We found that the higher video game addictive symptom level groups experience a U-shaped (ie, curvilinear) relationship between hedonic experience and intensity of play, whereas groups with lower video game addictive symptom levels exhibited no such relationship. The coefficients for the highest addictive symptom level group (ie, quartile 4) representing marginal effects for hours played per week and hours played per week squared, respectively, were significantly negative (P=.005) and significantly positive (P=.004). Figure 1 shows scatter plots; the trend curves illustrate the shift toward a U shape from quartile 1 to quartile 4.

The first (low) addiction quartile sample exhibited a trend curve with an inverted U shape. That is, sampled individuals in quartile 1 experienced first increasing, then decreasing hedonic experience with hours played. This trend curve shows that there were diminishing returns to video game play, a common result in microeconomic theory given standard utility functions. These sampled users reported increasing hedonic experience as they become involved in the challenges of a game but at a decreasing rate, until their hedonic experience reached a maximum and descended. The experience of diminishing returns tends to be associated with behavior moderation, as individuals experience negative reinforcement (declining hedonic experience) beyond a certain number of hours played. This observed negative reinforcement may be both a cause and symptom of low reported addiction levels. Moving to quartile 2, quartile 3, and quartile 4, the trend curve between hedonic experience and video game hours played per week became increasingly U shaped (in terms of both coefficient magnitude and coefficient significance level). For quartile 3 and quartile 4, this U-shaped relationship was significant, as reported previously.

From the regression coefficient output related to hours played per week, we can compute the "tipping point" at which, according to prior findings discussed in the introduction, the typical video game player exhibiting addictive symptoms attains the target dopamine response and subsequently faces increasing hedonic experience given additional video play. For quartile 3 players, the estimated minimum hedonic experience value was 26.25 hours per week. For quartile 4 players, the estimated minimum hedonic experience value was 37 hours per week. These postregression estimates suggest that quartile 4 players experienced a longer period of declining hedonic experience, consistent with tolerance theory (ie, "the process whereby increasing amounts of the particular activity are required to achieve the former effects" [12]). The increasing emergence of a U-shaped relationship for higher addiction quartile players suggests that higher-addiction video game players possess a different motivation and experience in video game play. They do not exhibit diminishing returns in play. Rather, they experience decreasing, then increasing hedonic experience. While seeking excessive dopamine release from video games, as found in the prior research, the hedonic experience for these users drops initially as the user plays more. The previous literature suggests that once the user achieves excessive dopamine release, however, the hedonic experience of quartile 3 and quartile 4 users increases. Such significant increases are observed in the regression results we observed. While low addiction quartile players play the game for its own merits, and thus experience a standard diminishing-returns response, the regression results are consistent with the finding that high addiction quartile users play the game to achieve excessive dopamine release and experience an initial "low" period of frustration followed by a "high" period.

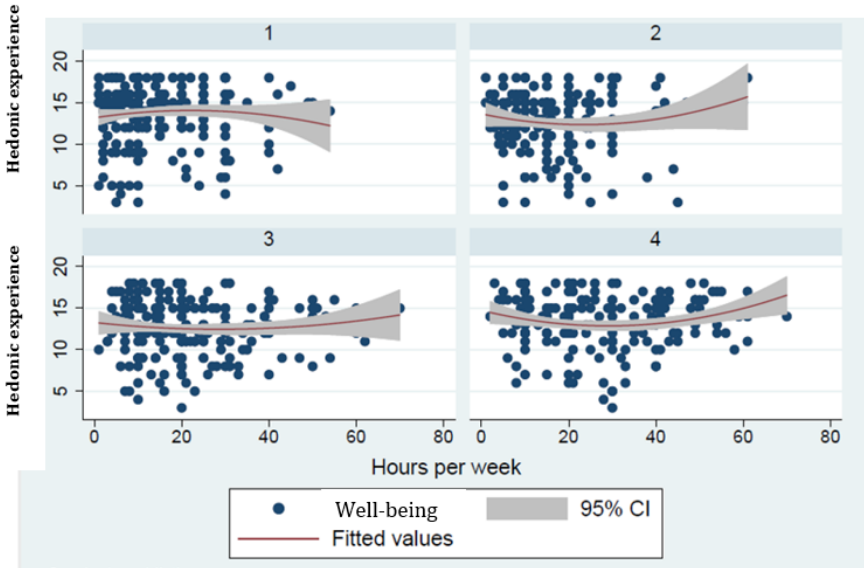
In general, there are two elements of addiction that represent theoretical constructs that are significantly consistent with the observed hedonic experience profiles: sensitization and tolerance. Sensitization indicates a hypersensitive reaction to video game exposure [19] and a heightened threshold of video game exposure to experience pleasure [12]. Tolerance makes it more difficult to experience hedonism and pleasure and achieve excessive dopamine release, driving higher addiction quartile users to more average hours of video game use. Sensitization (or amplification) of hedonic experience makes it more difficult for higher addiction quartile individuals to regulate video game use, even during periods in which it may be difficult or represent a high opportunity cost to play until achieving excessive dopamine release [32].

Table 3. Regression results, Internet Gaming Disorder Scale addiction score versus use intensity and subject controls.

Variable	Quartile 1, r (*t*)	Quartile 2, r (*t*)	Quartile 3, r (*t*)	Quartile 4, r (*t*)
Hours of video game play, linear regression (hours per week)	0.087 (0.146)	–0.096 (0.174)	–0.105 (0.062)	–0.148 (0.011)
Hours of video game play, quadratic regression (hours per week)	–0.001 (0.309)	0.002 (0.111)	0.002 (0.034)	0.002 (0.008)
Primary video game genre				
Multiplayer Online Battle Arena (0 or 1)	Reference	Reference	Reference	Reference
Sports (0 or 1)	–0.989 (0.217)	2.169 (0.034)	0.169 (0.805)	–0.617 (0.280)
First person shooting (0 or 1)	0.082 (0.906)	2.260 (0.009)	1.379 (0.031)	–0.954 (0.178)
Real time strategy (0 or 1)	0.006 (0.994)	–0.659 (0.753)	–0.917 (0.451)	–1.997 (0.113)
Action/adventure (0 or 1)	0.124 (0.855)	2.178 (0.013)	0.718 (0.253)	–1.697 (0.017)
Other (0 or 1)	0.829 (0.280)	2.202 (0.027)	2.558 (0.004)	–0.398 (0.745)
Plays sports (0 or 1)	0.255 (0.611)	–0.133 (0.833)	–0.292 (0.539)	–0.873 (0.576)
Employed (0 or 1)	0.001 (0.995)	0.034 (0.860)	–0.255 (0.105)	–0.068 (0.691)
Race				
White (0 or 1)	0 (0.999)	0 (0.999)	0 (0.999)	0 (0.999)
Black (0 or 1)	–1.434 (0.060)	–1.821 (0.029)	–0.497 (0.446)	0.211 (0.700)
American Indian or Alaska Native (0 or 1)	0.429 (0.890)	2.126 (0.206)	1.666 (0.550)	3.795 (0.191)
Asian or Pacific Islander (0 or 1)	–0.420 (0.651)	–2.571 (0.002)	0.825 (0.257)	–2.220 (0.001)
Other (0 or 1)	–2.198 (0.017)	–0.682 (0.626)	–0.364 (0.697)	–1.141 (0.292)
Not applicable (0 or 1)	N/A[a]	2.553 (0.437)	N/A	N/A
Education				
Less than high school (0 or 1)	0 (0.999)	N/A	0 (0.999)	0 (0.999)
High school (0 or 1)	0.500 (0.823)	0 (0.999)	–2.980 (0.345)	6.142 (0.053)
Some college (0 or 1)	0.116 (0.958)	1.247 (0.150)	–2.722 (0.344)	5.916 (0.056)
College (0 or 1)	0.424 (0.846)	2.022 (0.015)	–3.179 (0.313)	6.075 (0.049)
Some graduate school (0 or 1)	–0.138 (0.954)	0.057 (0.978)	–3.919 (0.240)	6.606 (0.036)
Graduate school (0 or 1)	1.087 (0.627)	3.119 (2.750)	–2.702 (0.397)	6.453 (0.040)
Marital status				
Married (0 or 1)	0 (0.999)	0 (0.999)	0 (0.999)	0 (0.999)
Single (0 or 1)	–.390 (0.001)	–0.380 (0.488)	–1.515 (0.001)	0.012 (0.979)
Widowed (0 or 1)	N/A	–2.834 (0.272)	1.422 (0.633)	0.563 (0.850)
Divorced (0 or 1)	–2.249 (0.016)	2.604 (0.112)	–3.380 (0.001)	–0.249 (0.855)
Separated (0 or 1)	–0.314 (0.886)	–1.279 (0.451)	0.251 (0.889)	–3.512 (0.101)
Married, spouse absent (0 or 1)	N/A	N/A	N/A	0.631 (0.834)
Not applicable (0 or 1)	–2.741 (0.213)	0.001 (0.999)	1.508 (0.294)	–0.131 (0.933)
Health condition (0 or 1, with 1 being no limiting health condition)	1.049 (0.027)	0.244 (0.755)	–0.794 (0.046)	–0.650 (0.051)
Physical condition (4-point scale)	1.449 (0.001)	1.064 (0.002)	0.907 (0.001)	0.461 (0.076)
Alcohol consumption per day (drinks per day)	0.231 (0.102)	–0.401 (0.032)	–0.191 (0.200)	–0.078 (0.168)
Constant	7.117 (0.013)	7.616(0.010)	17.050 (0.001)	11.429 (0.001)
lnsigma2 constant	2.172 (0.001)	2.303 (0.001)	2.122 (0.001)	2.050 (0.001)
Observations	249	179	209	188

[a]N/A: not applicable.

Figure 1. Regression-estimated relationship between hedonic experience and use intensity. Hedonic experience was measured using a 6-point, 3-item Likert scale; thus, total possible hedonic experience was 18 points.



Discussion

The results confirm our two main hypotheses. Namely, for video game players reporting higher behavioral addiction symptoms, time spent on video game play has a significant U-shaped relationship with hedonic experience. For those reporting lower behavioral addiction symptoms, time spent on video game play does not have a significant U-shaped relationship with hedonic experience, but rather a comparatively flat relationship. Hedonic experience is a key motive and consequence for video gaming [8] that can bring subjective happiness and pleasure when playing video games [9]. However, excessive video game play can increase the amount of video game play required to satisfy hedonic needs (ie, tolerance) [10] and lead to development of a hypersensitive reaction to video games (ie, sensitization) [17], affecting hedonic experience in video games and dependency on video games. We confirmed a U-shaped relationship between time spent on video games and hedonic experience among those reporting a higher tendency toward addictive symptoms; we did not see such a relationship for those with a lower tendency. Within the present context, this study is the first to empirically test this U-shaped relationship. We also computed the "tipping point" at which the typical video game user exhibiting addictive symptoms transitions from decreasing to increasing hedonic experience given additional video game play, consistent with what the literature has identified as achievement of target dopamine response. These postregression estimates suggest that higher addiction quartile players experience a longer period of declining hedonic experience, consistent with tolerance theory [10].

The first and second quartile respondents exhibited no substantial addictive symptoms. They are at low risk of developing video game addiction. They exhibited a more typical diminishing-returns response. They are social video gamers or experimental video gamers with limited video game addictive symptoms (excessive time spent, loss of money, and inability to pay attention to their daily responsibilities). Although they may exhibit some video game addictive symptoms, they do not meet the criteria for video game addictive disorder. However, King, Herd, and Delfabbro [33] state that "there is a consistent finding in these types of studies that normal and problem gamers both endorse many of the same motivations for gaming, with problem users simply tending to score much higher than casual users. Thus, the boundary between normal and maladaptive gaming motivations is not always clearly demarcated."

Nonetheless, results in the third and fourth quartiles show video gamers who display more addictive symptoms (ie, high amount of time spent, high level of tolerance, withdrawal, craving for the behavior, and negative impacts on family, social, and occupational responsibilities). They are frequent players, problematic players, or at-risk players. Their hedonic experience level drops at first (they show disappointment, frustration, and possibly display depressive symptoms), then rises (increasing returns to play). With high tolerance to video game exposure, they are frustrated due to their inability to attain the experience requirements of the game. This behavior is consistent with a study conducted by Kaptsis et al [34], in which they found that problem players exhibited withdrawal symptoms when they did not experience a certain number of requirements of the game. The experiencing of withdrawal symptoms explains the low position of the curve at low usage points for third and fourth quartile players. This indicates negative reactions when deprived of video game exposure. These discomforting withdrawal symptoms include irritability, depressive symptoms, and anxiety. The discomfort of withdrawal symptoms notably induces craving symptoms, such as the need to spend more time playing and the fear of missing specific gaming experiences.

The display of craving symptoms is consistent with a study conducted by Przybylski et al [35]. In this study, the researchers highlighted some psychological components of craving, such as fear of missing social play, novel gaming experiences, and gaming for escape or relaxation as motivating factors for the extensive use of time in gaming. These factors help to explain the rise in hedonic experience among problematic gamers in the third and fourth quartiles. According to King et al [33], this view of craving may explain why problem video game players engage in prolonged, intense, repetitive, or tedious gaming activities. This is consistent with results from our study showing that players with a higher tendency toward addictive symptoms also show higher levels of neglect of other daily activities, such as working or studying, eating, sleeping, and socializing. It is also consistent with multiple previous studies that have reported that problematic video gamers play video games for a prolonged period, skip school or work, experience problems with sleep, and have lower grades at school [36-40]. These symptoms have been observed across different age groups of video game players and other populations [41,42].

Conclusion

In this study, we conducted a survey of 835 individuals who regularly play video games to determine the relationship between intensity of video game play and hedonic experience of the player. We divided the sample into quartiles by self-reported video game addictive symptom level (using the IGDS) and conducted polynomial regressions separately for each quartile. We found that the higher video game addictive symptom level groups experienced a U-shaped (ie, curvilinear) relationship between hedonic experience and intensity of play, whereas groups with lower video game addictive symptom levels exhibited no such relationship. These results are consistent with sensitization and tolerance theories, which suggest that high-symptom groups are expected to experience frustration and disappointment until achieving excessive dopamine release, at which point their hedonic experience is expected to improve with additional play. Conversely, low-symptom groups experience no such fall-and-rise pattern. This result is consistent with the outcome that members of the latter group play the game for the direct experience, such that their hedonic experience is more directly related to events occurring in the game than to the increasingly elusive pursuit of excessive dopamine release. We also find that high-symptom groups spend substantially more time and money to support video gaming and are much more likely to engage in video gaming at the expense of other important activities, such as working, sleeping, and eating.

Limitations and Future Research

Although this study shows a novel relationship across all genres of video game play, it does not study the relationship categorically by genre of play. Future research based on longitudinal data can provide information on players and their well-being at different points in time, shedding light on microlevel changes in this relationship. Further, this study does not address cognitive or behavioral consequences of video game play [43]. Lastly, the study relies upon voluntary, self-reported data, which is subject to limitations related to honesty, introspective ability, and sampling (or self-selection) bias. Future studies might rely on random sampling to overcome some of these biases.

Acknowledgments

This research was funded by a Syracuse University CUSE grant (2019-2022).

Conflicts of Interest

None declared.

References

1. 2020 Essential Facts About the Computer and Video Game Industry. Entertainment Software Association. URL: https://www.theesa.com/resource/2020-essential-facts/ [accessed 2022-03-03]
2. Granic I, Lobel A, Engels RCME. The benefits of playing video games. Am Psychol 2014 Jan;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
3. Addictive behaviours: gaming disorder 2018. World Health Organization. URL: https://www.who.int/news-room/questions-and-answers/item/addictive-behaviours-gaming-disorder [accessed 2022-03-03]
4. Shi J, Renwick R, Turner NE, Kirsh B. Understanding the lives of problem gamers: The meaning, purpose, and influences of video gaming. Comput Hum Behav 2019 Aug;97:291-303. [doi: 10.1016/j.chb.2019.03.023]
5. Khazaal Y, Chatton A, Rothen S, Achab S, Thorens G, Zullino D, et al. Psychometric properties of the 7-item game addiction scale among french and German speaking adults. BMC Psychiatry 2016 May 10;16(1):132-132 [FREE Full text] [doi: 10.1186/s12888-016-0836-3] [Medline: 27160387]
6. Gros L, Debue N, Lete J, van de Leemput C. Video Game Addiction and Emotional States: Possible Confusion Between Pleasure and Happiness? Front Psychol 2019 Jan 27;10:2894-2894 [FREE Full text] [doi: 10.3389/fpsyg.2019.02894] [Medline: 32047450]
7. Abbasi A, Jamek ABSA. Playful-consumption experience of videogame-play influences consumer video-game engagement: a conceptual model. GBMR 2017;9:244-252 [FREE Full text]
8. Diener E, Oishi S, Richard RE. Subjective well-being: The science of happiness and life satisfaction. In: Lopez SJ, Snyder CR, editors. The Oxford Handbook of Positive Psychology (2 ed.). Oxford, UK: Oxford University Press; 2009:63-73.
9. Diener E, Lucas RE. Personality and subjective well-being. In: Kahneman D, Diener E, Schwarz N, editors. Well-being: Foundations of hedonic psychology. New York, NY: Russell Sage Foundation; 1999:213-213.
10. Li H, Liu Y, Xu X, Heikkilä J, van der Heijden H. Modeling hedonic is continuance through the uses and gratifications theory: An empirical study in online games. Comput Hum Behav 2015 Jul;48:261-272. [doi: 10.1016/j.chb.2015.01.053]

11. Kringelbach ML, Berridge KC. Towards a functional neuroanatomy of pleasure and happiness. Trends Cogn Sci 2009 Nov;13(11):479-487 [FREE Full text] [doi: 10.1016/j.tics.2009.08.006] [Medline: 19782634]
12. Griffiths M. A 'components' model of addiction within a biopsychosocial framework. J Subst Use 2009 Jul 12;10(4):191-197. [doi: 10.1080/14659890500114359]
13. Lemmens JS, Hendriks SJ. Addictive Online Games: Examining the Relationship Between Game Genres and Internet Gaming Disorder. Cyberpsychol Behav Soc Netw 2016 Apr;19(4):270-276. [doi: 10.1089/cyber.2015.0415] [Medline: 26959285]
14. Ahmed SH, Kenny PJ, Koob GF, Markou A. Neurobiological evidence for hedonic allostasis associated with escalating cocaine use. Nat Neurosci 2002 Jul 10;5(7):625-626. [doi: 10.1038/nn872] [Medline: 12055635]
15. Han DH, Lee YS, Yang KC, Kim EY, Lyoo IK, Renshaw PF. Dopamine genes and reward dependence in adolescents with excessive internet video game play. J Addict Med 2007 Sep;1(3):133-138. [doi: 10.1097/ADM.0b013e31811f465f] [Medline: 21768948]
16. Kim SH, Baik SH, Park CS, Kim SJ, Choi SW, Kim SE. Reduced striatal dopamine D2 receptors in people with Internet addiction. Neuroreport 2011 Jun 11;22(8):407-411. [doi: 10.1097/WNR.0b013e328346e16e] [Medline: 21499141]
17. Hou H, Jia S, Hu S, Fan R, Sun W, Sun T, et al. Reduced striatal dopamine transporters in people with internet addiction disorder. J Biomed Biotechnol 2012;2012:854524-854525 [FREE Full text] [doi: 10.1155/2012/854524] [Medline: 22505818]
18. Robinson TE, Berridge KC. Incentive-sensitization and addiction. Addiction 2001 Jan;96(1):103-114. [doi: 10.1046/j.1360-0443.2001.9611038.x] [Medline: 11177523]
19. Weinstein AM. Computer and video game addiction-a comparison between game users and non-game users. Am J Drug Alcohol Abuse 2010 Sep 15;36(5):268-276. [doi: 10.3109/00952990.2010.491879] [Medline: 20545602]
20. Berke JD, Hyman SE. Addiction, dopamine, and the molecular mechanisms of memory. Neuron 2000 Mar;25(3):515-532 [FREE Full text] [doi: 10.1016/s0896-6273(00)81056-9] [Medline: 10774721]
21. Bennette C, Vickers A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. BMC Med Res Methodol 2012 Feb 29;12(1):21-25 [FREE Full text] [doi: 10.1186/1471-2288-12-21] [Medline: 22375553]
22. Maggiore Q, Nigrelli S, Ciccarelli C, Grimaldi C, Rossi GA, Michelassi C. Nutritional and prognostic correlates of bioimpedance indexes in hemodialysis patients. Kidney Int 1996 Dec;50(6):2103-2108 [FREE Full text] [doi: 10.1038/ki.1996.535] [Medline: 8943496]
23. Kjos S, Peters RK, Xiang A, Henry OA, Montoro M, Buchanan TA. Predicting future diabetes in Latino women with gestational diabetes. Utility of early postpartum glucose tolerance testing. Diabetes 1995 May;44(5):586-591. [doi: 10.2337/diab.44.5.586] [Medline: 7729620]
24. Sousa AS, Guerra RS, Fonseca I, Pichel F, Ferreira S, Amaral TF. Financial impact of sarcopenia on hospitalization costs. Eur J Clin Nutr 2016 Sep 11;70(9):1046-1051. [doi: 10.1038/ejcn.2016.73] [Medline: 27167668]
25. Zhu S, Zhuang Y, Lee P, Li JC, Wong PWC. Leisure and Problem Gaming Behaviors Among Children and Adolescents During School Closures Caused by COVID-19 in Hong Kong: Quantitative Cross-sectional Survey Study. JMIR Serious Games 2021 May 07;9(2):e26808 [FREE Full text] [doi: 10.2196/26808] [Medline: 33960954]
26. Lemmens JS, Valkenburg PM, Gentile DA. The Internet Gaming Disorder Scale. Psychol Assess 2015 Jun;27(2):567-582. [doi: 10.1037/pas0000062] [Medline: 25558970]
27. 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]
28. Keyes CLM. Brief Description of the Mental Health Continuum Short Form (MHC-SF). Positive Emotions and Psychophysiology (PEP) Laboratory. 2009. URL: https://peplab.web.unc.edu/wp-content/uploads/sites/18901/2018/11/MHC-SFoverview.pdf [accessed 2021-06-10]
29. Snodgrass JG, Lacy MG, Dengah HF, Polzer ER, Else RJ, Arevalo JM, et al. Positive mental well-being and immune transcriptional profiles in highly involved videogame players. Brain Behav Immun 2019 Nov;82:84-92 [FREE Full text] [doi: 10.1016/j.bbi.2019.07.035] [Medline: 31376495]
30. Brevers D, Bechara A, Kilts CD, Antoniali V, Bruylant A, Verbanck P, et al. Competing Motivations: Proactive Response Inhibition Toward Addiction-Related Stimuli in Quitting-Motivated Individuals. J Gambl Stud 2018 Sep 24;34(3):785-806. [doi: 10.1007/s10899-017-9722-2] [Medline: 29067545]
31. Boer M, van den Eijnden RJ, Boniel-Nissim M, Wong S, Inchley JC, Badura P, et al. Adolescents' Intense and Problematic Social Media Use and Their Well-Being in 29 Countries. J Adolesc Health 2020 Jun;66(6S):S89-S99 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.02.014] [Medline: 32446614]
32. King DL, Herd MCE, Delfabbro PH. Tolerance in Internet gaming disorder: A need for increasing gaming time or something else? J Behav Addict 2017 Dec 01;6(4):525-533 [FREE Full text] [doi: 10.1556/2006.6.2017.072] [Medline: 29137493]
33. Kaptsis D, King DL, Delfabbro PH, Gradisar M. Withdrawal symptoms in internet gaming disorder: A systematic review. Clin Psychol Rev 2016 Feb;43:58-66. [doi: 10.1016/j.cpr.2015.11.006] [Medline: 26704173]
34. Przybylski AK, Murayama K, DeHaan CR, Gladwell V. Motivational, emotional, and behavioral correlates of fear of missing out. Comput Hum Behav 2013 Jul;29(4):1841-1848. [doi: 10.1016/j.chb.2013.02.014]

35. Greitemeyer T, Mügge DO. Video games do affect social outcomes: a meta-analytic review of the effects of violent and prosocial video game play. Pers Soc Psychol Bull 2014 May;40(5):578-589. [doi: 10.1177/0146167213520459] [Medline: 24458215]
36. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep Med Rev 2015 Jun;21:50-58 [FREE Full text] [doi: 10.1016/j.smrv.2014.07.007] [Medline: 25193149]
37. 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]
38. Mei S, Yau YH, Chai J, Guo J, Potenza MN. Problematic Internet use, well-being, self-esteem and self-control: Data from a high-school survey in China. Addict Behav 2016 Oct;61:74-79 [FREE Full text] [doi: 10.1016/j.addbeh.2016.05.009] [Medline: 27249805]
39. Saquib N, Saquib J, Wahid A, Ahmed AA, Dhuhayr HE, Zaghloul MS, et al. Video game addiction and psychological distress among expatriate adolescents in Saudi Arabia. Addict Behav Rep 2017 Dec;6:112-117 [FREE Full text] [doi: 10.1016/j.abrep.2017.09.003] [Medline: 29450245]
40. Wartberg L, Zieglmeier M, Kammerl R. Accordance of Adolescent and Parental Ratings of Internet Gaming Disorder and Their Associations with Psychosocial Aspects. Cyberpsychol Behav Soc Netw 2019 Apr;22(4):264-270. [doi: 10.1089/cyber.2018.0456] [Medline: 30801222]
41. Beard CL, Haas AL, Wickham RE, Stavropoulos V. Age of Initiation and Internet Gaming Disorder: The Role of Self-Esteem. Cyberpsychol Behav Soc Netw 2017 Jun;20(6):397-401. [doi: 10.1089/cyber.2017.0011] [Medline: 28622030]
42. Barrett SP, Boileau I, Okker J, Pihl RO, Dagher A. The hedonic response to cigarette smoking is proportional to dopamine release in the human striatum as measured by positron emission tomography and [11C]raclopride. Synapse 2004 Nov;54(2):65-71. [doi: 10.1002/syn.20066] [Medline: 15352131]
43. Choi E, Shin S, Ryu J, Jung K, Hyun Y, Kim J, et al. Association of Extensive Video Gaming and Cognitive Function Changes in Brain-Imaging Studies of Pro Gamers and Individuals With Gaming Disorder: Systematic Literature Review. JMIR Serious Games 2021 Jul 09;9(3):e25793-e25793 [FREE Full text] [doi: 10.2196/25793] [Medline: 34255648]

Abbreviations

IGDS: Internet Gaming Disorder Scale
ESA: Entertainment Software Association

Edited by N Zary; submitted 17.09.21; peer-reviewed by S Zhu, T Baranowski; comments to author 06.12.21; revised version received 25.01.22; accepted 22.04.22; published 09.06.22.

Please cite as:
Walia B, Kim J, Ijere I, Sanders S
Video Game Addictive Symptom Level, Use Intensity, and Hedonic Experience: Cross-sectional Questionnaire Study
JMIR Serious Games 2022;10(2):e33661
URL: https://games.jmir.org/2022/2/e33661
doi:10.2196/33661
PMID:35471995

Review

Outcomes, Measurement Instruments, and Their Validity Evidence in Randomized Controlled Trials on Virtual, Augmented, and Mixed Reality in Undergraduate Medical Education: Systematic Mapping Review

Lorainne Tudor Car[1,2], MD, MSc, PhD; Bhone Myint Kyaw[3], MBBS, MSc, PhD; Andrew Teo[1], PhD; Tatiana Erlikh Fox[3,4], MD, MSc; Sunitha Vimalesvaran[1], MD; Christian Apfelbacher[5,6], PhD; Sandra Kemp[7], PhD; Niels Chavannes[8], PhD

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[2]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom
[3]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[4]Department of Internal Medicine, Onze Lieve Vrouwen Gasthuis, Amsterdam, Netherlands
[5]Institute of Social Medicine and Health Systems Research, Otto von Guericke University Magdeburg, Magdegurg, Germany
[6]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore
[7]Faculty of Health Sciences, Curtin Medical School, Curtin University, Bentley, Australia
[8]Department of Public Health and Primary Care, Leiden University Medical Centre, Leiden, Netherlands

Corresponding Author:
Lorainne Tudor Car, MD, MSc, PhD
Lee Kong Chian School of Medicine
Nanyang Technological University
Clinical Sciences Building
11 Mandalay Road
Singapore, 308232
Singapore
Phone: 65 69041258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Extended reality, which encompasses virtual reality (VR), augmented reality (AR), and mixed reality (MR), is increasingly used in medical education. Studies assessing the effectiveness of these new educational modalities should measure relevant outcomes using outcome measurement tools with validity evidence.

Objective: Our aim is to determine the choice of outcomes, measurement instruments, and the use of measurement instruments with validity evidence in randomized controlled trials (RCTs) on the effectiveness of VR, AR, and MR in medical student education.

Methods: We conducted a systematic mapping review. We searched 7 major bibliographic databases from January 1990 to April 2020, and 2 reviewers screened the citations and extracted data independently from the included studies. We report our findings in line with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Results: Of the 126 retrieved RCTs, 115 (91.3%) were on VR and 11 (8.7%) were on AR. No RCT on MR in medical student education was found. Of the 115 studies on VR, 64 (55.6%) were on VR simulators, 30 (26.1%) on screen-based VR, 9 (7.8%) on VR patient simulations, and 12 (10.4%) on VR serious games. Most studies reported only a single outcome and immediate postintervention assessment data. Skills outcome was the most common outcome reported in studies on VR simulators (97%), VR patient simulations (100%), and AR (73%). Knowledge was the most common outcome reported in studies on screen-based VR (80%) and VR serious games (58%). Less common outcomes included participants' attitudes, satisfaction, cognitive or mental load, learning efficacy, engagement or self-efficacy beliefs, emotional state, competency developed, and patient outcomes. At least one form of validity evidence was found in approximately half of the studies on VR simulators (55%), VR patient simulations (56%), VR serious games (58%), and AR (55%) and in a quarter of the studies on screen-based VR (27%). Most studies used assessment methods that were implemented in a nondigital format, such as paper-based written exercises or in-person assessments where examiners observed performance (72%).

Conclusions: RCTs on VR and AR in medical education report a restricted range of outcomes, mostly skills and knowledge. The studies largely report immediate postintervention outcome data and use assessment methods that are in a nondigital format. Future RCTs should include a broader set of outcomes, report on the validity evidence of the measurement instruments used, and explore the use of assessments that are implemented digitally.

(JMIR Serious Games 2022;10(2):e29594) doi:10.2196/29594

KEYWORDS

virtual reality; augmented reality; mixed reality; outcomes; extended reality; digital education; randomized controlled trials; medical education; measurement instruments

Introduction

Background

Extended reality (ER) encompasses immersive technologies within the reality-virtuality continuum, such as virtual reality (VR), augmented reality (AR), and mixed reality (MR). The use of ER technologies is becoming more common in medical education. These technologies offer a wide range of educational opportunities within different medical specialties. VR is a technology that renders a fully computer-generated 3D multimedia environment in real time. It supports a first-person active-learning experience through immersion, that is, a perception of the digital world as real. VR can be integrated with other educational approaches such as virtual patients or serious games. VR patient simulations are interactive computer simulations of real-life clinical scenarios for the purpose of medical education. VR serious games incorporate gaming concepts such as different levels of difficulties, rewards, or feedback within the computer-generated 3D environment.

AR is a technology in which the real-world environment is enhanced by computer-generated virtual imagery information. In AR, virtual objects are projected over the real-world environment. MR is a hybrid technology that merges the features of VR and AR. In MR, virtual objects become a part of the real word. ER technologies can be displayed through desktop computers, mobile devices, and large screens or projected on the walls. They can be purely screen based or also involve the use of joysticks, probes, gloves, simulators, and other forms of haptic devices.

Effectiveness of VR

Our systematic review on the effectiveness of VR for health professions education showed that VR may improve postintervention knowledge and skills outcomes compared with traditional education (ie, nondigital education) or other types of digital education such as online or offline digital education [1]. Data for other outcomes were limited. Systematic reviews of randomized controlled trials (RCTs) remain the gold standard for evidence on the effectiveness of interventions. However, the heterogeneity of participants, interventions, comparison interventions, and outcomes reported in the individual studies can limit the trustworthiness of the systematic review findings and preclude a meta-analysis. Similarly, differences in measurement instruments and types of validity evidence can lead to unreliable conclusions [2]. The choice of digital education outcomes can be influenced by different factors, including types of digital education, the curriculum, and the field of study [3,4]. The process of measuring digital education outcomes can be achieved with a wide variety of measurement instruments, including multiple-choice questions, structured essays, and structured direct observations with checklists for ratings [5]. Measurement instruments used in research need to have validity evidence. Validity is defined as “the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests” [6]. Validity evidence for measurement instruments is important to ensure that the instruments reliably measure what they purport to measure and to support the interpretation of assessment data. However, reporting of validity evidence of measurement instruments in health professions education literature is still suboptimal, ranging from 34.6% in studies on continuing medical education to 64% in studies on technology-enhanced health professions simulation training [7,8].

The use of measurement instruments without validity evidence severely undermines the credibility of the research results [9]. ER is increasingly used in medical education, and studies in this field should evaluate diverse outcomes using outcome measurement instruments with validity evidence. Our aim is to support this by mapping the current choice of outcomes, measurement instruments, and the prevalence of measurement instruments with validity evidence in RCTs on the use of ER in undergraduate and preregistration medical education.

Methods

Methodology, Definitions, and Eligibility Criteria

We performed this systematic review in line with the Cochrane gold standard systematic review methodology and report it according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standards of quality for reporting systematic reviews [10,11]. In this review, we aim to answer the following research questions:

1. Which outcomes (eg, knowledge, skills, attitudes, and behavior) are assessed and reported in RCTs on the effectiveness of VR, AR, and MR in undergraduate and preregistration medical education?
2. What type of measurement instruments were used in RCTs on the use of VR, AR, and MR in undergraduate and preregistration medical education?

3. What proportion of RCTs on the use of VR, AR, and MR in undergraduate medical education report validity evidence for the measurement instruments used, and how was the evidence reported?

We included studies meeting the following eligibility criteria:

1. RCTs
2. Studies on students participating in preregistration or undergraduate medical education in any geographical or educational setting
3. Studies evaluating any type of blended (ie, a combination of extended and nondigital, traditional education) or full ER technology, including VR, AR, and MR
4. Studies comparing VR with control interventions such as classroom-based learning, no intervention, and other types of digital and blended education

We defined different ER technologies as per Textbox 1. Preregistration or undergraduate medical education was defined in line with the World Health Organization (WHO) definition as "any type of initial study leading to a qualification that (i) is recognized by the relevant governmental or professional bodies of the country where the study was conducted and (ii) enables its holder primary entry into the healthcare workforce" [12]. Studies were excluded if they focused on traditional and complementary medicine as defined by WHO (as such education is not included in most medical schools) and used study designs other than an RCT [13].

Textbox 1. Descriptions and classification of different types of virtual reality (VR), augmented reality (AR), and mixed reality (MR).

Types of extended reality modalities in medical education

- VR is a technology that allows the user to explore and manipulate computer-generated 2D or 3D, multimedia sensory environments in real time [14]. The VR environment is the computer-generated representation of a real or artificial environment that can be interacted with by external involvement, allowing for a first-person active-learning experience through immersion [15].
- Screen-based VR interventions are computer-based 3D software applications delivered either through computer screens or head-mounted displays (ie, VR headsets). This type of VR in medical education mostly includes 3D models of organs and VR worlds.
- VR simulators or psychomotor skills trainers encompass use of VR technology and physical probes or objects that help the learners to connect with the objects from the VR environment and convey feedback or tactile sensation to the learners.
- VR patient simulation refers to the interactive computer simulations of real-life clinical scenarios in VR for the purpose of medical training, education, or assessment [16]. They include virtual patients represented by computer-generated 2D or 3D characters or avatars.
- VR serious gaming or gamification intervention involves gaming concepts such as different levels of difficulties, rewards, feedback, and so on, within the computer-generated VR environment for learning purposes.
- AR is a technology that allows a live real-time direct or indirect real-world environment to be augmented or enhanced by computer-generated virtual imagery information (eg, smart, virtually enhanced glasses). Computer-generated information is overlaid on the real-world environment. AR is distinct from VR in which only a computer-generated image is supplied to the user [17].
- MR is a hybrid technology that merges the features of VR and AR [18]. In MR, physical and virtual or digital objects are displayed together and the features of *virtuality* and reality are merged for the learners [19].

Electronic Searches

We developed a comprehensive search strategy for MEDLINE (Ovid), Embase (Elsevier), Cochrane Central Register of Controlled Trials (Wiley), PsycINFO (Ovid), Education Resources Information Center (Ovid), CINAHL (EBSCO), and Web of Science Core Collection (Thomson Reuters). Databases were searched from January 1990 until April 2020 without language restrictions.

We used 1990 as the starting year for our search because before 1990, the use of computers was uncommon for educational use. We used the MEDLINE strategy presented in Multimedia Appendix 1. This was adapted to search the other databases with the help of a librarian (Ms Yasmin Munro). To identify unpublished studies, we searched the International Clinical Trials Registry Platform Search Portal and metaRegister of Controlled Trials. We also checked reference lists of relevant systematic reviews and potentially eligible studies against the inclusion criteria.

Search results across different databases were compiled using EndNote X8 software (Clarivate), and duplicate records were removed. In all, two pairs of two reviewers (BMK, AT, TEF, and SV) independently screened the studies, extracted the data, and carried out data analysis. Any disagreements were resolved by a discussion between the 2 reviewers, with a third reviewer acting as an arbiter if needed. The PRISMA flow diagram was used to report the selection and inclusion of studies [10].

Data Extraction

The data for each of the included studies were independently extracted and managed by 2 reviewers using a structured data recording form, which included information about the study characteristics such as reference of the study, country of the study, the WHO region of the study, name of measurement instrument, description of measurement instrument, types of outcomes reported, assessment category of measurement instrument [5], assessment method of measurement instrument, types of participants, sample size, raters of the instrument, procedure of identifying the raters, and training of the raters for the instruments [20]. We recorded all information relating to validity evidence sources and measurement properties that were reported directly in the articles [5,6]. We also recorded any validity evidence recorded indirectly; for example, through a reference to a validation study focusing on a particular measurement instrument. If the studies presented more than one outcome measure, relevant details of the second outcome measure were also recorded. The data extraction form was piloted and amended according to feedback received. We contacted the study authors for further data in case of missing information.

Data Analysis and Synthesis

We analyzed and synthesized the data as follows: (1) we ascertained the types of primary and secondary outcome measurement instruments; (2) we classified and mapped the data according to types of outcomes (eg, knowledge, skills, attitudes, satisfaction, or competencies); intervention (eg, VR vs classroom-based learning and VR vs serious gaming); year of medical studies (ie, first year, second year, or final year), types of measurement instruments (eg, written exercises [surveys with only multiple-choice questions and surveys with other types of questions and essays] vs in-person assessment where an examiner observed performance [eg, global ratings, structured direct observation, and objective structured clinical examinations]); assessment delivery mode (ie, digital vs classroom-based assessment); and discipline (eg, laparoscopic surgery, anatomy, and internal medicine); and (3) we determined the proportion of RCTs on the use of VR, AR, and MR in undergraduate medical education using measurement instruments with sufficient validity evidence in relation to the goal of the measurements (*validity evidence*). The aim of this study is to comprehensively document outcomes and measurement instruments rather than to synthesize data about the effect of the interventions [6]. Therefore, we did not undertake a risk-of-bias assessment of the studies because it was not relevant to the objectives of this review.

We assessed the validity evidence of the measurement instruments as reported in the cited validation studies using the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) taxonomy of measurement properties [21]. The COSMIN taxonomy outlines three measurement properties or validity evidence domains: reliability, validity, and responsiveness. The reliability domain encompasses measurement properties such as internal consistency, reliability, and measurement error. The domain validity contains the measurement properties such as content validity (including face validity), construct validity (including structural validity, hypotheses testing, and cross-cultural validity and measurement invariance), and criterion validity [21].

Digital assessments were defined as assessments that were delivered exclusively using digital technology (ie, PCs, laptops, mobile phones, and tablets) and included online surveys, questionnaires, computer scoring, or the use of software metrics such as time to completion, number of errors, path length, and so on. Assessments in which digital tools (eg, video recordings or Microsoft PowerPoint presentations) were used to facilitate classroom-based assessment, such as written exercises or in-person observation by the examiners, were not categorized as digital assessments.

Ethics Approval

This systematic mapping review is an analysis of published studies and as such, did not require an ethics approval.

Results

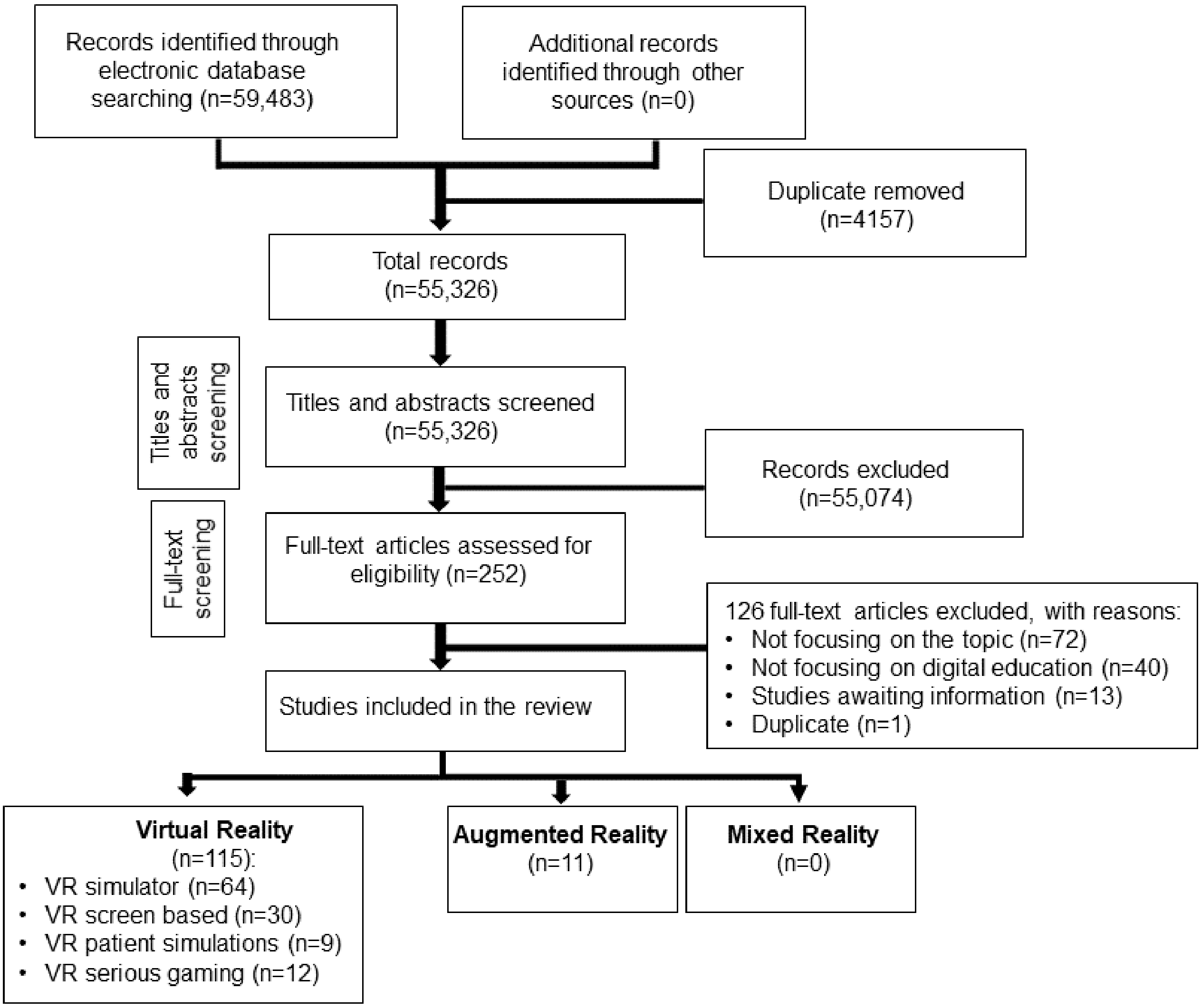
Study Characteristics

The searches identified 59,483 records through electronic databases, of which we included 126 (0.21%) RCTs. Of the 126 RCTS, 115 (91.3%) assessed different forms of VR, whereas 11 (8.7%) focused on AR simulations (Figure 1). We did not find any study evaluating the use of MR in medical student education.

Of the 115 included articles focusing on VR-based training for medical student education, 64 (55.7%) focused on VR-based psychomotor skills training [22-85], 30 (26.1%) on screen-based VR [86-115], 9 (7.8%) on VR patient simulations [116-124], and 12 (10.4%) on VR serious gaming and gamification [125-136]. Only 8.7% (11/126) of the included studies focused on AR simulations [137-147] and none focused on MR training in medical student education. The included studies were published between 1997 and 2020. Most of the studies were from high-income countries, except for 8.7% (11/126) of the studies, which were conducted in low- and middle-income countries [35,36,72,75,105,114,126,127,132,134,139]. Of the 126 studies, 31 (24.6%) cited validation studies for the measurement instruments used [23,25,27,30-32,34-36,47,48, 52,58,60,63-65,70,72,78,79,82,84,92,101,118-120,126,128,133] (Multimedia Appendices 2 and 3).

Participants included medical students from the first to sixth year of medical schools (N=9010). The studies compared the use of VR and AR training (either stand-alone intervention or blended with traditional, nondigital learning) with traditional, nondigital learning or a different form of VR and AR training or other forms of digital education such as online digital education or offline digital education. Of the 64 studies focused on the effects of VR simulators for medical student education, 61 (95%) were delivered in a university setting, whereas 3 (5%) were conducted in a hospital setting [37,72,74].

Figure 1. Study flow diagram. VR: virtual reality.

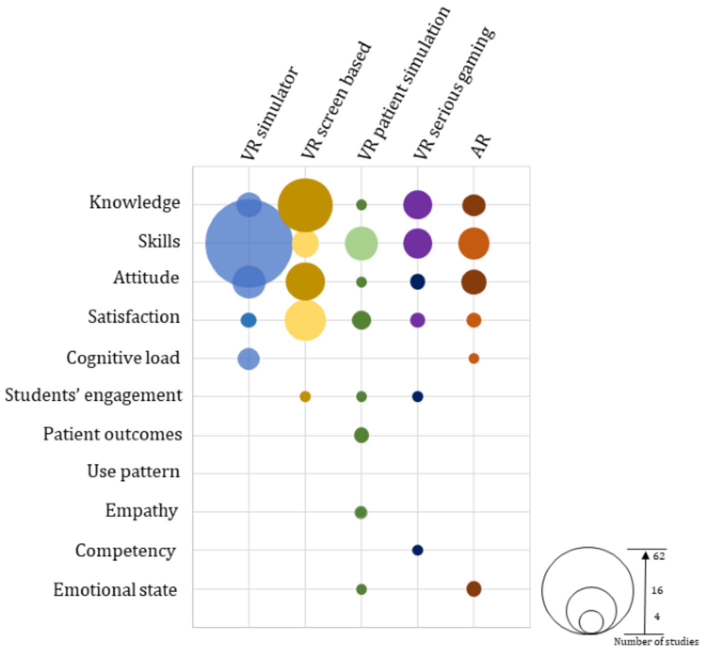


VR Simulators

Of the 115 VR studies, 64 (55.6%) with 3132 medical students evaluated the effects of VR simulators in medical student education [22-85]. The studies included first year to sixth year medical students and were published between 2001 and 2020. In terms of the topic or area of study, 53% (34/64) of the studies focused on laparoscopic surgery [22,24,27,31,35-38,40,41,45-48,50-52,54,56-66,69,78,81-83]; 16% (10/64) on surgery [25,28,53,55,67,68,71,74,76,77]; 8% (5/64) on orthopedic surgery [39,42,73,79,84]; 8% (5/64) on ureteroscopy [30,33,34,80,85]; 5% (3/64) each on ophthalmology [26,70,75] and intravenous cannulation [29,32,72], and 2% (1/64) each on endoscopy [49], colonoscopy [23], shoulder-joint clinical anatomy [44], and empathic communication skills [43].

For the outcomes, 97% (62/64) of the studies reported on participants' postintervention skills [22-43,45-53,55-85], 8% (5/64) on knowledge [28,37,44,54,65], 14% (9/64) on attitudes toward the intervention [31,32,44,48,54,65,66,71,75], 3% (2/64) on satisfaction [68,71] and 6% (4/64) on cognitive load [25,27,39,63](Figure 2). Of the 62 studies that reported on participants' postintervention skills, 11 (18%) reported change score from baseline for the skills outcome [25,50,56,58,68,73,76-78,80,85] and 1 (2%) reported change score from baseline for the satisfaction outcome [68]. Regarding retention, 7.8% (5/64) of the studies assessed skills retention at 2-4 weeks after the intervention [25,31,33,40,83]. The remaining studies did not report retention outcomes.

Figure 2. Types of reported outcomes in virtual reality (VR)– and augmented reality (AR)–based training.



For modes of assessment, 46.8% (30/64) of the studies used paper-based written assessments or in-person assessments (ie, nondigital) using checklists by the examiners [24,25,31-37,39,46,47,51-55,58,60,65,68,70-72,75,79-82,84]; 31% (20/64) used digital assessments such as software-based metrics (eg, time spent on training, number of errors, total path length, motion analysis, or checklists) [22,23,26,29,40-42,45,49,50,56,57,59,61,62,64,69,73,78,85]; 11% (7/64) used a combination of digital assessments using software-based metrics, paper-based written assessments, or in-person assessments by supervising examiners [27,38,43,44,48,63,66] and 2% (1/64) used both paper-based written assessments and in-person assessments using checklists [37]. In 10.1% (7/64) of the studies, the mode of assessment was unspecified [28,30,67,74,76,77,83].

For validity evidence, 54.6% (35/64) of the studies reported a single form of validity evidence (mostly either internal consistency or reliability) for the measurement instruments largely used for assessment of skills [22,23,25,27,30-37,39,40,47,48,51-55,58,60,63-66,68,70,72,78-80,82,84] (Multimedia Appendices 2 and 3). The remaining studies did not provide any information on the validity of assessment tools used for measuring the outcomes. Of the 64 studies, 23 (36%) referenced pertinent measurement instrument validation studies, largely used for assessment of skills (mostly either internal consistency or reliability) for the measurement instruments largely used for assessment of skills [23,25,27,30-32,34-36,47,48,52,58,60,63-65,70,72,78,79,82,84]. Of the measurement properties, these studies mostly reported internal consistency and reliability, followed by structural validity and hypotheses testing.

Screen-Based VR

Of the 115 VR studies, 30 (26.1%) studies with 2409 medical students evaluated the effect of screen-based or nontechnical training for medical students [86-115]. The studies included first year to sixth year medical students and were published between 1997 and 2020. In terms of the topic or area of study, 37% (11/30) of the studies focused on anatomy [87,91,95-98,100,102,104,106,114]; 17% (5/30) on ophthalmology [93,109,112,113,115]; 17% (5/30) on surgery [88,90,92,101,105]; 6% (2/30) each on patient examination [99,108] and one study each (3%,1/30) on operating room introduction [107], biomechanics of the spine [89], histology [111], trauma [94], traumatic head injury [86], radiology [103], and genetics [110].

For the outcomes, 80% (24/30) of the studies reported on participants' postintervention knowledge [89,91,93-107,109-115], 17% (5/30) on skills [88,92,99,101,107], 40% (12/30) on attitudes toward topics and interventions [86,87,90,91,95,97,102-104,107,108,115], 47% (14/30) on satisfaction [87,89,91-93,97,98,100, 102,105,109,112-114] and 3% (1/30) on students' learning engagement [89] (Figure 2). Of the 24 studies assessing knowledge, 5 (2%) also reported change score from baseline [101,104,105,113,114]. Similarly, 20% (1/5) of the studies assessing skills [101], 17% (2/12) of the studies assessing attitude toward the intervention [90,104], and 21% (3/14) of the studies assessing satisfaction [105,113,114] also reported change score from baseline. Regarding retention, only a single study assessed retention at 12 months after the intervention [112]. The remaining studies did not report outcomes at the follow-up stages.

Most of the studies (21/30, 70%) used paper-based written assessments [86,87,89-91,93,95,97,98,100,102-104,108-115]. Other forms of assessment included in-person assessments by an examiner [88], digital assessment in the form of questionnaires and ratings [94,105,106], combined paper-based written and in-person assessments [92,99,101,107], and a paper-based written assessment with questions delivered in the form of a PowerPoint presentation [96].

Of the 30 studies, 8 (27%) reported at least one form of validity evidence (mostly reliability) for the measurement instruments that were largely used to assess skills [88,91,92, 98,99,101,107,108]. Of these 8 studies, 2 (25%) referenced measurement instrument validation studies, both focusing on skills assessment and reporting on their reliability [92,101].

VR Patient Simulations

Of the 115 VR studies, 9 (7.8%) with 782 medical students evaluated the effect of VR-based patient simulations in medical student education simulations [116-124]. Of these 9 studies, 4 (44%) focused on communication skills [117-119,124]; 2 (22%) on pediatric life support [121,122]; and 1 (11%) each on clinical reasoning [123], internal medicine [116], and suicide risk assessment [120] (Figure 2).

For the outcomes, 11% (1/9) of the studies reported on participants' postintervention knowledge [122], 100% (9/9) on skills [116-124], 33% (3/9) on students' satisfaction [119,120,123], 22% (2/9) on patient-related outcomes (eg, patients' satisfaction) [119,120], and 11% (1/9) each on attitudes toward the intervention [124], engagement [123], mood changes or emotional state [124], and empathetic behavior [117]. None of the studies reported change score from baseline or retention data.

For mode of assessment, most of the studies used in-person assessments by an examiner [116-120,123,124] or paper-based written assessments [119,120,122,123]. Of the 9 studies, 2 (22%) used both paper-based written and in-person assessments by an examiner [119,120]; 1 (11%) used both digital assessments consisting of virtual patients and scoring and in-person assessment by an examiner [116]; and, finally, 1 (11%) used a combined assessment of digital assessment in the form of a survey, in-person assessment by an examiner, and paper-based written assessment for different outcomes [123].

Of the 9 studies, 5 (56%) reported at least one form of validity evidence (mostly internal consistency and reliability) for the measurement instruments used to assess skills [116-120] (Multimedia Appendices 2 and 3). Of these 5 studies, 3 (60%) referenced measurement instrument validation studies: 67% (2/3) focused on assessment of patient satisfaction [119,120] and 33% (1/3) on skills [118]. The measurement properties mentioned in the referenced validation studies were internal consistency and reliability, followed by internal validity.

VR Serious Gaming and Gamification

Of the 115 studies, 12 (10.4%) with 743 medical students evaluated the effects of VR serious gaming and gamification in medical student education [125-136]. The studies included participants from the first to fifth year of studies and were published between 2008 and 2020. Regarding the topic or area of study, 25% (3/12) of the studies focused on surgery [126,129,136] and 8% (1/12) each on acute medicine [131], advanced life support [132], basic life support [127], engagement and self-efficacy beliefs [128], geriatric medicine [130], laparoscopy [135], pediatrics [133], primary care [134], and urology [125].

For the outcomes, 58% (7/12) of the studies reported on participants' postintervention knowledge [125,127,129, 130,132-134], 58% (7/12) on skills [126,127,129, 131,132,135,136], 17% (2/12) on attitudes toward the intervention and toward the outcomes [125,132], 17% (2/12) on satisfaction [133,134], 8% (1/12) on competencies [130] and 8% (1/12) on engagement and self-efficacy belief [128](Figure 2). Of the 7 studies assessing participants' skills, 1 (14%) also reported change score from baseline [126]. Overall, 25% (3/12) of the studies assessed retention [126,133,134]. Of these 3 studies, 2 (67%) assessing the knowledge outcome also assessed retention from 4 to 6 weeks after the intervention [133,134] and 1 (33%) assessing the skills outcome also assessed retention at 3 weeks after the intervention [126].

For the assessment methods, most of the included studies used paper-based written assessments [125,130], in-person assessments by supervising clinicians [126,131,135,136], or both assessment methods [127,129,132]. Of the 12 studies, 1 (8%) used digital assessments in the form of a questionnaire in addition to paper-based written assessment [134], 1 (8%) used only digital assessments in the form of a questionnaire [133], and the mode of assessment in 1 (8%) was not mentioned [128].

Of the 12 studies, 7 (58%) reported at least one form of validity evidence (mostly internal consistency and reliability) for the measurement instruments that were mainly used to assess knowledge [125,126,128-130,133,134] (Multimedia Appendices 2 and 3). Of these 7 measurement instruments, 4 (57%) were focused on knowledge, 2 (29%) on skills, 2 (29%) on satisfaction, and 1 (14%) each on cognitive load and self-efficacy beliefs. Of the 7 studies, 3 (43%) referenced a measurement instrument validation study [126,128,133]. The reported measurement properties included internal consistency (for the skills, engagement, and satisfaction measurement instrument), reliability (for the skills and engagement measurement instrument), structural validity (for the skills and satisfaction measurement instrument), and hypothesis (for the skills measurement instrument).

AR Interventions

Of the 126 studies, 11 (8.7%) with 448 medical students used an AR intervention to assess the outcomes [137-147]. The studies included first year to fourth year medical students and were published between 2013 and 2020. The studies covered different topics, including arthroplasty [142], facet joint injection [143], needle insertion [147], general medicine [144], forensic medicine [137], ophthalmology [140], surgery [141,145], laparoscopy [146], and anatomy [138,139].

The reported outcomes included participants' postintervention knowledge [137-139,144], skills [138,140-143,145-147], attitudes toward learning experience or intervention [137,140-142,144], satisfaction [138,146], emotional state , [137,144] and cognitive load [139] (Figure 2). Most studies used paper-based written assessments [137-139,144] or in-person assessments by examiners [143,147] or both approaches [140,142,146]. Of the 11 studies, 1 (9%) used both digital and paper-based written assessments [141] and 1 (9%) used digital assessment in the form of software-based metrics [145]. Of the 8 studies assessing a skills outcome, 2 (25%) also reported change score from baseline [138,145]. Similarly, of the 6 studies assessing knowledge and satisfaction, 1 (17%) also reported change score from baseline [138]. In terms of retention, only 25% (1/4) of the studies assessing knowledge also reported retention 2 weeks after the intervention [144].

Of the 11 studies, 6 (55%) reported at least one form of validity evidence (mostly internal consistency) for a variety of measurement instruments used [137-140,144,145]. These measurement instruments were used to assess knowledge in 18% (2/11) of the studies, attitudes in 18% (2/11), and emotional state in 18% (2/11), whereas in 9% (1/11) of the studies each, skills, cognitive load, and visuospatial assessment were assessed. None of the studies provided references for validation of the instruments used to measure the outcomes.

MR Interventions

None of the included studies assessed the effectiveness of MR interventions in medical student education.

Discussion

Principal Findings

In this review, we assessed and mapped the choice of outcomes, measurement instruments, and the prevalence of measurement instruments with validity evidence in RCTs on the use of ER technologies in undergraduate medical education. Among the 126 included studies, we found 115 (91.3%) RCTs on different forms of VR, 11 (8.7%) articles on AR simulations, and no RCTs on MR in medical student education. The included studies often reported only a single outcome and immediate postintervention assessments. The types of reported outcomes varied across different types of VR and AR simulations. Participants' skills were the most common outcomes measured in studies on VR simulators, VR patient simulations, and AR. Participants' knowledge was the most common outcome measured in studies on screen-based VR and VR serious games. Other more commonly reported primary outcomes were participants' attitudes toward the intervention or topic and satisfaction with the intervention. More than half of the studies on VR simulators, VR patient simulations, VR serious gaming, and AR as well as only a quarter of the studies on screen-based VR reported at least one form of validity evidence. The most common validity evidence for the measurement instruments used were internal consistency and reliability. Most of the studies used nondigital assessment methods such as paper-based written or in-person assessments by an examiner.

Comparison With Existing Literature

There is a lack of standardization regarding the choice of outcomes and assessments in RCTs focusing on ER for medical student education. The findings are in line with published reviews focusing on the effectiveness of digital education for pre- and postregistration health professionals [1,131,148].

Our review shows a diversity of outcomes and measurement instruments used in trials on ER in medical education. Reporting of a limited set of outcomes, immediate postintervention data, and the use of measurement instruments lacking validity evidence is common in RCTs on different digital health professions education modalities. However, the choice of appropriate outcomes as well as robust measurement instruments to assess these outcomes is essential when designing trials. It is also important that the chosen outcomes are relevant to key stakeholders who will be able to influence policy and practice. This can be achieved through the development and use of an agreed standardized collection of outcomes and measurement instruments [21].

Strengths and Limitations

In our review, we used a comprehensive search strategy for 7 major bibliographic databases and gray literature sources without language limitations to identify relevant studies. We covered the search period starting from 1990 onward to include all available RCTs on VR-, AR-, and MR-based trainings in medical student education. We performed the screening and data extraction in parallel and independently to ensure reliability of our findings.

There are also some limitations to our study. We performed a descriptive analysis and mapping of outcomes and validity evidence for the measurement instruments used. A more in-depth analysis of the types of validity evidence used was not feasible because of limited information in the included studies. We aimed to complement this by searching for, and including, additional information on validity evidence from validation studies referenced in the included studies. However, information provided in these referenced validation studies was also often limited. We acknowledge that some of the mentioned measurement instruments may have validity evidence not reported in the included RCT papers or for which no validity study was referenced. Furthermore, the reporting of validity evidence in the included RCTs and validation studies may be incomplete and not reflect all validity evidence for a particular measurement instrument. Finally, to determine the validity evidence for the measurement instruments used in the included trials, we used COSMIN, an established taxonomy of measurement properties. Although COSMIN was originally developed for health outcome measurement instruments, it is also applicable to other types of outcomes. However, there are other validity frameworks that were developed primarily for education and may be more appropriate for future analysis of medical education outcomes [9,149].

Future Recommendations

Future studies should aim to include a broader set of outcomes, report change score from baseline, and assess learning retention. They should also aim to use measurement instruments with validity evidence. We list those used in the included trials in Multimedia Appendix 3. Most of the measurement instruments with validity evidence were used to assess participants' skills. There is a need for greater use or adaptation of existing measurement instruments with validity evidence and potentially also development of new ones assessing other relevant outcomes such as attitudes and satisfaction. In addition, digital technology offers diverse and potentially more efficient approaches to assessment and should be more extensively explored and applied in this area. This is particularly relevant given the pervasive and sudden shift to remote teaching because of the COVID-19 pandemic.

Conclusions

Studies on the use of VR and AR in undergraduate medical education often report a limited set of outcomes, mostly knowledge and skills, and usually immediate postintervention assessment data. The use of measurement instruments with validity evidence for outcomes other than skills is limited, as is the use of digital forms of assessment. Future studies should report a broader set of outcomes, change score from baseline, and retention data, as well as use measurement instruments with validity evidence.

Acknowledgments

The authors would like to acknowledge funding support from Nanyang Technological University, Singapore. The authors are also grateful to Ms Yasmin Munro for her assistance with our search strategy.

Authors' Contributions

LTC conceived the idea for the review. BMK, AT, and TEF screened the studies. BMK, AT, TEF, and SV extracted and analyzed the data from the eligible studies. BMK and LTC wrote the review, and LTC provided methodological guidance. SK, CA, and NC critically revised the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MEDLINE (Ovid) search strategy.
[DOCX File , 21 KB - games_v10i2e29594_app1.docx]

Multimedia Appendix 2
Characteristics of the included studies.
[DOCX File , 47 KB - games_v10i2e29594_app2.docx]

Multimedia Appendix 3
Types and number of reported outcomes and measurement instruments with validity in the included studies.
[DOCX File , 27 KB - games_v10i2e29594_app3.docx]

References

1. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, 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 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
2. George PP, Papachristou N, Belisario JM, Wang W, Wark PA, Cotic Z, et al. Online eLearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010406 [FREE Full text] [doi: 10.7189/jogh.04.010406] [Medline: 24976965]
3. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]
4. Harden RM. Outcome-Based Education: the future is today. Med Teach 2007 Sep;29(7):625-629. [doi: 10.1080/01421590701729930] [Medline: 18236247]
5. Epstein RM. Assessment in medical education. N Engl J Med 2007 Jan 25;356(4):387-396. [doi: 10.1056/NEJMra054784] [Medline: 17251535]
6. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC, US: American Educational Research Association; 2014.
7. Ratanawongsa N, Thomas PA, Marinopoulos SS, Dorman T, Wilson LM, Ashar BH, et al. The reported validity and reliability of methods for evaluating continuing medical education: a systematic review. Acad Med 2008 Mar;83(3):274-283. [doi: 10.1097/ACM.0b013e3181637925] [Medline: 18316877]
8. Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Technology-enhanced simulation to assess health professionals: a systematic review of validity evidence, research methods, and reporting quality. Acad Med 2013 Jun;88(6):872-883. [doi: 10.1097/ACM.0b013e31828ffdcf] [Medline: 23619073]
9. Kane MT. Validating the Interpretations and Uses of Test Scores. Journal of Educational Measurement 2013 Mar 14;50(1):1-73. [doi: 10.1111/jedm.12000]

10. 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 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
11. 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]
12. eLearning for undergraduate health professional education: A systematic review informing a radical transformation of health workforce development. World Health Organization. 2015. URL: https://www.who.int/hrh/documents/14126-eLearningReport.pdf [accessed 2022-04-03]
13. International Standard Classification of Education: Fields of education and training 2013 (ISCED-F 2013). Unesco Institute for Statistics. 2014. URL: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-fields-of-education-and-training-2013-detailed-field-descriptions-2015-en.pdf [accessed 2022-04-06]
14. Strangman N, Hall T, Meyer A. Virtual Reality/Computer Simulations and the Implications for UDL Implementation. National Center on Accessing the General Curriculum. 2003. URL: https://ccie.ucf.edu/wp-content/uploads/sites/8/2018/02/VirtualRealityUDL.pdf [accessed 2022-04-06]
15. Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. Cyberpsychol Behav 2003 Aug;6(4):389-395. [doi: 10.1089/109493103322278772] [Medline: 14511451]
16. Ellaway R, Candler C, Greene P, Smothers V. An Architectural Model for MedBiquitous Virtual Patients. MedBiquitous. 2006. URL: http://groups.medbiq.org/medbiq/display/VPWG/MedBiquitous+Virtual+Patient+Architecture [accessed 2022-04-06]
17. Zhu E, Hadadgar A, Masiello I, Zary N. Augmented reality in healthcare education: an integrative review. PeerJ 2014;2:e469 [FREE Full text] [doi: 10.7717/peerj.469] [Medline: 25071992]
18. Tepper OM, Rudy HL, Lefkowitz A, Weimer KA, Marks SM, Stern CS, et al. Mixed Reality with HoloLens: Where Virtual Reality Meets Augmented Reality in the Operating Room. Plast Reconstr Surg 2017 Nov;140(5):1066-1070. [doi: 10.1097/PRS.0000000000003802] [Medline: 29068946]
19. Flavián C, Ibáñez-Sánchez S, Orús C. The impact of virtual, augmented and mixed reality technologies on the customer experience. Journal of Business Research 2019 Jul;100(4):547-560. [doi: 10.1016/j.jbusres.2018.10.050] [Medline: 23725941]
20. Law GC, Apfelbacher C, Posadzki PP, Kemp S, Tudor Car L. Choice of outcomes and measurement instruments in randomised trials on eLearning in medical education: a systematic mapping review protocol. Syst Rev 2018 May 17;7(1):75 [FREE Full text] [doi: 10.1186/s13643-018-0739-0] [Medline: 29776434]
21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010 Jul;63(7):737-745. [doi: 10.1016/j.jclinepi.2010.02.006] [Medline: 20494804]
22. Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. Am J Surg 2006 Jan;191(1):128-133. [doi: 10.1016/j.amjsurg.2005.10.014] [Medline: 16399123]
23. Ahad S, Boehler M, Schwind CJ, Hassan I. The effect of model fidelity on colonoscopic skills acquisition. A randomized controlled study. J Surg Educ 2013 May;70(4):522-527. [doi: 10.1016/j.jsurg.2013.02.010] [Medline: 23725941]
24. Ahlberg G, Heikkinen T, Iselius L, Leijonmarck C, Rutqvist J, Arvidsson D. Does training in a virtual reality simulator improve surgical performance? Surg Endosc 2002 Jan;16(1):126-129. [doi: 10.1007/s00464-001-9025-6] [Medline: 11961622]
25. Andersen SAW, Konge L, Cayé-Thomasen P, Sørensen MS. Learning Curves of Virtual Mastoidectomy in Distributed and Massed Practice. JAMA Otolaryngol Head Neck Surg 2015 Oct;141(10):913-918. [doi: 10.1001/jamaoto.2015.1563] [Medline: 26334610]
26. Bergqvist J, Person A, Vestergaard A, Grauslund J. Establishment of a validated training programme on the Eyesi cataract simulator. A prospective randomized study. Acta Ophthalmol 2014 Nov 13;92(7):629-634 [FREE Full text] [doi: 10.1111/aos.12383] [Medline: 24612448]
27. Bjerrum F, Sorensen JL, Konge L, Rosthøj S, Lindschou J, Ottesen B, et al. Randomized trial to examine procedure-to-procedure transfer in laparoscopic simulator training. Br J Surg 2016 Jan;103(1):44-50. [doi: 10.1002/bjs.9966] [Medline: 26511775]
28. Bowyer CMW, Liu AV, Bonar JP. Validation of SimPL -- a simulator for diagnostic peritoneal lavage training. Stud Health Technol Inform 2005;111:64-67. [Medline: 15718700]
29. Bowyer MW, Pimentel EA, Fellows JB, Scofield RL, Ackerman VL, Horne PE, et al. Teaching intravenous cannulation to medical students: comparative analysis of two simulators and two traditional educational approaches. Stud Health Technol Inform 2005;111:57-63. [Medline: 15718699]
30. Brunckhorst O, Shahid S, Aydin A, McIlhenny C, Khan S, Raza S, et al. MP22-09 DEVELOPMENT AND VALIDATION OF AN INTEGRATED SKILLS CURRICULUM WITHIN URETEROSCOPY– A RANDOMISED CONTROLLED TRIAL. Journal of Urology 2015 Apr;193(4S):1018-1025. [doi: 10.1016/j.juro.2015.02.1021] [Medline: 19616797]

31. Van Bruwaene S, Schijven MP, Napolitano D, De Win G, Miserez M. Porcine cadaver organ or virtual-reality simulation training for laparoscopic cholecystectomy: a randomized, controlled trial. J Surg Educ 2015 Oct;72(3):483-490. [doi: 10.1016/j.jsurg.2014.11.015] [Medline: 25555673]
32. Brydges R, Carnahan H, Rose D, Rose L, Dubrowski A. Coordinating Progressive Levels of Simulation Fidelity to Maximize Educational Benefit. Academic Medicine 2010 Dec 13;85(5):806-812. [doi: 10.1097/acm.0b013e3181d7aabd] [Medline: 23493565]
33. Bube S, Dagnaes-Hansen J, Mahmood O, Rohrsted M, Bjerrum F, Salling L, et al. Simulation-based training for flexible cystoscopy - A randomized trial comparing two approaches. Heliyon 2020 Jan;6(1):e03086 [FREE Full text] [doi: 10.1016/j.heliyon.2019.e03086] [Medline: 31922043]
34. Chou DS, Abdelshehid C, Clayman RV, McDougall EM. Comparison of results of virtual-reality simulator and training model for basic ureteroscopy training. J Endourol 2006 Apr;20(4):266-271 [FREE Full text] [doi: 10.1089/end.2006.20.266] [Medline: 16646655]
35. da Cruz JAS, Dos Reis ST, Cunha Frati RM, Duarte RJ, Nguyen H, Srougi M, et al. Does Warm-Up Training in a Virtual Reality Simulator Improve Surgical Performance? A Prospective Randomized Analysis. J Surg Educ 2016 Oct;73(6):974-978. [doi: 10.1016/j.jsurg.2016.04.020] [Medline: 27233673]
36. da Cruz JAS, Sandy NS, Passerotti CC, Nguyen H, Antunes AA, Dos Reis ST, et al. Does training laparoscopic skills in a virtual reality simulator improve surgical performance? J Endourol 2010 Nov;24(11):1845-1849. [doi: 10.1089/end.2010.0328] [Medline: 20958203]
37. De La Garza JR, Schmidt MW, Kowalewski K, Benner K, Müller PC, Kenngott HG, et al. Does rating with a checklist improve the effect of E-learning for cognitive and practical skills in bariatric surgery? A rater-blinded, randomized-controlled trial. Surg Endosc 2019 May;33(5):1532-1543. [doi: 10.1007/s00464-018-6441-4] [Medline: 30209607]
38. Eldred-Evans D, Grange P, Cheang A, Yamamoto H, Ayis S, Mulla M, et al. Using the mind as a simulator: a randomized controlled trial of mental training. J Surg Educ 2013 Apr 23;70(4):544-551. [doi: 10.1016/j.jsurg.2013.04.003] [Medline: 23725944]
39. Frithioff A, Frendø M, Mikkelsen PT, Sørensen MS, Andersen SAW. Ultra-high-fidelity virtual reality mastoidectomy simulation training: a randomized, controlled trial. Eur Arch Otorhinolaryngol 2020 May;277(5):1335-1341. [doi: 10.1007/s00405-020-05858-3] [Medline: 32067096]
40. Fu Y, Cavuoto L, Qi D, Panneerselvam K, Arikatla VS, Enquobahrie A, et al. Characterizing the learning curve of a virtual intracorporeal suturing simulator VBLaST-SS©. Surg Endosc 2020 Jul;34(7):3135-3144 [FREE Full text] [doi: 10.1007/s00464-019-07081-6] [Medline: 31482354]
41. Ganai S, Donroe JA, St Louis MR, Lewis GM, Seymour NE. Virtual-reality training improves angled telescope skills in novice laparoscopists. Am J Surg 2007 Feb;193(2):260-265. [doi: 10.1016/j.amjsurg.2005.11.019] [Medline: 17236859]
42. Gasco J, Patel A, Ortega-Barnett J, Branch D, Desai S, Kuo Y, et al. Virtual reality spine surgery simulation: an empirical study of its usefulness. Neurological Research 2014 May 20;36(11):968-973. [doi: 10.1179/1743132814y.0000000388] [Medline: 26841278]
43. Guetterman TC, Sakakibara R, Baireddy S, Kron FW, Scerbo MW, Cleary JF, et al. Medical Students' Experiences and Outcomes Using a Virtual Human Simulation to Improve Communication Skills: Mixed Methods Study. J Med Internet Res 2019 Nov 27;21(11):e15459-e15459 [FREE Full text] [doi: 10.2196/15459] [Medline: 31774400]
44. Hariri S, Rawn C, Srivastava S, Youngblood P, Ladd A. Evaluation of a surgical simulator for learning clinical anatomy. Med Educ 2004 Aug;38(8):896-902. [doi: 10.1111/j.1365-2929.2004.01897.x] [Medline: 15271051]
45. Hiemstra E, Terveer EM, Chmarra MK, Dankelman J, Jansen FW. Virtual reality in laparoscopic skills training: is haptic feedback replaceable? Minim Invasive Ther Allied Technol 2011 May;20(3):179-184. [doi: 10.3109/13645706.2010.532502] [Medline: 21438717]
46. Hyltander A, Liljegren E, Rhodin PH, Lönroth H. The transfer of basic skills learned in a laparoscopic simulator to the operating room. Surg Endosc 2002 Sep;16(9):1324-1328 [FREE Full text] [doi: 10.1007/s00464-001-9184-5] [Medline: 11988802]
47. Johnston TJ, Tang B, Alijani A, Tait I, Steele RJ, Ker J, Surgical Simulation Group at the University of Dundee. Laparoscopic surgical skills are significantly improved by the use of a portable laparoscopic simulator: results of a randomized controlled trial. World J Surg 2013 May;37(5):957-964. [doi: 10.1007/s00268-013-1945-5] [Medline: 23404485]
48. Kanumuri P, Ganai S, Wohaibi EM, Bush RW, Grow DR, Seymour NE. Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices. JSLS 2008;12(3):219-226 [FREE Full text] [Medline: 18765042]
49. Karabanov AN, Irmen F, Madsen KH, Haagensen BN, Schulze S, Bisgaard T, et al. Getting to grips with endoscopy - Learning endoscopic surgical skills induces bi-hemispheric plasticity of the grasping network. Neuroimage 2019 Apr 01;189:32-44. [doi: 10.1016/j.neuroimage.2018.12.030] [Medline: 30583066]
50. Kothari SN, Kaplan BJ, DeMaria EJ, Broderick TJ, Merrell RC. Training in laparoscopic suturing skills using a new computer-based virtual reality simulator (MIST-VR) provides results comparable to those with an established pelvic trainer system. J Laparoendosc Adv Surg Tech A 2002 Jun 11;12(3):167-173 [FREE Full text] [doi: 10.1089/10926420260188056] [Medline: 12184901]

51. Kowalewski K, Minassian A, Hendrie JD, Benner L, Preukschas AA, Kenngott HG, et al. One or two trainees per workplace for laparoscopic surgery training courses: results from a randomized controlled trial. Surg Endosc 2019 May;33(5):1523-1531. [doi: 10.1007/s00464-018-6440-5] [Medline: 30194644]
52. Krogh C, Konge L, Bjurström J, Ringsted C. Training on a new, portable, simple simulator transfers to performance of complex bronchoscopy procedures. The Clinical Respiratory Journal 2012 Aug 20;7(3):237-244. [doi: 10.1111/j.1752-699x.2012.00311.x] [Medline: 15718699]
53. Lee CR, Rho SY, Han SH, Moon Y, Hwang SY, Kim YJ, et al. Comparison of Training Efficacy Between Custom-Made Skills Simulator (CMSS) and da Vinci Skills Simulators: A Randomized Control Study. World J Surg 2019 Nov;43(11):2699-2709. [doi: 10.1007/s00268-019-05108-6] [Medline: 31399794]
54. Lesch H, Johnson E, Peters J, Cendán JC. VR Simulation Leads to Enhanced Procedural Confidence for Surgical Trainees. J Surg Educ 2020 Feb;77(1):213-218 [FREE Full text] [doi: 10.1016/j.jsurg.2019.08.008] [Medline: 31466895]
55. Lindquist NR, Leach M, Simpson MC, Antisdel JL. Evaluating Simulator-Based Teaching Methods for Endoscopic Sinus Surgery. Ear Nose Throat J 2019 Sep 20;98(8):490-495 [FREE Full text] [doi: 10.1177/0145561319844742] [Medline: 31018690]
56. Loukas C, Lahanas V, Kanakis M, Georgiou E. The Effect of Mixed-Task Basic Training in the Acquisition of Advanced Laparoscopic Skills. Surg Innov 2015 Aug;22(4):418-425. [doi: 10.1177/1553350614556365] [Medline: 25377214]
57. Loukas C, Nikiteas N, Schizas D, Lahanas V, Georgiou E. A head-to-head comparison between virtual reality and physical reality simulation training for basic skills acquisition. Surg Endosc 2012 Sep 27;26(9):2550-2558. [doi: 10.1007/s00464-012-2230-7] [Medline: 22476832]
58. Lucas S, Tuncel A, Bensalah K, Zeltser I, Jenkins A, Pearle M, et al. Virtual reality training improves simulated laparoscopic surgery performance in laparoscopy naïve medical students. J Endourol 2008 May;22(5):1047-1051. [doi: 10.1089/end.2007.0366] [Medline: 18643722]
59. Madan AK, Frantzides CT. Prospective randomized controlled trial of laparoscopic trainers for basic laparoscopic skills acquisition. Surg Endosc 2007 Feb;21(2):209-213. [doi: 10.1007/s00464-006-0149-6] [Medline: 17122975]
60. McDougall EM, Kolla SB, Santos RT, Gan JM, Box GN, Louie MK, et al. Preliminary study of virtual reality and model simulation for learning laparoscopic suturing skills. J Urol 2009 Sep 05;182(3):1018-1025. [doi: 10.1016/j.juro.2009.05.016] [Medline: 19616797]
61. Mulla M, Sharma D, Moghul M, Kailani O, Dockery J, Ayis S, et al. Learning basic laparoscopic skills: a randomized controlled study comparing box trainer, virtual reality simulator, and mental training. J Surg Educ 2012 Sep 5;69(2):190-195. [doi: 10.1016/j.jsurg.2011.07.011] [Medline: 22365864]
62. Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A. Laparoscopic virtual reality and box trainers: is one superior to the other? Surg Endosc 2004 Mar 1;18(3):485-494. [doi: 10.1007/s00464-003-9043-7] [Medline: 14752633]
63. Muresan C, Lee TH, Seagull J, Park AE. Transfer of training in the development of intracorporeal suturing skill in medical student novices: a prospective randomized trial. Am J Surg 2010 Oct;200(4):537-541. [doi: 10.1016/j.amjsurg.2009.12.018] [Medline: 20451174]
64. Nehme J, Sodergren MH, Sugden C, Aggarwal R, Gillen S, Feussner H, et al. A randomized controlled trial evaluating endoscopic and laparoscopic training in skills transfer for novices performing a simulated NOTES task. Surg Innov 2013 Dec 1;20(6):631-638. [doi: 10.1177/1553350613480854] [Medline: 23493565]
65. Nickel F, Brzoska JA, Gondan M, Rangnick HM, Chu J, Kenngott HG, et al. Virtual reality training versus blended learning of laparoscopic cholecystectomy: a randomized controlled trial with laparoscopic novices. Medicine (Baltimore) 2015 May;94(20):e764-e769 [FREE Full text] [doi: 10.1097/MD.0000000000000764] [Medline: 25997044]
66. Oussi N, Enochsson L, Henningsohn L, Castegren M, Georgiou E, Kjellin A. Trainee Performance After Laparoscopic Simulator Training Using a Blackbox versus LapMentor. J Surg Res 2020 Jun;250(3):1-11. [doi: 10.1016/j.jss.2019.12.039] [Medline: 32014696]
67. Patel A, Koshy N, Ortega-Barnett J, Chan HC, Kuo Y, Luciano C, et al. Neurosurgical tactile discrimination training with haptic-based virtual reality simulation. Neurol Res 2014 Dec;36(12):1035-1039. [doi: 10.1179/1743132814Y.0000000405] [Medline: 24984771]
68. Plana NM, Rifkin WJ, Kantar RS, David JA, Maliha SG, Farber SJ, et al. A Prospective, Randomized, Blinded Trial Comparing Digital Simulation to Textbook for Cleft Surgery Education. Plast Reconstr Surg 2019 Jan;143(1):202-209 [FREE Full text] [doi: 10.1097/PRS.0000000000005093] [Medline: 30325894]
69. Schlosser K, Alkhawaga M, Maschuw K, Zielke A, Mauner E, Hassan I. Training of laparoscopic skills with virtual reality simulator: a critical reappraisal of the learning curve. Eur Surg 2007 Jun 19;39(3):180-184. [doi: 10.1007/s10353-006-0292-2] [Medline: 17710487]
70. Selvander M, Åsman P. Virtual reality cataract surgery training: learning curves and concurrent validity. Acta Ophthalmol 2012 Aug;90(5):412-417 [FREE Full text] [doi: 10.1111/j.1755-3768.2010.02028.x] [Medline: 21054818]
71. Solyar A, Cuellar H, Sadoughi B, Olson TR, Fried MP. Endoscopic Sinus Surgery Simulator as a teaching tool for anatomy education. Am J Surg 2008 Jul;196(1):120-124. [doi: 10.1016/j.amjsurg.2007.06.026] [Medline: 18374890]

72. Sotto JAR, Ayuste EC, Bowyer MW, Almonte JR, Dofitas RB, Lapitan MCM, et al. Exporting simulation technology to the Philippines: a comparative study of traditional versus simulation methods for teaching intravenous cannulation. Stud Health Technol Inform 2009;142:346-351. [Medline: 19377182]
73. Sugand K, Akhtar K, Khatri C, Cobb J, Gupte C. Training effect of a virtual reality haptics-enabled dynamic hip screw simulator. Acta Orthop 2015 Oct 10;86(6):695-701 [FREE Full text] [doi: 10.3109/17453674.2015.1071111] [Medline: 26168925]
74. Suh IH, Mukherjee M, Park S, Oleynikov D, Siu K. Enhancing robot-assisted fundamental surgical proficiency using portable virtual simulator. Surgical Endoscopy and Other Interventional Techniques 2010;24(1):S686-S687.
75. Sun W, Konjg J, Li XY, Zhang JS. Application of operational simulation training system in the training of ophthalmic students. Int Eye Sci 2014;14(9):1567-1569.
76. Tanoue K, Ieiri S, Konishi K, Yasunaga T, Okazaki K, Yamaguchi S, et al. Effectiveness of endoscopic surgery training for medical students using a virtual reality simulator versus a box trainer: a randomized controlled trial. Surg Endosc 2008 Apr;22(4):985-990. [doi: 10.1007/s00464-007-9554-8] [Medline: 17710487]
77. Tanoue K, Yasunaga T, Konishi K, Okazaki K, Ieiri S, Kawabe Y, et al. Effectiveness of training for endoscopic surgery using a simulator with virtual reality: Randomized study. International Congress Series 2005 May 27;1281(11):515-520 [FREE Full text] [doi: 10.1016/j.ics.2005.03.367] [Medline: 31774400]
78. Torkington J, Smith SG, Rees BI, Darzi A. Skill transfer from virtual reality to a real laparoscopic task. Surg Endosc 2001 Oct 01;15(10):1076-1079. [doi: 10.1007/s004640000233] [Medline: 11727073]
79. Unger B, Tordon B, Pisa J, Hochman JB. Importance of Stereoscopy in Haptic Training of Novice Temporal Bone Surgery. Stud Health Technol Inform 2016;220:439-445. [Medline: 27046619]
80. Wilhelm DM, Ogan K, Roehrborn CG, Cadeddu JA, Pearle MS. Assessment of basic endoscopic performance using a virtual reality simulator. Journal of the American College of Surgeons 2002 Nov;195(5):675-681. [doi: 10.1016/s1072-7515(02)01346-7] [Medline: 31399794]
81. Youngblood PL, Srivastava S, Curet M, Heinrichs WL, Dev P, Wren SM. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. J Am Coll Surg 2005 Apr;200(4):546-551. [doi: 10.1016/j.jamcollsurg.2004.11.011] [Medline: 15804468]
82. Zeltser IS, Bensalah K, Tuncel A, Lucas SM, Jenkins A, Pearle MS. Training on the virtual reality laparoscopic simulator improves performance of an unfamiliar live surgical laparoscopic procedure: a randomized, controlled trial. J Endourol 2007;21(Suppl 1):A137.
83. Zhang L, Sankaranarayanan G, Arikatla VS, Ahn W, Grosdemouge C, Rideout JM, et al. Characterizing the learning curve of the VBLaST-PT(©) (Virtual Basic Laparoscopic Skill Trainer). Surg Endosc 2013 Oct;27(10):3603-3615 [FREE Full text] [doi: 10.1007/s00464-013-2932-5] [Medline: 23572217]
84. Zhao YC, Kennedy G, Yukawa K, Pyman B, O'Leary S. Can virtual reality simulator be used as a training aid to improve cadaver temporal bone dissection? Results of a randomized blinded control trial. Laryngoscope 2011 Apr;121(4):831-837. [doi: 10.1002/lary.21287] [Medline: 21433021]
85. Neumann E, Mayer J, Russo GI, Amend B, Rausch S, Deininger S, et al. Transurethral Resection of Bladder Tumors: Next-generation Virtual Reality Training for Surgeons. Eur Urol Focus 2019 Sep;5(5):906-911. [doi: 10.1016/j.euf.2018.04.011] [Medline: 29802051]
86. Alverson D, Saiki SJ, Kalishman S, Lindberg M, Mennin S, Mines J, et al. Medical students learn over distance using virtual reality simulation. Simul Healthc 2008;3(1):10-15. [doi: 10.1097/SIH.0b013e31815f0d51] [Medline: 19088637]
87. Battulga B, Konishi T, Tamura Y, Moriguchi H. The effectiveness of an interactive 3-dimensional computer graphics model for medical education. Interact J Med Res 2012 Jul 09;1(2):e2 [FREE Full text] [doi: 10.2196/ijmr.2172] [Medline: 23611759]
88. Blumstein G, Zukotynski B, Cevallos N, Ishmael C, Zoller S, Burke Z, et al. Randomized Trial of a Virtual Reality Tool to Teach Surgical Technique for Tibial Shaft Fracture Intramedullary Nailing. J Surg Educ 2020;77(4):969-977 [FREE Full text] [doi: 10.1016/j.jsurg.2020.01.002] [Medline: 32035854]
89. Courteille O, Ho J, Fahlstedt M, Fors U, Felländer-Tsai L, Hedman L, et al. Face validity of VIS-Ed: a visualization program for teaching medical students and residents the biomechanics of cervical spine trauma. Stud Health Technol Inform 2013;184:96-102. [Medline: 23400137]
90. Deladisma AM, Gupta M, Kotranza A, Bittner JG, Imam T, Swinson D, et al. A pilot study to integrate an immersive virtual patient with a breast complaint and breast examination simulator into a surgery clerkship. Am J Surg 2009 Jan;197(1):102-106. [doi: 10.1016/j.amjsurg.2008.08.012] [Medline: 19101251]
91. Drapkin ZA, Lindgren KA, Lopez MJ, Stabio ME. Development and assessment of a new 3D neuroanatomy teaching tool for MRI training. Anat Sci Educ 2015;8(6):502-509. [doi: 10.1002/ase.1509] [Medline: 25573020]
92. Flores R, DeMoss P, Klene C, Havlik RJ, Tholpady S. Digital Animation versus Textbook in Teaching Plastic Surgery Techniques to Novice Learners. Plastic and Reconstructive Surgery 2013;132(1):101e-109e. [doi: 10.1097/prs.0b013e3182910aa9]
93. Glittenberg CGO. Methods and advantages of the use of computer-assisted 3-D-design and multimedia teaching programs in the demonstration and education of the neuroophthalmological basics of the oculomotor system. Spektrum der Augenheilkunde 2003;17(6):242-246.

94. Gutiérrez F, Pierce J, Vergara VM, Coulter R, Saland L, Caudell TP, et al. The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. Stud Health Technol Inform 2007;125:155-160. [Medline: 17377256]
95. Hampton BS, Sung VW. Improving medical student knowledge of female pelvic floor dysfunction and anatomy: a randomized trial. Am J Obstet Gynecol 2010 Jun;202(6):601.e1-601.e8. [doi: 10.1016/j.ajog.2009.08.038] [Medline: 20430356]
96. Hisley KC, Anderson LD, Smith SE, Kavic SM, Tracy JK. Coupled physical and digital cadaver dissection followed by a visual test protocol provides insights into the nature of anatomical knowledge and its evaluation. Anat Sci Educ 2008 Jan;1(1):27-40. [doi: 10.1002/ase.4] [Medline: 19177376]
97. Hopkins R, Regehr G, Wilson TD. Exploring the Changing Learning Environment of the Gross Anatomy Lab. Academic Medicine 2011 Aug;86(7):883-888. [doi: 10.1097/acm.0b013e31821de30f] [Medline: 27307270]
98. Hu A, Shewokis PA, Ting K, Fung K. Motivation in computer-assisted instruction. Laryngoscope 2016 Aug;126 Suppl 6:S5-S13. [doi: 10.1002/lary.26040] [Medline: 27307270]
99. Kalet A, Song H, Sarpel U, Schwartz R, Brenner J, Ark T, et al. Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module. Medical Teacher 2012 Aug 23;34(10):833-839. [doi: 10.3109/0142159x.2012.706727] [Medline: 21412990]
100. Keedy AW, Durack JC, Sandhu P, Chen EM, O'Sullivan PS, Breiman RS. Comparison of traditional methods with 3D computer models in the instruction of hepatobiliary anatomy. Anat Sci Educ 2011 Sep;4(2):84-91. [doi: 10.1002/ase.212] [Medline: 21412990]
101. Khatib M, Hald N, Brenton H, Barakat MF, Sarker SK, Standfield N, et al. Validation of open inguinal hernia repair simulation model: a randomized controlled educational trial. Am J Surg 2014 Aug;208(2):295-301. [doi: 10.1016/j.amjsurg.2013.12.007] [Medline: 24581993]
102. Kockro RA, Amaxopoulou C, Killeen T, Wagner W, Reisch R, Schwandt E, et al. Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. Ann Anat 2015 Sep;201:91-98. [doi: 10.1016/j.aanat.2015.05.006] [Medline: 26245861]
103. Lorenzo-Alvarez R, Rudolphi-Solero T, Ruiz-Gomez MJ, Sendra-Portero F. Medical Student Education for Abdominal Radiographs in a 3D Virtual Classroom Versus Traditional Classroom: A Randomized Controlled Trial. AJR Am J Roentgenol 2019 Sep;213(3):644-650. [doi: 10.2214/AJR.19.21131] [Medline: 31287725]
104. Maresky HS, Oikonomou A, Ali I, Ditkofsky N, Pakkal M, Ballyk B. Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. Clin Anat 2019 Mar;32(2):238-243. [doi: 10.1002/ca.23292] [Medline: 30295333]
105. Motsumi MJ, Bedada AG, Ayane G. The role of Moodle-based surgical skills illustrations using 3D animation in undergraduate training. AJHPE 2019 Dec 12;11(4):149. [doi: 10.7196/ajhpe.2019.v11i4.1189]
106. Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. Med Educ 2006 Nov;40(11):1081-1087. [doi: 10.1111/j.1365-2929.2006.02611.x] [Medline: 17054617]
107. Patel V, Aggarwal R, Osinibi E, Taylor D, Arora S, Darzi A. Operating room introduction for the novice. Am J Surg 2012 Feb;203(2):266-275. [doi: 10.1016/j.amjsurg.2011.03.003] [Medline: 21703594]
108. Persky S, Eccleston CP. Medical student bias and care recommendations for an obese versus non-obese virtual patient. Int J Obes (Lond) 2011 May;35(5):728-735 [FREE Full text] [doi: 10.1038/ijo.2010.173] [Medline: 20820169]
109. Prinz A, Bolz M, Findl O. Advantage of three dimensional animated teaching over traditional surgical videos for teaching ophthalmic surgery: a randomised study. Br J Ophthalmol 2005 Nov;89(11):1495-1499 [FREE Full text] [doi: 10.1136/bjo.2005.075077] [Medline: 16234460]
110. Schutte B, de Goeij T, de Grave W, Koehorst AM. The Effects of Visual Genetics on the Learning of Students in a Problem Based Curriculum. In: Advances in Medical Education. Dordrecht, The Netherlands: Springer; 1997:336-338.
111. Scoville SA, Buskirk TD. Traditional and virtual microscopy compared experimentally in a classroom setting. Clin Anat 2007 Jul;20(5):565-570. [doi: 10.1002/ca.20440] [Medline: 17109440]
112. Succar T, Grigg J. A new vision for teaching ophthalmology in the medical curriculum: The virtual ophthalmology clinic. ascilite. 2010. URL: https://ascilite.org/conferences/sydney10/procs/Succar-poster.pdf [accessed 2022-04-06]
113. Succar T, Zebington G, Billson F, Byth K, Barrie S, McCluskey P, et al. The impact of the Virtual Ophthalmology Clinic on medical students' learning: a randomised controlled trial. Eye (Lond) 2013 Oct;27(10):1151-1157 [FREE Full text] [doi: 10.1038/eye.2013.143] [Medline: 23867718]
114. Yi X, Ding C, Xu H, Huang T, Kang D, Wang D. Three-Dimensional Printed Models in Anatomy Education of the Ventricular System: A Randomized Controlled Study. World Neurosurg 2019 May;125:e891-e901. [doi: 10.1016/j.wneu.2019.01.204] [Medline: 30743037]
115. Glittenberg C, Binder S. Using 3D computer simulations to enhance ophthalmic training. Ophthalmic Physiol Opt 2006 Jan;26(1):40-49. [doi: 10.1111/j.1475-1313.2005.00358.x] [Medline: 16390481]
116. Botezatu M, Hult H, Tessma MK, Fors UGH. Virtual patient simulation for learning and assessment: Superior results in comparison with regular course exams. Med Teach 2010;32(10):845-850. [doi: 10.3109/01421591003695287] [Medline: 20854161]

117. Deladisma AM, Cohen M, Stevens A, Wagner P, Lok B, Bernard T, Association for Surgical Education. Do medical students respond empathetically to a virtual patient? Am J Surg 2007 Jun;193(6):756-760. [doi: 10.1016/j.amjsurg.2007.01.021] [Medline: 17512291]
118. Dickerson R, Johnsen K, Raij A, Lok B, Stevens A, Bernard T, et al. Virtual patients: assessment of synthesized versus recorded speech. Stud Health Technol Inform 2006;119:114-119. [Medline: 16404028]
119. Foster A, Chaudhary N, Kim T, Waller JL, Wong J, Borish M, et al. Using Virtual Patients to Teach Empathy: A Randomized Controlled Study to Enhance Medical Students' Empathic Communication. Simul Healthc 2016 Jun;11(3):181-189. [doi: 10.1097/SIH.0000000000000142] [Medline: 26841278]
120. Foster A, Chaudhary N, Murphy J, Lok B, Waller J, Buckley PF. The Use of Simulation to Teach Suicide Risk Assessment to Health Profession Trainees-Rationale, Methodology, and a Proof of Concept Demonstration with a Virtual Patient. Acad Psychiatry 2015 Dec;39(6):620-629. [doi: 10.1007/s40596-014-0185-9] [Medline: 25026950]
121. Lehmann R, Lutz T, Helling-Bakki A, Kummer S, Huwendiek S, Bosse HM. Animation and interactivity facilitate acquisition of pediatric life support skills: a randomized controlled trial using virtual patients versus video instruction. BMC Med Educ 2019 Jan 05;19(1):7 [FREE Full text] [doi: 10.1186/s12909-018-1442-5] [Medline: 30611273]
122. Lehmann R, Thiessen C, Frick B, Bosse HM, Nikendei C, Hoffmann GF, et al. Improving Pediatric Basic Life Support Performance Through Blended Learning With Web-Based Virtual Patients: Randomized Controlled Trial. J Med Internet Res 2015 Jul 02;17(7):e162 [FREE Full text] [doi: 10.2196/jmir.4141] [Medline: 26139388]
123. McCoy L. Virtual patient simulations for medical education: Increasing clinical reasoning skills through deliberate practice. ERIC. 2015. URL: https://eric.ed.gov/?id=ED569014 [accessed 2022-04-06]
124. O'Rourke SR, Branford KR, Brooks TL, Ives LT, Nagendran A, Compton SN. The Emotional and Behavioral Impact of Delivering Bad News to Virtual versus Real Standardized Patients: A Pilot Study. Teach Learn Med 2020;32(2):139-149. [doi: 10.1080/10401334.2019.1652180] [Medline: 31437006]
125. Boeker M, Andel P, Vach W, Frankenschmidt A. Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. PLoS One 2013;8(12):e82328 [FREE Full text] [doi: 10.1371/journal.pone.0082328] [Medline: 24349257]
126. de Araujo TB, Silveira FR, Souza DLS, Strey YTM, Flores CD, Webster RS. Impact of video game genre on surgical skills development: a feasibility study. J Surg Res 2016 Mar;201(1):235-243. [doi: 10.1016/j.jss.2015.07.035] [Medline: 26850208]
127. 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]
128. Hedman L, Schlickum M, Felländer-Tsai L. Surgical novices randomized to train in two video games become more motivated during training in MIST-VR and GI Mentor II than students with no video game training. Stud Health Technol Inform 2013;184:189-194. [Medline: 23400154]
129. Kolga Schlickum M, Hedman L, Enochsson L, Kjellin A, Felländer-Tsai L. Transfer of systematic computer game training in surgical novices on performance in virtual reality image guided surgical simulators. Stud Health Technol Inform 2008;132:210-215. [Medline: 18391288]
130. Lagro J, van de Pol MHJ, Laan A, Huijbregts-Verheyden FJ, Fluit LCR, Olde Rikkert MGM. A randomized controlled trial on teaching geriatric medical decision making and cost consciousness with the serious game GeriatriX. J Am Med Dir Assoc 2014 Dec;15(12):957.e1-957.e6. [doi: 10.1016/j.jamda.2014.04.011] [Medline: 24913210]
131. Middeke A, Anders S, Raupach T, Schuelper N. Transfer of Clinical Reasoning Trained With a Serious Game to Comparable Clinical Problems: A Prospective Randomized Study. Simul Healthc 2020 Apr;15(2):75-81. [doi: 10.1097/SIH.0000000000000407] [Medline: 32044851]
132. 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 Mar 09;22(3):e16987 [FREE Full text] [doi: 10.2196/16987] [Medline: 32149711]
133. Sward KA, Richardson S, Kendrick J, Maloney C. Use of a Web-based game to teach pediatric content to medical students. Ambul Pediatr 2008;8(6):354-359. [doi: 10.1016/j.ambp.2008.07.007] [Medline: 19084784]
134. Tubelo RA, Portella FF, Gelain MA, de Oliveira MMC, de Oliveira AEF, Dahmer A, et al. Serious game is an effective learning method for primary health care education of medical students: A randomized controlled trial. Int J Med Inform 2019 Oct;130:103944. [doi: 10.1016/j.ijmedinf.2019.08.004] [Medline: 31442848]
135. Boyle E, Kennedy AM, Traynor O, Hill ADK. Training surgical skills using nonsurgical tasks--can Nintendo Wii™ improve surgical performance? J Surg Educ 2011;68(2):148-154. [doi: 10.1016/j.jsurg.2010.11.005] [Medline: 21338974]
136. Chien JH, Suh IH, Park S, Mukherjee M, Oleynikov D, Siu K. Enhancing fundamental robot-assisted surgical proficiency by using a portable virtual simulator. Surg Innov 2013 Apr;20(2):198-203. [doi: 10.1177/1553350612458545] [Medline: 22956399]
137. Albrecht U, Folta-Schoofs K, Behrends M, von Jan U. Effects of mobile augmented reality learning compared to textbook learning on medical students: randomized controlled pilot study. J Med Internet Res 2013 Aug 20;15(8):e182 [FREE Full text] [doi: 10.2196/jmir.2497] [Medline: 23963306]

138. Bogomolova K, van der Ham IJM, Dankbaar MEW, van den Broek WW, Hovius SER, van der Hage JA, et al. The Effect of Stereoscopic Augmented Reality Visualization on Learning Anatomy and the Modifying Effect of Visual-Spatial Abilities: A Double-Center Randomized Controlled Trial. Anat Sci Educ 2020 Sep;13(5):558-567. [doi: 10.1002/ase.1941] [Medline: 31887792]
139. Küçük S, Kapakin S, Göktaş Y. Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. Anat Sci Educ 2016 Oct;9(5):411-421. [doi: 10.1002/ase.1603] [Medline: 26950521]
140. Leitritz MA, Ziemssen F, Suesskind D, Partsch M, Voykov B, Bartz-Schmidt KU, et al. Critical evaluation of the usability of augmented reality ophthalmoscopy for the training of inexperienced examiners. Retina 2014 Apr;34(4):785-791. [doi: 10.1097/IAE.0b013e3182a2e75d] [Medline: 24670999]
141. Lemke M, Lia H, Gabinet-Equihua A, Sheahan G, Winthrop A, Mann S, et al. Optimizing resource utilization during proficiency-based training of suturing skills in medical students: a randomized controlled trial of faculty-led, peer tutor-led, and holography-augmented methods of teaching. Surg Endosc 2020 Apr;34(4):1678-1687. [doi: 10.1007/s00464-019-06944-2] [Medline: 31286252]
142. Logishetty K, Western L, Morgan R, Iranpour F, Cobb JP, Auvinet E. Can an Augmented Reality Headset Improve Accuracy of Acetabular Cup Orientation in Simulated THA? A Randomized Trial. Clin Orthop Relat Res 2019 May;477(5):1190-1199 [FREE Full text] [doi: 10.1097/CORR.0000000000000542] [Medline: 30507832]
143. Moult E, Ungi T, Welch M, Lu J, McGraw RC, Fichtinger G. Ultrasound-guided facet joint injection training using Perk Tutor. Int J Comput Assist Radiol Surg 2013 Sep;8(5):831-836. [doi: 10.1007/s11548-012-0811-5] [Medline: 23329279]
144. Noll C, von Jan U, Raap U, Albrecht U. Mobile Augmented Reality as a Feature for Self-Oriented, Blended Learning in Medicine: Randomized Controlled Trial. JMIR Mhealth Uhealth 2017 Sep 14;5(9):e139 [FREE Full text] [doi: 10.2196/mhealth.7943] [Medline: 28912113]
145. Sugand K, Wescott RA, Carrington R, Hart A, van Duren BH. Training and Transfer Effect of FluoroSim, an Augmented Reality Fluoroscopic Simulator for Dynamic Hip Screw Guidewire Insertion: A Single-Blinded Randomized Controlled Trial. J Bone Joint Surg Am 2019 Sep 04;101(17):e88. [doi: 10.2106/JBJS.18.00928] [Medline: 31483408]
146. Vera AM, Russo M, Mohsin A, Tsuda S. Augmented reality telementoring (ART) platform: a randomized controlled trial to assess the efficacy of a new surgical education technology. Surg Endosc 2014 Dec;28(12):3467-3472. [doi: 10.1007/s00464-014-3625-4] [Medline: 24962856]
147. Yeo CT, Ungi T, Leung R, Moult E, Sargent D, McGraw R, et al. Augmented reality assistance in training needle insertions of different levels of difficulty. 2018 Presented at: SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling; March 13, 2018; Houston, Texas, US.
148. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital Education in Health Professions: The Need for Overarching Evidence Synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
149. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. Adv Simul (Lond) 2016;1:31 [FREE Full text] [doi: 10.1186/s41077-016-0033-y] [Medline: 29450000]

Abbreviations

AR: augmented reality
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
ER: extended reality
MR: mixed reality
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
VR: virtual reality
WHO: World Health Organization

Edited by N Zary; submitted 13.04.21; peer-reviewed by R Lundin, S Gallagher; comments to author 13.09.21; revised version received 20.09.21; accepted 15.12.21; published 13.04.22.

Please cite as:
Tudor Car L, Kyaw BM, Teo A, Fox TE, Vimalesvaran S, Apfelbacher C, Kemp S, Chavannes N
Outcomes, Measurement Instruments, and Their Validity Evidence in Randomized Controlled Trials on Virtual, Augmented, and Mixed Reality in Undergraduate Medical Education: Systematic Mapping Review
JMIR Serious Games 2022;10(2):e29594
URL: https://games.jmir.org/2022/2/e29594
doi:10.2196/29594
PMID:35416789

Review

Virtual Reality Applications in Chronic Pain Management: Systematic Review and Meta-analysis

Lisa Goudman[1,2,3,4,5*], PhD; Julie Jansen[1,2*], MSc; Maxime Billot[6], PhD; Nieke Vets[1,2], MSc; Ann De Smedt[2,3,7], MD, PhD; Manuel Roulaud[6], PhD; Philippe Rigoard[6,8,9], MD, PhD; Maarten Moens[1,2,3,4,10], MD, PhD

[1]Department of Neurosurgery, Universitair Ziekenhuis Brussel, Jette, Belgium
[2]STIMULUS, Vrije Universiteit Brussel, Jette, Belgium
[3]Center for Neurosciences, Vrije Universiteit Brussel, Jette, Belgium
[4]Pain in Motion Research Group, Department of Physiotherapy, Human Physiology and Anatomy, Faculty of Physical Education and Physiotherapy, Jette, Belgium
[5]Research Foundation—Flanders, Brussel, Belgium
[6]PRISMATICS, Poitiers University Hospital, Poitiers, France
[7]Department of Physical Medicine and Rehabilitation, Universitair Ziekenhuis Brussel, Jette, Belgium
[8]Department of Spine Surgery & Neuromodulation, Poitiers University Hospital, Poitiers, France
[9]Pprime Institute, University of Poitiers, Chasseneuil-du-Poitou, France
[10]Department of Radiology, Universitair Ziekenhuis Brussel, Jette, Belgium
[*]these authors contributed equally

Corresponding Author:
Lisa Goudman, PhD
STIMULUS
Vrije Universiteit Brussel
Laarbeeklaan 103
Jette, 1090
Belgium
Phone: 32 2477 5514
Email: lisa.goudman@gmail.com

Abstract

Background: Virtual reality (VR) is a computer technology that immerses a user in a completely different reality. The application of VR in acute pain settings is well established. However, in chronic pain, the applications and outcome parameters influenced by VR are less clear.

Objective: This review aimed to systematically identify all outcome parameters that are reported in relation to VR in patients with chronic pain.

Methods: A total of 4 electronic databases (PubMed, Scopus, Web of Science, and Embase) were searched for relevant studies. Multilevel random-effect meta-analyses were performed, whereby the standardized mean difference was chosen as the effect size to denote the difference between measurements before and after a VR intervention.

Results: The initial database search identified 1430 studies, of which 41 (2.87%) were eventually included in the systematic review. Evidence has been found for the effects of VR on pain, functioning, mobility, functional capacity, psychological outcomes, quality of life, neuropsychological outcomes, and physical sensations. The overall effect size (a total of 194 effect sizes from 25 studies) based on a three level meta-analysis was estimated at 1.22 (95% CI 0.55-1.89; z=3.56; P<.001), in favor of improvements after a VR intervention. When categorizing effect sizes, the overall effect sizes were reported as follows: 1.60 (95% CI 0.83-2.36; z=4.09; P<.001) for the effect of VR on pain (n=31), 1.40 (95% CI 0.13-2.67; z=2.17; P=.03) for functioning (n=60), 0.49 (95% CI −0.71 to 1.68; z=0.80; P=.42) for mobility (n=24), and 0.34 (95% CI −1.52 to 2.20; z=0.36; P=.72) for functional capacity (n=21).

Conclusions: This systematic review revealed a broad range of outcome variables influenced by an intervention of VR technology, with statistically significant pain relief and improvements in functioning. These findings indicate that VR not only has applications in acute pain management but also in chronic pain settings, whereby VR might be able to become a promising first-line intervention as complementary therapy for patients with chronic pain.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42021227016; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=227016

(JMIR Serious Games 2022;10(2):e34402) doi:10.2196/34402

KEYWORDS

virtual reality; chronic pain; systematic review; multilevel meta-analysis; immersive technologies; clinical outcomes; mobile phone

Introduction

Background

After the recent revision of the definition of *pain* in 2020 by the International Association for the Study of Pain, *chronic primary pain* is defined as pain in 1 or more anatomical regions that (1) persists or recurs for >3 months and (2) is associated with significant emotional distress (eg, anxiety, anger, frustration, or depressed mood) or significant functional disability (interference in activities of daily life and participation in social roles), and (3) the symptoms are not better accounted for by another diagnosis [1]. Approximately 19% of adults complain of chronic pain [2], resulting in significant physical and mental burden as well as significant economic and social consequences [3,4]. Pain management for chronic pain conditions that are incorporated within chronic primary pain often entails a multidisciplinary or interdisciplinary pain management program relying on a biopsychosocial approach [5], mostly with a focus on functional restoration [6]. For the management of long-term pain, public opinion is currently in strong favor of self-management strategies as a first-line effective strategy to engage patients in actively managing their own health status [7-9]. The safety and cost-effectiveness of self-management programs have been proven; nevertheless, the effect sizes are small and not sustained in the long term [7,10]. In addition, the limited efficacy of pharmacotherapy in treating chronic pain and the long-term side effects of these pharmacological treatment options [11,12] have put a premium on novel nonpharmacological therapy options for chronic pain, thereby creating room for promising new pain management strategies such as virtual reality (VR) [13,14].

VR is characterized by an artificial computer-generated environment created to replace real-world sensory inputs [15]. This technology has evolved rapidly over the last 2 decades [16] and can be operationally described as simulations that use a combination of interaction devices and sensory display systems [17]. Within VR applications, an important distinction can be made between immersive and nonimmersive media, whereby they differ based on the participant's point of view and the experience that is produced during the application (ie, difference in spatial presence [18]) [19]. With immersive technology, participants view the full panorama, which enables the creation of a high sense of presence and immersion as if the participant is essentially inside the created environment [19]. 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 [20].

In both acute and chronic pain settings, the main idea of VR is to create distraction from the painful region [21]. VR can affect pain perception through immersive virtual environments, by occupying finite attentional resources, and by blocking external stimulation associated with the real environment and painful stimuli [22]. As distraction interventions work by competing for attention otherwise directed toward painful stimuli, pain tolerance and pain thresholds have been shown to increase under VR conditions [23]. Moreover, pain intensity, anxiety, and time spent thinking about pain decreased after VR distraction [24]. VR is thought to be more effective than traditional methods of distraction (eg, pleasant imagining, rhythmic cognitive activities, external focus of attention, and neutral imagining [25]) because of its immersive property, encompassing a patient's visual and auditory processing and even physical actions, which, in theory, demand more attention [26]. In addition to distraction as an underlying analgesic effect of VR, the long-term use of VR is expected to induce neuroplastic changes in the sensory and motor brain regions [27].

Objectives

Several systematic reviews have stressed the effectiveness of VR in the management of acute pain associated with medical procedures, wound debridement, and experimental pain [28-31]. The implementation of VR in chronic pain settings is still in its infancy compared with the widely accepted use of VR in acute pain settings [16]. Within the context of chronic pain, VR could be applied as an analgesic intervention and distraction method or could address pain-related behaviors [32]. Nevertheless, there is no consensus yet on which outcome measures VR has a positive effect in chronic pain settings. Therefore, this systematic review and meta-analysis aimed to evaluate the effect of VR on several outcome parameters related to the application of VR in patients with chronic pain.

Methods

Protocol and Registration

The study protocol was registered prospectively with PROSPERO (CRD42021227016). This systematic review and meta-analysis was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [33].

Search Strategy

A search strategy based on the PICO (Patient or Population, Intervention, Comparison, and Outcome; evidence-based search strategy focusing on patients or populations, interventions, comparisons, and outcomes) framework was developed [34]. The following research question was constructed: *The effects of virtual reality (intervention) on multiple outcome measurements (outcome) in patients with chronic pain* *(population)*. The component *control* was not defined for this question. The final search strategy was built by combining free and Medical Subject Headings terms. Within *Population*, *Intervention*, and *Outcome*, search terms were combined using the Boolean term OR. The Boolean term AND was used between the complete search terms for *Population*, *Intervention*, and *Outcome*. The search strategy was conducted using the following databases: PubMed, Embase, Scopus, and Web of Science from their outset until December 2020. No limitations were applied to the search. The full search strategy for PubMed is outlined in Multimedia Appendix 1.

Eligibility Criteria

Studies exploring the effects of VR in patients with chronic pain were eligible for inclusion. Studies with both adults and children were eligible if the study participants experienced chronic pain. Chronic pain was defined as pain that lasted for >3 months [35], including patients with primary and secondary chronic pain conditions. Studies in which the effect of VR was explored in healthy participants or in acute pain settings were excluded. For this intervention, there were no restrictions on the VR devices, and all types of VR were permitted, including (but not limited to) studies with head-mounted displays, video games, displays with body motion sensors, Nintendo Wii consoles, Xbox, PlayStation, and computers. Studies that did not use VR to explore its effects on patients with chronic pain were excluded. There were no limitations in the outcome measurements because the goal was to explore the effects of VR on all types of outcome variables that were recorded in this setting. Studies reporting in languages other than English, Dutch, French, or German were excluded. Publications available only in abstract form, conference abstracts, expert opinions, letters to the editor, study protocols, reviews, meta-analyses, or meeting reports were considered not suitable for inclusion.

Study Selection

After deduplication using the EndNote X9 reference manager (Clarivate), 2 reviewers (JJ and VN) independently screened all retrieved articles for titles and abstracts using the Rayyan web-based software (Rayyan Systems Inc) [36]. Subsequently, the same 2 reviewers performed the full-text screening (independently). If discrepancies occurred, consensus was sought through consultation and discussion with a third independent reviewer (ADS).

Data Extraction

All relevant information concerning the possible effects of VR in patients with chronic pain was synthesized in an a priori constructed evidence table. The following items were extracted from each of the remaining articles: author, publication year, country, study design, sample size (including sex distribution), underlying pathology, VR application (including duration, type of VR, and device), and reported outcome measurements. Data extraction was performed by one reviewer (JJ) and checked for correctness by another (LG). Any discrepancies were discussed in a consensus meeting with all the reviewers.

Risk of Bias Assessment

A modified version of the Downs and Black checklist was used to assess the quality of the included studies [37]. This instrument, consisting of 27 items, was developed to evaluate the methodological quality of several study designs, including randomized trials, nonrandomized trials, and observational studies [37]. All items were categorized into 5 subscales: reporting, external validity, bias, confounding, and power. Each item was given a score of 0 (*no*) or 1 (*yes*), except for item 5, where a score of 1 meant the item was partially presented and a score of 2 if a complete description was presented. The answer option *not applicable* was also available. The scoring of item 27, which refers to the power of the study, had been modified in the sense that it received a score of 1 if a power calculation was performed and 0 otherwise [38,39]. After scoring each individual item, all included studies were categorized as having poor, fair, good, or excellent quality based on the total score for further data synthesis. A total score of ≤14 out of 28 was considered poor quality, 15-19 was considered fair, 20-25 was considered good, and 26-28 was considered excellent methodological quality [40].

Statistical Procedure for the Meta-analysis

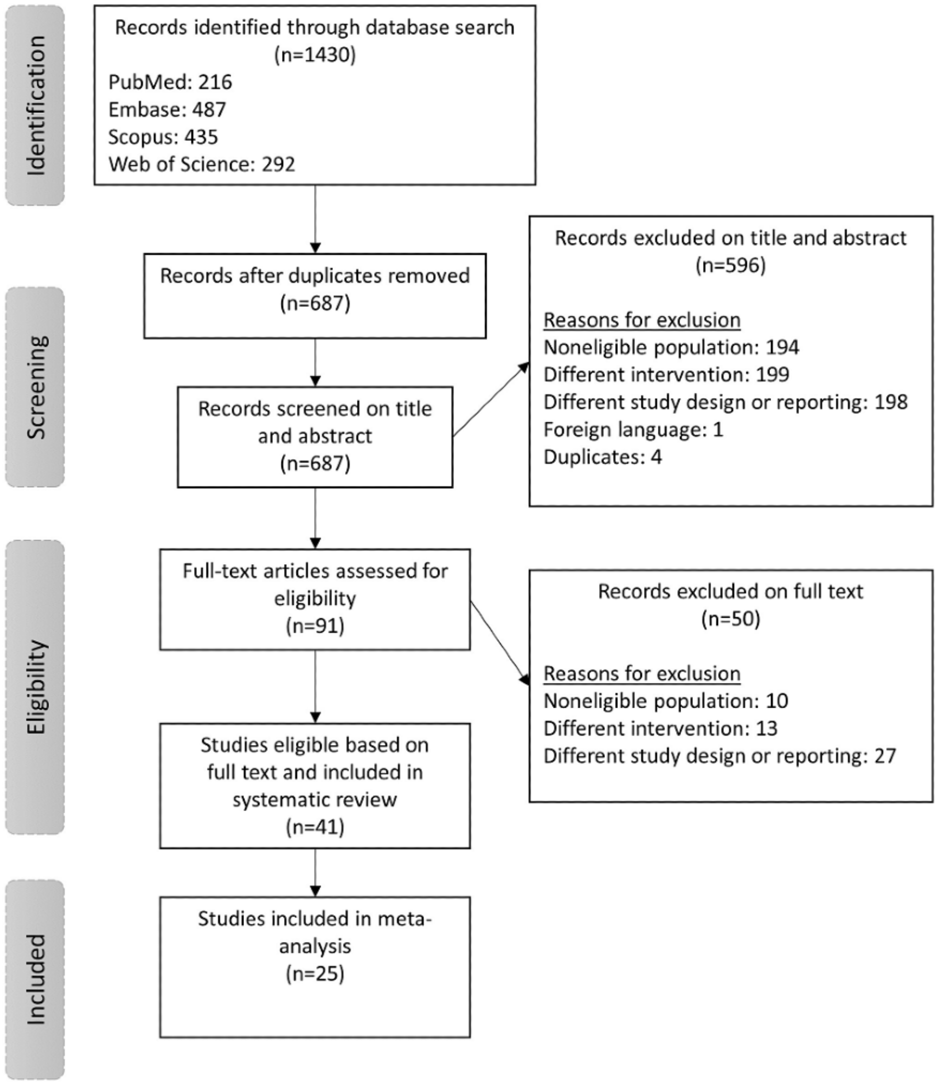
The standardized mean difference was chosen as the effect size to compare the differences between measurements before and after the VR intervention, calculated as the difference in sample means after VR minus before VR, divided by the SD before VR (g gain).nOnly sample SDs of the measurements before VR, instead of pooled SDs, were used in the calculation because they are not influenced by VR effects and are therefore more likely to be consistent across studies [41]. Correlations between measurements were estimated to be 0.5. As standardized mean difference does not correct for differences in the direction of the scale, the mean values from outcomes that have a higher score when they reveal an improvement are multiplied by −1 to ensure that all the scales point in the same direction. Thus, the effect size measures are positive if the data indicate a desirable effect of VR. Several studies included more than one outcome measurement, wherefore a random-effect 3-level meta-analysis was performed [42]. Thus, a 3-level meta-analysis model was fitted with sampling variance at the first level, within-study variance at the second level, and between-study variance at the third level [42]. Meta-analyses were performed for all effect sizes combined and for effect sizes categorized according to the type of outcome measurements (ie, pain, functioning, mobility, and functional capacity). Meta-analyses were performed using R Studio (R Foundation for Statistical Computing) version 1.4.1106 (R version 4.0). P≤.05 was considered statistically significant. Figures were created based on the code provided by Fernández-Castilla et al [43] to visually present the results of meta-analyses of multiple outcomes. Only studies with sufficient data were included in this meta-analysis. If insufficient information was provided, the authors were contacted to provide more in-depth data.

Results

Study Selection

A total of 1430 articles were identified through 4 database searches (Figure 1). After removing duplicates, 687 (48.04%) articles were selected for screening. After screening titles and abstracts, 13.2% (91/687) of articles remained eligible for full-text screening. Studies were excluded based on study design or reporting (n=198), no patients with chronic pain (n=194), no VR as intervention (n=199), duplicates (n=4), or other languages (n=1). The percentage of agreement on title and abstract screening between both reviewers was 92.8%. After full-text screening, 41 articles were included in this systematic review. The primary reasons for study exclusion were no patients with chronic pain (n=10), no VR as intervention (n=13), and a different study design or reporting (n=27). The percentage agreement between both reviewers for full-text screening was 95.7%. Data from 25 studies were included in the meta-analysis.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Study Characteristics

A comprehensive overview of the characteristics of the included studies is presented in Table 1. The earliest study included in this review was published in 2008 and the most recent study, in 2020. In terms of study design, the following studies were included: 18 quasi-experimental study designs, 16 randomized controlled trials, and 7 case series.

In terms of population, most studies (n=11) investigated adults with chronic pain without further specification of the specific type of chronic pain [44-54]. A total of 7 studies evaluated the effect of VR in patients with fibromyalgia [55-61], 6 studied patients with chronic low back pain [62-67], 5 studied patients with chronic neck pain [68-72], 4 studies evaluated VR in patients with (upper limb) complex regional pain syndrome [73-76], and 3 studied patients with phantom limb pain [77-79]. In addition, 1 study investigated pediatric patients with chronic pain [80], and 1 study investigated both adults and adolescents [73]. A total of 1232 patients were enrolled in 41 studies, of which 784 (63.64%) were women.

With regard to the VR application, the duration of the sessions ranged from 2×2 minutes with 30 seconds rest between sessions [64] up to sessions of 1 hour [60]. The frequency varied from 1 session [64] to 2 sessions per week for 24 weeks [60].

Table 1. Characteristics of the included studies.

Authors	Country	Design	Participants, N (number of female participants)	Population	Duration of the intervention	VR[a] device and application	Type of VR	Outcome measurements
Alemanno et al [62]	Italy	QE[b]	20 (11)	Adults with chronic low back pain	1 hour per session for 12 sessions over 4-6 weeks	VR Rehabilitation System, Khymeia, Italy	Nonimmersive	Roland and Morris Disability Questionnaire, Repetition Index, SF-36 Short Form Health Survey, NRS[c] Pain, McGill Pain Questionnaire, Brief Pain Inventory, Beck Depression Inventory, and neuropsychological evaluations
Amin et al [44]	Canada	RCT[d]	30 (13)	Adult patients with chronic pain	2×10 minutes	Oculus Rift DK2 and a Cardboard VR to play In-Mind	Immersive	Present and retroactive pain intensity
Botella et al [55]	Spain	QE	6 (6)	Adult patients with fibromyalgia	7 weeks with ten 2-hour sessions: sessions 1 through 6 were delivered twice a week for 3 weeks, whereas sessions 7 through 10 were delivered weekly for 4 weeks	2 PCs, a large projection screen, 2 projectors, a wireless pad, and a speaker system. Application used: EMMA[e] World	Nonimmersive	Beck Depression Inventory II, Positive and Negative Affect Schedule, Chronic Pain Coping Inventory, Fibromyalgia Impact Questionnaire, and VR Satisfaction Scale
Brown et al [63]	United States	RCT	45 (27)	Adults with chronic low back pain, receiving spinal injections	One 5-minute session	VR immersive format in the Oculus GoÒ headset	Immersive	Anxiety thermometer, NRS pain, Patient-Reported Outcomes Measurement Information System, Modified Oswestry Disability Index, and VR Symptom Questionnaire
Chau et al [73]	United States	CS[f]	8 (7)	Upper limb complex regional pain syndrome (adults and 1 adolescent)	10 sessions (1-3 per week) with 45-60 minutes each session	HTC Vive VR System (virtual 3D kitchen environment)	Immersive	Short Form McGill Pain Questionnaire, VAS[g], Wong-Baker Faces Scale, and Subjective feedback
Collado-Mateo et al [56]	Spain	RCT	83 (83)	Adult patients with fibromyalgia	Twice a week for 1 hour per session over 8 weeks	An exergame called VirtualEx-FM based on Microsoft Kinect	Nonimmersive	Timed up and go test, functional reach, Clinical Test of Sensory Integration of Balance, fear of falling using VAS
Darnall et al [81]	United States	RCT	88 (ND[h])	Adults with chronic nonmalignant low back pain or fibromyalgia	21 days	Oculus Go VR Headset (visual biofeedback in relaxation training)	Immersive	Defense and Veterans Pain Rating Scale; pain interference on activity, mood, sleep, and stress; pain catastrophizing scale; pain self-efficacy; global impression of change; satisfaction with treatment; and motion sickness and nausea

Authors	Country	Design	Participants, N (number of female participants)	Population	Duration of the intervention	VR[a] device and application	Type of VR	Outcome measurements
Fowler et al [45]	United States	QE	16 (3)	Veterans with chronic pain	19-day program with 20 minutes per VR session	Oculus Rift and Samsung Oculus Gear VR	Immersive	VR feasibility, Pain Outcomes Questionnaire-VA, Fear of Daily Activities, questionnaire for kinesiophobia, Pain Catastrophizing Scale, and patient-specific functional scale
Garcia-Palacios et al [57]	Spain	RCT	61 (61)	Adults with fibromyalgia	Six 2-hour group sessions delivered twice a week	2 PC computers with an EMMA VR environment	Nonimmersive	Fibromyalgia Impact Questionnaire, Brief Pain Inventory, Chronic Pain Coping Inventory, Beck Depression Inventory II, Quality of Life Index, and acceptability and satisfaction
Garrett et al [46]	Canada	CS	8 (6)	Adults with chronic pain conditions	1 month with 12 sessions of 30 minutes each	Oculus Rift DK2 1100 field of view stereoscopic HMD[i]	Immersive	NRS pain, Brief Pain Inventory, Self-Administered Leeds Assessment of Neuropathic Symptoms and Signs, cybersickness, and individual interviews
Griffin et al [80]	United States	QE	17 (13)	Pediatric patients with chronic pain	1-8 sessions of 30 minutes (once a week)	HTC Vive VR system	Immersive	Presence, child daily questionnaire, and interviews
Gromala et al [47]	Canada	RCT	13 (7)	Adult patients with chronic pain	1 moment	Technology's DeepStream VR with virtual meditative walk	Immersive	NRS pain
Guarino et al [48]	Italy	CS	11 (8)	Adult patients with chronic pain	8 VR sessions of 30 minutes, 2 times a week	VR scenarios were run in a PC, and the environments were visualized on a monitor	Nonimmersive	McGill Pain Questionnaire, Brief Pain Inventory Severity and Interference, State trait Anxiety Inventory, Beck Depression Inventory, VAS, and subjective units of distress scales
Harvie et al [68]	Australia	CS	12 (3)	Adults with chronic neck pain	36-70 days with 10 minutes of VR twice a day	Samsung Gear VR system with Motor Offset Visual Illusion	Immersive	Pain threshold and NRS pain
Herrero et al [58]	Spain	QE	40 (40)	Adults with fibromyalgia	3 sessions of 20 minutes each	2 PC computers, a 3.4-m screen made of reflective material, 2 projectors, and a Dolby 7.1 surround sound audio system with EMMA	Nonimmersive	Mood State, NRS pain, NRS fatigue, NRS motivation, NRS self-efficacy, and NRS of several emotions (joy, sadness, anger, surprise, anxiety, calmness, and vigor/energy)

Authors	Country	Design	Participants, N (number of female participants)	Population	Duration of the intervention	VR[a] device and application	Type of VR	Outcome measurements
House et al [82]	United States	QE	12 (12)	Adults with persistent pain in shoulder and arm following postsurgical breast cancer	20-50 minutes twice week for 8 weeks	BrightArm Duo technology with 3D custom integrative rehabilitation games	Nonimmersive	NRS pain, upper limb range of motion, Beck Depression Inventory II, Neurophysiological Assessment Battery, Revised Hopkins Verbal Learning Test, Revised Brief Visuospatial Memory Test, and Trail Making Test
Jin et al [49]	Canada	RCT	20 (16)	Adults with chronic pain	1 moment, 35-45 minutes per participant with 10-minute VR	Oculus Rift DK2 with Cryoslide game	Immersive	Pain intensity (VAS) and distraction
Jones et al [50]	United States	QE	30 (20)	Adults with chronic pain conditions	Single 5-minute exposure to VR	Oculus rift DK2 and Deepstream VR with cool!	Immersive	NRS pain, engagement, and side effects
Matamala-Gomez et al [74]	Spain	QE	19 (14)	Adults with neuropathic chronic pain in the upper limb	Single session of 55 minutes	HMD rift development kit 2, Oculus	Immersive	Ownership, agency, mental representation, NRS pain, and VR questionnaire
Matheve et al [64]	Belgium	RCT	84 (54)	Adults with chronic low back pain	2×2 minutes with 30 seconds of rest in between	Valedo Pro, Hocom	Nonimmersive	NRS pain, Roland and Morris Disability Questionnaire, Pain Catastrophizing Scale, Tampa Scale for Kinesiophobia, pain intensity, time spent thinking about pain, and pelvic tilts
Monteiro et al [65]	Portugal	RCT	34 (34)	Adults with chronic low back pain	8 weeks with sessions 3 times a week for 90 minutes	Nintendo Wii motion and Wii balance board	Nonimmersive	NRS pain, balance, sit-to-stand test, and profile of mood states
Mortensen et al [59]	Denmark	QE	15 (15)	Adults with fibromyalgia	15 sessions of 30 minutes	Motion-controlled video games, Wii, Ps3, and Xbox Kinect	Nonimmersive	Pain VAS, Brief Fatigue Inventory, Activities of Daily Life Questionnaire, test of playfulness, and interviews
Mouraux et al [76]	United States and Belgium	QE	22 (12)	Adults with chronic neuropathic pain in unilateral upper extremity	5 sessions of 20 minutes over 1 week	3D augmented reality system with a 3D display (Kit Nvidia 3D Vision) and 3D camera (Xbox 360 Kinect)	Nonimmersive	Pain VAS, McGill Pain Questionnaire, and Douleur Neuropathique 4 Questions
Ortiz-Catalan et al [77]	Sweden and Slovenia	QE	14 (ND)	Adults with chronic intractable phantom limb pain	Twice per week, 12 sessions of 2 hours each	Neuromotus and Integrum AB	Nonimmersive	NRS intensity, frequency, duration, and quality of phantom limb pain; Pain Rating Index Scale; Short form of McGill questionnaire; and interviews

Authors	Country	Design	Participants, N (number of female participants)	Population	Duration of the intervention	VR[a] device and application	Type of VR	Outcome measurements
Pamment and Aspell [51]	United Kingdom	QE	18 (12)	Adults with chronic pain	4 conditions, each lasting 2 minutes	HMD (WRAP 1200, Vuzix) connected to a video camera	Immersive	McGill Pain Questionnaire and illusion and control questions
Phoon Nguyen et al [83]	Australia	QE	9 (7)	Adults with burning mouth syndrome	1 session of 3 experimental conditions of 1 minute each	MIRAGE-mediated reality system; computer screen displaying live digitally manipulated video feed of their own face	Nonimmersive	Pain VAS, Wong-Baker Faces Pain Rating Scale, and VAS burning pain/sensation
Rezaei et al [69]	Iran	RCT	44 (ND)	Adults with chronic neck pain	8 training sessions over 4 weeks with 21 minutes each	Cervigame head mouse extreme	Nonimmersive	Pain VAS, Neck Disability Index, and Y-balance test
Rutledge et al [78]	United States	QE	14 (1)	Adults with phantom limb pain	57 treatment sessions with 40-60 minutes per session	VR treatment based on mirror therapy with an Oculus Rift Headset	Immersive	Phantom Limb Pain Questionnaire, Trinity Amputation and Prosthetic Experience Scale-Short Form-12, Patient Health Questionnaire-9, Posttraumatic Stress Disorder Checklist-Military version, and Present Questionnaire
Sarig-Bahat et al [70]	Australia	RCT	32 (22)	Adults with chronic neck pain	4-6 supervised intervention sessions for 30 minutes each over a period of 5 weeks+home training sessions of 30 minutes for at least three times a week	HMD with a Wrap 1200VR by Vuzix	Immersive	VAS pain, Neck Disability Index, Tampa Scale for Kinesiophobia, static and functional balance, satisfaction, global perceived effect, and range of motion
Sarig-Bahat et al [71]	Australia and Israel	RCT	90 (63)	Adults with chronic neck pain	5 minutes, 4 times a day, 4 days per week, for 4 weeks	Oculus Rift DK1 HMD with 3D motion tracking	Immersive	Neck disability Index, global perceived effect, VAS pain, self-rated health status in the European life quality questionnaire, velocity, Tampa scale for kinesiophobia
Sato et al [75]	Japan	CS	5 (4)	Adults with complex regional pain syndrome	5-8 weeks with 1 session each week	FASTRAK and cyberglove, PC desktop with a CyberGlove as hand input, FASTRAK as real-time position and motion tracker, and computer screen	Nonimmersive	VAS pain and range of motion

Authors	Country	Design	Participants, N (number of female participants)	Population	Duration of the intervention	VR[a] device and application	Type of VR	Outcome measurements
Shahrbanian et al [52]	Canada	QE	12 (5)	Adult stroke patients with chronic pain	1 moment with 3-5 minutes of each VR condition	HMD Kaiser Optical System with Nvidia Quadro FX 4500 graphics card	Immersive	Pain threshold, engagement, VAS mood, and NRS pain
Solca et al [84]	Switzerland	QE	15 (5)	Adults with chronic leg pain with a spinal cord implant	1 moment	Oculus Rift CV1 with RealiSM software	Immersive	Analgesia and embodiment
Tejera et al [72]	Spain	RCT	44 (23)	Adults with nonspecific chronic neck pain	2 sessions per week for 4 weeks	VR Vox Play glasses used with an HMD clamping system	Immersive	VAS pain, pain pressure threshold, temporal summation, range of motion, neck disability index, pain catastrophizing scale, Tampa scale for kinesiophobia, fear-avoidance beliefs questionnaire, and Pain Anxiety Symptoms Scale
Thomas et al [66]	United States	RCT	53 (ND)	Adults with chronic low back pain	3 consecutive days	Samsung 3D shutter glasses with Vizard software	Nonimmersive	Changes in lumbar spine flexion, VAS expectations of pain and harm, and game experience survey
Tong et al [79]	China	CS	5 (0)	Adults with phantom limb pain	10 sessions over 6 weeks	Immersive room-scale VR system and HMD from HTC Vive with Unity 3D	Immersive	Short Form McGill Pain Questionnaire, VAS pain, NRS embodiment, NRS ownership, and hospital anxiety and depression scale
Trujillo et al [67]	United States	CS	2 (0)	Adults with chronic low back pain	7 sessions with 2 sessions per week of 30-45 minutes each	Virtual embodiment training with Virtual Embodiment Training (KVET); HTC Vive with a VR HMD	Immersive	Simulator Sickness Questionnaire; VAS pain and Pain Catastrophizing Scale
Villafaina et al [60]	Spain	RCT	55 (55)	Adults with fibromyalgia	2 sessions of 1 hour per week for 24 weeks	VirtualEx-FM	Nonimmersive	Chair-stand test, 10-step stair test, 6-minute walk test, Fibromyalgia impact questionnaire, and International Physical Activity Questionnaire
Villafaina et al [61]	Spain	RCT	55 (55)	Adults with fibromyalgia	2 sessions of 1 hour per week for 24 weeks	VirtualEx-FM	Nonimmersive	Electroencephalography
Wiederhold et al [53]	United States	QE	31 (ND)	Adults with chronic pain	5-minute pain focus session followed by a 20-minute intervention session	Mobile phone VR therapy and an HMD	Immersive	Simple Descriptive Pain Intensity Scale, Numerical Pain Intensity Scale, VAS pain, and physiological measures (heart rate, peripheral skin temperature, respiration, and skin conductance)

Authors	Country	Design	Participants, N (number of female participants)	Population	Duration of the intervention	VR[a] device and application	Type of VR	Outcome measurements
Wiederhold et al [54]	Belgium and United States	QE	40 (ND)	Adults with chronic pain	15-minute exposure session	VR exposure while wearing HMD	Immersive	Pain reduction, pain focus, and skin temperature

[a]VR: virtual reality.

[b]QE: quasi-experimental.

[c]NRS: Numeric Rating Scale.

[d]RCT: randomized controlled trial.

[e]EMMA: Engaging Media for Mental Health Applications.

[f]CS: case series.

[g]VAS: visual analog scale.

[h]ND: not displayed.

[i]HMD: head-mounted display.

Risk of Bias

The range of the total scores on the Downs and Black checklist varied between 8 out of 28 and 25 out of 28. Overall, 5 (12%) of the 41 included studies scored poor on the risk of bias assessment (total score ≤14), 17 (42%) had a fair score (total score between 15 and 19), and 19 (46%) had good quality (total score between 20 and 25). The total scores on the risk of bias assessment are presented in Table 2. Multimedia Appendix 2 presents the full results of the risk of bias assessment.

In the 18 studies with a quasi-experimental design, low scores were found for the external validity subscale. Only 6 (33%) of the 18 studies had an accurate score on the item of whether patients were representative of the entire population from which they were recruited, and the item concerning representativeness of staff, places, and facilities where the patients were treated was only considered representative of the treatment most patients received in 7 (39%) of 18 studies. Several items that evaluated internal validity performed poorly, including the attempt to blind study participants, which was only evaluated efficiently in 1 (6%) of the 18 studies. No study has attempted to blind outcome assessors. Recruiting patients over the same period was only scored satisfactory in 3 (17%) studies, no studies randomized patients, and adjustments for confounding were only performed in 6 (33%) of 18 studies.

In the 16 randomized controlled trials, several items evaluating internal validity received low scores. Of the 16 studies, only 1 (6%) study provided information on attempts to blind study participants, whereas 5 (31%) studies provided information on attempts to blind outcome assessors. Information regarding adjustments for confounding was provided in only 7 (44%) studies.

A total of 7 case series were included; only 2 (29%) studies reported the estimates of random variability. For the 3 items evaluating external validity, most studies performed poorly with representativeness ranging from 14% (1/7) to 43% (3/7). Of the 7 studies, on the subscale for internal validity (bias), only 1 (14%) study attempted to blind the study participants, and none attempted to blind the outcome assessor. Patients were not recruited in the same period within a study, and only 1 (14%) of 7 studies used randomization (no concealed randomization). Adjustments for confounding were explored in 2 (29%) studies, and a power calculation was reported in 3 (43%) of the 7 studies.

Table 2. Total score on the risk of bias assessment.

Study[a]	Total score (out of 28)	Category
Alemanno et al [62]	22	Good
Amin et al [44]	19	Fair
Botella et al [55]	16	Fair
Brown et al [63]	24	Good
Chau et al [73]	16	Fair
Collado-Mateo et al [56]	24	Good
Darnall et al [81]	25	Good
Fowler et al [45]	21	Good
Garcia-Palacios et al [57]	23	Good
Garrett et al [46]	18	Fair
Griffin et al [80]	20	Good
Gromala et al [47]	18	Fair
Guarino et al [48]	10	Poor
Harvie et al [68]	20	Good
Herrero et al [58]	18	Fair
House et al [82]	19	Fair
Jin et al [49]	18	Fair
Jones et al [50]	16	Fair
Matamala-Gomez et al [74]	19	Fair
Matheve et al [64]	24	Good
Monteiro et al [65]	21	Good
Mortensen et al [59]	18	Fair
Mouraux et al [76]	14	Poor
Ortiz-Catalan et al [77]	18	Fair
Pamment and Aspell [51]	20	Good
Phoon Nguyen et al [83]	21	Good
Rezaei et al [69]	25	Good
Rutledge et al [78]	18	Fair
Sarig-Bahat et al [70]	23	Good
Sarig-Bahat et al [71]	21	Good
Sato et al [75]	12	Poor
Shahrbanian et al [52]	17	Fair
Solca et al [84]	19	Fair
Tejera et al [72]	23	Good
Thomas et al [66]	25	Good
Tong et al [79]	16	Fair
Trujillo et al [67]	17	Fair
Villafaina et al [60]	23	Good
Villafaina et al [61]	22	Good
Wiederhold et al [53]	11	Poor
Wiederhold et al [54]	8	Poor

[a]Each study was scored on all 27 items of the modified Downs and Black checklist. On the basis of the total score, all included studies were categorized as presenting poor, fair, good, or excellent quality. A total score of ≤14 out of 28 was considered poor quality, 15-19 was considered fair, 20-25 was considered good, and 26-28 was considered excellent methodological quality.

Interventions

A total of 23 studies used immersive VR techniques, and 18 used non-immersive techniques. Within the category of immersive techniques, VR games (7/23, 30%), mindfulness-based interventions (6/23, 28%), practical exercises (6/23, 28%), and visual illusions (4/23, 17%) were used. In the category of nonimmersive techniques, of the 18 studies, 6 (33%) used exergames, 3 (17%) used an avatar or exoskeleton, and 9 (50%) studies used a television or PC screen. A more detailed description of immersive techniques is provided in further sections, followed by an in-depth explanation of nonimmersive VR techniques.

When VR was applied as a game, a broad range of VR games were used, including shooter games [44], a game of grasping where participants had to stomp fruit [80], a game with sliding in an icy cave during which participants should hit creatures with snowballs [49], a game in which the user travels through a landscape with interaction [50], a game with a visualization of a red airplane that could be controlled by head motion and the user could hit targets [70,71], and a game of pushing a ball of the table and shooting a basketball [79].

When VR was used in the context of mindfulness-based or relaxation treatments, several applications were used, including a 5-minute relaxation video [63], sessions to support patients in learning self-management skills based on cognitive behavior therapy principles in which the VR headset was used for visual biofeedback in relaxation training [81], exploratory environments [46,47], interaction with simulation graphics and exploration of virtual worlds [53], and relaxing environments and sounds [54].

In addition, immersive VR was also used to practice exercises in a kitchen environment in which participants had to perform tasks representative of daily activities [73]. In addition, minimal exercises such as neck and head movements and larger exercises such as torso and upper extremity movements were provided [45,68]. Participants could bicycle on a pedaller through VR environments [78], practice neck exercises through the illusion of diving with sounds of the sea [72], and perform exercises based on the principles of graded motor imagery to relearn associations to pain and improve their function [67].

Finally, other possibilities of immersive VR were seeing an illusion of the affected body part through an avatar [74], a visualization of their own back and synchronously or asynchronously tapping with a wooden stick [51], experiencing hot and cold stimuli through a snow world environment and a canyon environment, respectively [52], and a visual illumination of a circumscribed skin region in the VR corresponding with Spinal Cord Stimulation [84].

In contrast to immersive applications, several studies have used nonimmersive VR techniques. The studies (n=6) that made use of exergames consisted of a variety of exercises, such as aerobic sessions; postural control and coordination; exercises to improve mobility skills, fitness, ability, and balance [56,60,61]; Wii Fit Plus workouts [65]; a sports game package [59]; and an exercise in which a real-time 3D image of the individual's moving, nonaffected body part was presented, and participants played a game to touch a few targets with the hand or fingers of the virtual affected upper extremity [76].

An avatar or exoskeleton was used in 3 studies as virtual rehabilitation and could help to teach patients to execute correct movements with the painful body parts to regain a correct body image [62] or with games for unimanual and bimanual motor, emotive, and cognitive training [82] or with pelvic tilt exercises where the VR was used as a game [64].

A television or PC screen was used in the following situations (n=9): support of a group cognitive behavior therapy containing specific content for developing relaxation and mindfulness skills, as an adjunct to the activity pacing component, with education, activity management, and relapse prevention, to induce positive emotions and promote motivation, self-efficacy, and behavior activation [48,55,57,58]. Other authors used a television or PC screen to practice motor execution by games, such as racing cars using phantom movements and matched random target postures of a virtual arm [66,77], a rabbit attempting to reach carrots and avoid obstacles [69], or a target-oriented motor control task where hand exercises consisted of reaching out, grasping, transferring, and placing [75]. In another study, patients with burning mouth syndrome watched their tongues on a computer screen with illusions and performed tongue movements [83].

Instruments of Outcome Measurements

Multimedia Appendix 3 provides a complete summary of the measurement instruments used to evaluate these outcomes. Of the 41 included studies, 35 (85%) evaluated pain-related outcomes [44-54,57-59,62-73,75-81,83,84]. Psychological outcomes, such as kinesiophobia and fear, mood, satisfaction, expectations of pain, pain focus, time spent thinking about pain, self-efficacy, emotions, motivation, stress, catastrophizing, acceptability, global impression of change, ownership, and agency, were measured in 19 studies [45,48,52,55-59,62-65,67,70-72,74,81,82]. Functional outcomes, including disability, physical comfort, strength, fitness, and sleep, were measured in 15 studies [45,55-57,59,60,62,64,65,69-72,77,82]. Functional capacity (evaluated by measuring balance, repetition index, step test, and composite value) was measured in 5 studies [56,62,65,69,70]. Mobility (range of motion) was measured in 4 studies [45,62,71,72]. Neuropsychological functions were measured in 2 studies [61,82] using resting brain dynamics with electroencephalography, Brief Visuospatial Memory Test-Revised, Neurophysiological Assessment Battery, Revised Hopkins Verbal Learning Test, Trail Making Test, and other neuropsychological evaluations. Quality of life was measured in 3 studies [57,62,71]. Other sensations were measured in 5 studies [52,54,68,72,78]. The experience of VR technology, such as presence, simulator sickness, physical comfort, feasibility, safety, effort put in the game, and satisfaction, was measured in 18 of the included studies [45,46,50-52,55,57,59,63,66,67,70,74,78-81,84] using the VR Satisfaction Scale, VR Symptom Questionnaire, satisfaction with treatment (numeric rating scale), motion and nausea (numeric rating scale), test of playfulness, illusion and control questions, questionnaire of embodiment, and a game experience survey.

Meta-analysis of Outcome Measurements

In most (21/25, 84%) studies, more than one outcome measurement was used to evaluate the effect of VR in patients with chronic pain, resulting in 194 effect sizes from 25 studies. Figure 2 [45,49,52,55-60,62-65,69-73,75-79,82,84] presents the summary forest plots, where each line represents all the outcomes from a particular study. If 0 was included in the CI, the reported effect size was not statistically significant at the 5% level. The overall effect size based on the 3‐level meta‐analysis was estimated at 1.22, with an SE of 0.34 and a 95% CI of 0.55-1.89 (z=3.56; P<.001). This indicates that VR intervention in patients with chronic pain had a positive effect on the outcome measurements used. Caterpillar plots are visualized in Figure 3 to provide a general view of the distribution of all effect sizes (Figure 3A) and study effect sizes (Figure 3B). Funnel plots are scatter plots in which effect sizes are plotted against the SE with small SEs at the top of the graph. Funnel plots for this meta-analysis are displayed in Figure 4, demonstrating an asymmetrical plot, which might be interpreted as an indication of publication bias.

The type of VR intervention (binary factor: immersive vs nonimmersive) was included as an additional covariate in the analysis; however, the type of VR intervention was not statistically significant ($F_{1,192}$=0.88; P=.35). Given these results, it can be concluded that the overall effect of VR on several outcome measurements is not moderated by the type of VR intervention. In addition, the type of pain (primary or secondary) and the objective of VR (exercise, virtual illusion, distraction, or cognitive therapy) were added as potential moderators to the analyses, whereby the omnibus test for the moderator analysis was not statistically significant for the type of pain ($F_{1,180}$=0.03; P=.86), nor for the objective of VR ($F_{1,190}$=0.59; P=.62).

When specifically focusing on effect sizes that measure the effect of VR on pain (n=31), an overall effect size of 1.60 (SE 0.39, 95% CI 0.83-2.36) was revealed, favoring VR interventions to decrease pain (z=4.09; P<.001; Figure 5A). For functioning (n=60), an effect size of 1.40 (SE 0.65, 95% CI 0.13-2.67) was calculated (z=2.17; P=.03; Figure 5B). For mobility (n=24; Figure 5C) and functional capacity (n=21; Figure 5D), effect sizes of 0.49 (SE 0.61, 95% CI −0.71 to 1.68; z=0.80, P=.42) and 0.34 (SE 0.95, 95% CI −1.52 to 2.20; z=0.36; P=.72) were revealed, respectively.

Figure 2. Summary forest plot for the effect of virtual reality on all outcome measurements in patients with chronic pain. Each line presents the results of 1 study (potentially including multiple effect sizes). The meta-analytic mean, with the corresponding 95% CI, is presented with a black dot and black line. This black 95% CI represents the total study precision. The additional CI in gray is based on the sampling variance of individual observed effect sizes of the study to obtain a visual contribution of the study sample size on the total study precision. The thickness of the gray CI is proportional to the number of effect sizes within studies. The number of effect sizes per study is presented as "J" on the right side of the figure [45,49,52,55-60,62-65,69-73,75-79,82,83].

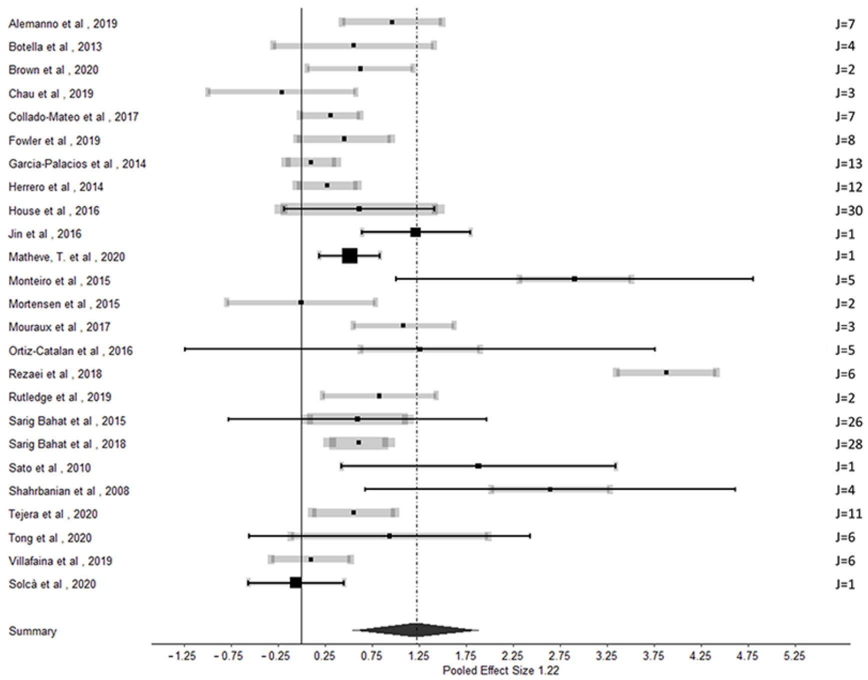


Figure 3. Caterpillar plots with all effect sizes (A) and study effect sizes (B). The overall effect size is presented in red.

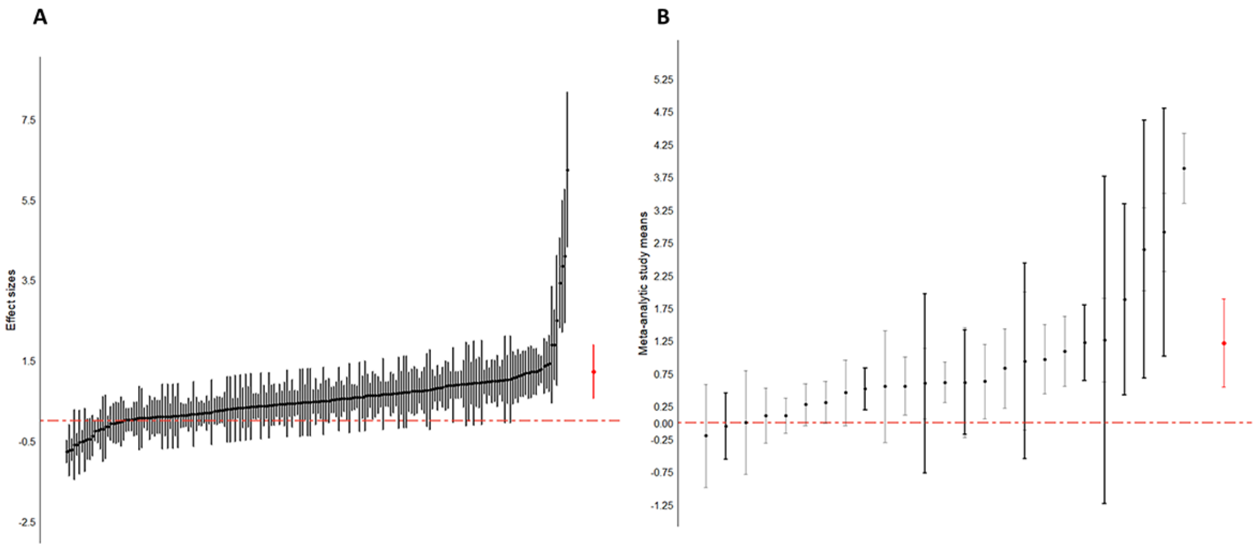


Figure 4. A funnel plot of all effect sizes (A) and a study funnel plot (B).

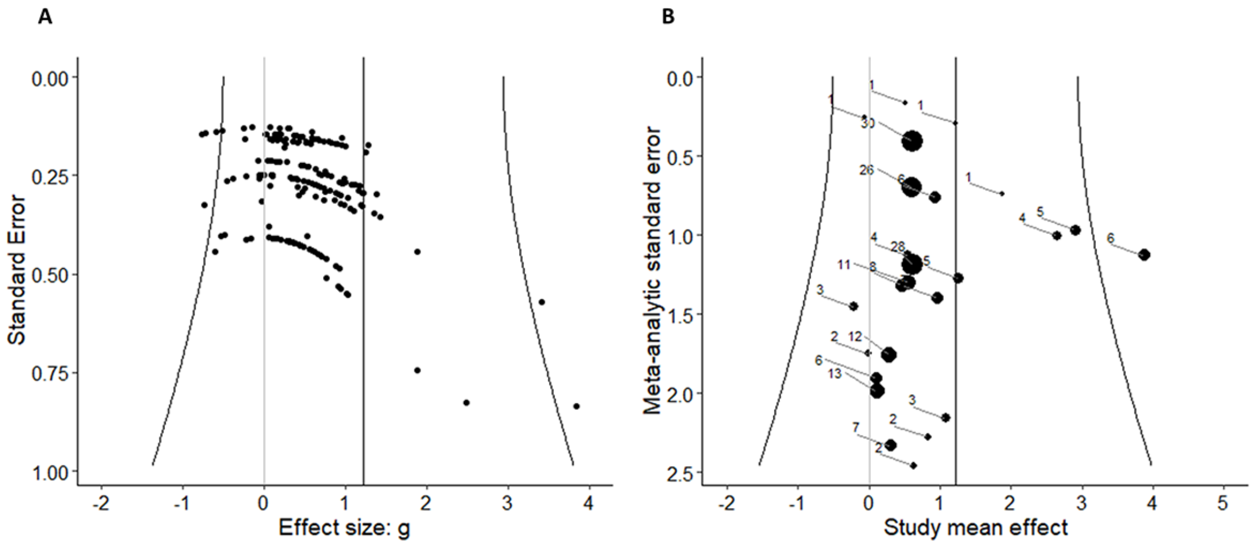
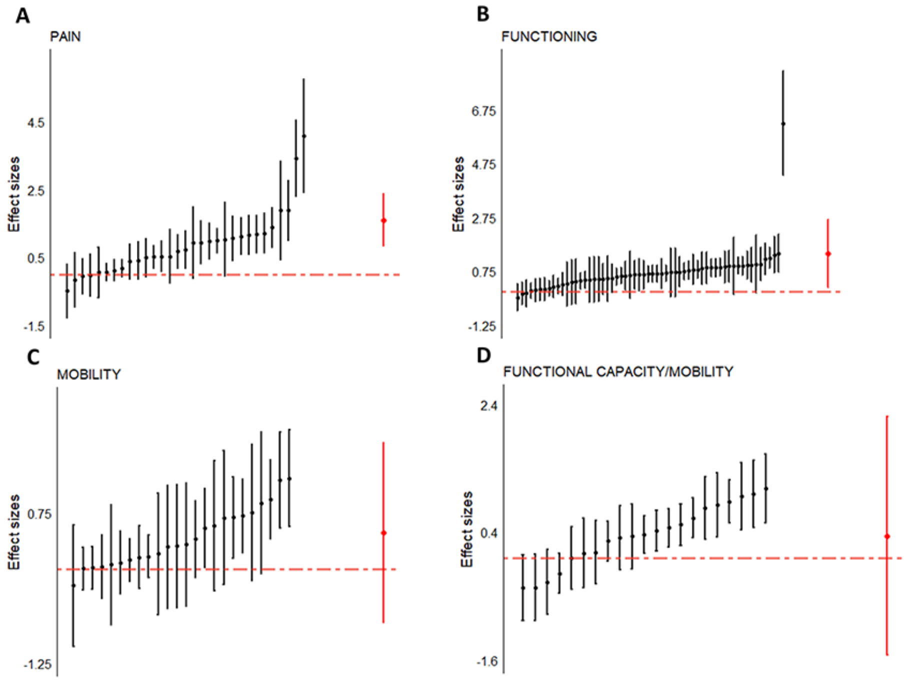


Figure 5. Reported effect sizes from the studies exploring (A) pain, (B) functioning, (C) mobility, and (D) functional capacity. Significant effects were found for pain and functioning outcome variables. Effect sizes from individual studies are presented in black, and the overall effect size is shown in red.



Discussion

Principal Findings

The main aim of this systematic review was to systematically identify all outcome parameters that are reported in relation to VR in the context of chronic pain management. This review revealed a broad range of outcome variables that are influenced by an intervention of VR technology, with statistically significant results on pain relief and improvements in functioning. These findings indicate that VR not only has applications in acute pain management but also in chronic pain. Over the last decade, medicine has shifted toward the use of innovative tools such as VR. The use of VR technology has expanded from the entertainment industry to clinical medicine and has proven to be cost-effective and efficient [14]. The effects of VR technology have been explored in different settings in medicine, ranging from surgical education to pain management [14]. The use of VR technology in surgical training has become an essential prerequisite for junior physicians before they can actively participate in real surgical operations [85]. Compared with VR technology, trainees who have received conventional surgical training have been found to perform surgery substantially slower and are more likely to cause injury, damage tissue, or fail to progress with surgery [86].

Comparison With Existing Literature

The management of pain in acute care settings often relies on pharmacological treatments to decrease pathophysiological responses [87]. However, the emergence of VR technology has brought another path for the management of acute pain and has been investigated for the management of burn-induced injuries [14]. As a result, VR technology provides analgesia with minimal side effects [88] and minimal impact on the physical hospital environment [88]. Another approach to the use of VR technology is to augment hypnosis, where patients report lower levels of pain and anxiety [89,90]. In addition to the use of morphine for pain reduction, VR for relaxation has shown to be very effective [91]. In addition, in chronic pain management, the applications of VR technology are very broad, including distraction, cognitive therapy, exercise, and enabling virtual limb movements [32].

Within the field of VR, one of the major classifications is based on the applied materials to induce a virtual experience. Both studies with 3D settings with a computer-generated simulation of 3D environment that makes use of a head-mounted display to entirely replace the real world and computer-generated simulations of 2D environment viewed on a computer or a wall-mounted screen were included [32]. A systematic review concluded that immersive VR technology is more likely to generate pain than nonimmersive (2D) VR technology [30]. As the price of immersive VR equipment continues to decrease, this approach is becoming very affordable and could potentially become a self-management tool for pain relief used by patients in ambulatory and home settings [14]. In this meta-analysis, the type of VR (immersive vs nonimmersive) was not a statistically significant confounder, indicating that VR in general has a beneficial effect on the reported outcome measurements. Further studies are required to confirm these results in chronic pain settings.

Patients affected by chronic pain aim to increase participation in daily living, ranging from work duties and family commitments to social activities and leisure time [92]. Therefore, a complete view of the patient’s functioning and health-related quality of life is needed rather than exclusively assessing the degree of pain relief [93]. A possible framework for organizing and documenting information on functioning and disability is the International Classification of Functioning, Disability, and Health (ICF), which is a classification of health and health-related outcomes [94]. The model conceptualizes functioning as a dynamic interaction between a person's health condition, personal factors, and environmental factors, as the functioning and disability of an individual occur in this context [95]. Therefore, the ICF provides a multi-perspective, biopsychosocial approach, which is reflected in the multidimensional model [95]. Multiple factors of the ICF model have been found in the reported outcome measurements of this systematic review, whereby mainly bodily functions and structures (eg, pain intensity and sleep), functional disability (eg, walking and balance control), and psychological distress (eg, depression, anxiety, and catastrophizing) were represented. The meta-analyses revealed a statistically significant positive effect of VR on pain-related (95% CI 0.83-2.36; z=4.09; P<.001) and functioning-related (95% CI 0.13-2.67; z=2.17; P=.03) outcome measurements. These results are in line with the decrease in pain intensity and disability by VR compared with proprioceptive training and lumbar stabilization exercises in patients with chronic neck pain and chronic low back pain, respectively, as observed in a recent meta-analysis [96]. No statistically significant results were found for functional capacity and mobility, whereby a number of effect sizes were indicative of no effect of VR on these outcome measurements. However, this finding does not exclude the potential added value of VR with another rehabilitation strategy to work on these concepts. Recently, it was suggested that manual therapy in combination with VR could define an entire mind-body intervention that relies on psychological, interoceptive, and exteroceptive stimulations for rebalancing sensorimotor integration and distorted perceptions, including visual and body images [97], denoting the complementary action of several conservative treatments.

Strengths and Limitations

The main strength of this literature review was the addition of a 3-level meta-analysis to consider the dependency between effect sizes. In the case of dependency, effect sizes are correlated, which inevitably leads to inflation of information and overconfidence in the results of a *standard* meta-analysis [42]. To avoid this inflation, several approaches have been used, among which the selection of only one effect size per study, relying on an average effect size for dependent effect sizes, or dependency could simply be ignored, and the analysis was performed as if the effect sizes were independent [98]. As the aim of this review was to evaluate the effect of VR on all outcome measurements in relation to chronic pain, the authors did not find the stated methods well suited, wherefore a multilevel approach was found to be the most appropriate. Furthermore, this systematic review has certain strengths that secure a minimal risk of potential bias: double-blind screening of the literature, including consensus meetings when needed, and consultation with several databases with the aim of fully representing the existing literature.

Certain limitations should be taken into account as well when interpreting the results of this study. None of the included studies qualified for excellent methodological quality. Of the 41 studies, most of the studies were scored as good on the risk of bias assessment (n=19, 46%), 17 (41%) scored as fair, and only 5 (12%) studies scored as poor. In many of the included studies, the authors mainly focused on the primary outcome variables; in this systematic review, all outcome variables were taken into account (primary and secondary outcome measurements); hence, the results were not always described in full detail for secondary outcomes. Finally, much heterogeneity is present in studies with VR, especially in terms of study design, underlying chronic pain etiologies, VR materials, and VR applications.

Conclusions

This systematic review explored the outcome measurements that are influenced by VR in patients with chronic pain. A broad range of outcome variables was revealed, whereby an intervention using VR technology can induce pain relief and improvements in functioning. These findings indicate that VR not only has applications in acute pain management but also in chronic pain settings, whereby VR technology might be able to become a promising first-line intervention as a complementary therapy for patients with chronic pain.

Conflicts of Interest

MM received speaker fees from Medtronic Nevro, and Saluda outside of the submitted work. LG is a postdoctoral research fellow funded by the Research Foundation Flanders, Belgium (project number: 12ZF622N). STIMULUS (Research and Teaching Neuromodulation Universitair Ziekenhuis Brussel/Vrije Universiteit Brussel) has received research grants from Medtronic.

Multimedia Appendix 1
The full search strategy for PubMed.
[DOCX File , 14 KB - games_v10i2e34402_app1.docx]

Multimedia Appendix 2
Risk of bias table based on the modified version of the Downs and Black Checklist.
[DOCX File , 87 KB - games_v10i2e34402_app2.docx]

Multimedia Appendix 3
Overview of the effects of virtual reality on several outcome measurements for each study.

[DOCX File , 152 KB - games_v10i2e34402_app3.docx]

References

1. Nicholas M, Vlaeyen JW, Rief W, Barke A, Aziz Q, Benoliel R, IASP Taskforce for the Classification of Chronic Pain. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019 Jan;160(1):28-37. [doi: 10.1097/j.pain.0000000000001390] [Medline: 30586068]
2. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006 May;10(4):287-333. [doi: 10.1016/j.ejpain.2005.06.009] [Medline: 16095934]
3. Suso-Ribera C, Yakobov E, Carriere JS, García-Palacios A. The impact of chronic pain on patients and spouses: consequences on occupational status, distribution of household chores and care-giving burden. Eur J Pain 2020 Oct;24(9):1730-1740. [doi: 10.1002/ejp.1616] [Medline: 32533892]
4. Naiditch N, Billot M, Moens M, Goudman L, Cornet P, Le Breton D, et al. Persistent spinal pain syndrome type 2 (PSPS-T2), a social pain? Advocacy for a social gradient of health approach to chronic pain. J Clin Med 2021 Jun 25;10(13):2817 [FREE Full text] [doi: 10.3390/jcm10132817] [Medline: 34202362]
5. Wippert PM, Drießlein D, Beck H, Schneider C, Puschmann AK, Banzer W, et al. The feasibility and effectiveness of a new practical multidisciplinary treatment for low-back pain: a randomized controlled trial. J Clin Med 2019 Dec 31;9(1):115 [FREE Full text] [doi: 10.3390/jcm9010115] [Medline: 31906224]
6. Gatchel RJ, Okifuji A. Evidence-based scientific data documenting the treatment and cost-effectiveness of comprehensive pain programs for chronic nonmalignant pain. J Pain 2006 Nov;7(11):779-793. [doi: 10.1016/j.jpain.2006.08.005] [Medline: 17074616]
7. Elbers S, Wittink H, Pool JJ, Smeets RJ. The effectiveness of generic self-management interventions for patients with chronic musculoskeletal pain on physical function, self-efficacy, pain intensity and physical activity: a systematic review and meta-analysis. Eur J Pain 2018 Oct;22(9):1577-1596 [FREE Full text] [doi: 10.1002/ejp.1253] [Medline: 29845678]
8. Damush TM, Kroenke K, Bair MJ, Wu J, Tu W, Krebs EE, et al. Pain self-management training increases self-efficacy, self-management behaviours and pain and depression outcomes. Eur J Pain 2016 Aug;20(7):1070-1078. [doi: 10.1002/ejp.830] [Medline: 26849410]
9. Billot M, Daycard M, Rigoard P. Self-Reiki, consideration of a potential option for managing chronic pain during pandemic COVID-19 period. Medicina (Kaunas) 2021 Aug 25;57(9):867 [FREE Full text] [doi: 10.3390/medicina57090867] [Medline: 34577790]
10. Du S, Hu L, Dong J, Xu G, Chen X, Jin S, et al. Self-management program for chronic low back pain: a systematic review and meta-analysis. Patient Educ Couns 2017 Jan;100(1):37-49. [doi: 10.1016/j.pec.2016.07.029] [Medline: 27554077]
11. Teater D. The psychological and physical side effects of pain medication. National Safety Council. 2015. URL: https://www.nsc.org/getmedia/0113f259-d2c5-4a3e-abca-f05299f65ec2/adv-rx-side-effects-wp.pdf [accessed 2021-10-10]
12. Cheatle MD. Prescription opioid misuse, abuse, morbidity, and mortality: balancing effective pain management and safety. Pain Med 2015 Oct;16 Suppl 1:S3-S8. [doi: 10.1111/pme.12904] [Medline: 26360931]
13. 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]
14. Li L, Yu F, Shi D, Shi J, Tian Z, Yang J, et al. Application of virtual reality technology in clinical medicine. Am J Transl Res 2017 Sep 15;9(9):3867-3880 [FREE Full text] [Medline: 28979666]
15. Chuan A, Zhou JJ, Hou RM, Stevens CJ, Bogdanovych A. Virtual reality for acute and chronic pain management in adult patients: a narrative review. Anaesthesia 2021 May;76(5):695-704. [doi: 10.1111/anae.15202] [Medline: 32720308]
16. Keefe FJ, Huling DA, Coggins MJ, Keefe DF, Rosenthal ZM, Herr NR, et al. Virtual reality for persistent pain: a new direction for behavioral pain management. Pain 2012 Nov;153(11):2163-2166 [FREE Full text] [doi: 10.1016/j.pain.2012.05.030] [Medline: 22770840]
17. Trost Z, Zielke M, Guck A, Nowlin L, Zakhidov D, France CR, et al. The promise and challenge of virtual gaming technologies for chronic pain: the case of graded exposure for low back pain. Pain Manag 2015;5(3):197-206. [doi: 10.2217/pmt.15.6] [Medline: 25971643]
18. Kober SE, Kurzmann J, Neuper C. Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: an EEG study. Int J Psychophysiol 2012 Mar;83(3):365-374. [doi: 10.1016/j.ijpsycho.2011.12.003] [Medline: 22206906]
19. 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 Nov 14;10:2509 [FREE Full text] [doi: 10.3389/fpsyg.2019.02509] [Medline: 31798492]
20. Repetto C, Germagnoli S, Triberti S, Riva G. Learning into the wild: a protocol for the use of 360° video for foreign language learning. In: Cipresso P, Serino S, Ostrovsky Y, Baker JT, editors. Pervasive Computing Paradigms for Mental Health. Cham, Switzerland: Springer; 2018:56-63.
21. Abiodun AO, Adesina MA. Virtual reality: a breakthrough in pain management? World News Nat Sci 2019;23:24-33.
22. Wismeijer AA, Vingerhoets AJ. The use of virtual reality and audiovisual eyeglass systems as adjunct analgesic techniques: a review of the literature. Ann Behav Med 2005 Dec;30(3):268-278. [doi: 10.1207/s15324796abm3003_11] [Medline: 16336078]

23. Hoffman HG, Patterson DR, Carrougher GJ, Nakamura D, Moore M, Garcia-Palacios A, et al. The effectiveness of virtual reality pain control with multiple treatments of longer durations: a case study. Int J Hum Comput Interact 2001 Mar 21;13(1):1-12. [doi: 10.1207/S15327590ijhc1301_1]
24. Rutter CE, Dahlquist LM, Weiss KE. Sustained efficacy of virtual reality distraction. J Pain 2009 Apr;10(4):391-397 [FREE Full text] [doi: 10.1016/j.jpain.2008.09.016] [Medline: 19231295]
25. Johnson MH. How does distraction work in the management of pain? Curr Pain Headache Rep 2005 Apr;9(2):90-95. [doi: 10.1007/s11916-005-0044-1] [Medline: 15745617]
26. Mahrer NE, Gold JI. The use of virtual reality for pain control: a review. Curr Pain Headache Rep 2009 Apr;13(2):100-109. [doi: 10.1007/s11916-009-0019-8] [Medline: 19272275]
27. Cheung KL, Tunik E, Adamovich SV, Boyd LA. Neuroplasticity and virtual reality. In: Weiss PL, Keshner EA, Levin MF, editors. Virtual Reality for Physical and Motor Rehabilitation. New York, NY, USA: Springer; 2014:5-24.
28. Garrett B, Taverner T, Masinde W, Gromala D, Shaw C, Negraeff M. A rapid evidence assessment of immersive virtual reality as an adjunct therapy in acute pain management in clinical practice. Clin J Pain 2014 Dec;30(12):1089-1098. [doi: 10.1097/AJP.0000000000000064] [Medline: 24535053]
29. Shahrbanian S, Ma X, Korner-Bitensky N, Simmonds MJ. Scientific evidence for the effectiveness of virtual reality for pain reduction in adults with acute or chronic pain. Stud Health Technol Inform 2009;144:40-43. [Medline: 19592727]
30. Malloy KM, Milling LS. The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clin Psychol Rev 2010 Dec;30(8):1011-1018. [doi: 10.1016/j.cpr.2010.07.001] [Medline: 20691523]
31. Chan E, Foster S, Sambell R, Leong P. Clinical efficacy of virtual reality for acute procedural pain management: a systematic review and meta-analysis. PLoS One 2018 Jul 27;13(7):e0200987 [FREE Full text] [doi: 10.1371/journal.pone.0200987] [Medline: 30052655]
32. Austin PD. The analgesic effects of virtual reality for people with chronic pain: a scoping review. Pain Med 2022 Jan 03;23(1):105-121. [doi: 10.1093/pm/pnab217] [Medline: 34260724]
33. 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. Int J Surg 2021 Apr;88:105906. [doi: 10.1016/j.ijsu.2021.105906] [Medline: 33789826]
34. O'Sullivan D, Wilk S, Michalowski W, Farion K. Using PICO to align medical evidence with MDs decision making models. Stud Health Technol Inform 2013;192:1057. [Medline: 23920831]
35. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. Pain 2015 Jun;156(6):1003-1007 [FREE Full text] [doi: 10.1097/j.pain.0000000000000160] [Medline: 25844555]
36. 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]
37. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun;52(6):377-384 [FREE Full text] [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
38. Trac MH, McArthur E, Jandoc R, Dixon SN, Nash DM, Hackam DG, et al. Macrolide antibiotics and the risk of ventricular arrhythmia in older adults. CMAJ 2016 Apr 19;188(7):E120-E129 [FREE Full text] [doi: 10.1503/cmaj.150901] [Medline: 26903359]
39. Korakakis V, Whiteley R, Tzavara A, Malliaropoulos N. The effectiveness of extracorporeal shockwave therapy in common lower limb conditions: a systematic review including quantification of patient-rated pain reduction. Br J Sports Med 2018 Mar;52(6):387-407. [doi: 10.1136/bjsports-2016-097347] [Medline: 28954794]
40. Hooper P, Jutai JW, Strong G, Russell-Minda E. Age-related macular degeneration and low-vision rehabilitation: a systematic review. Can J Ophthalmol 2008 Apr;43(2):180-187. [doi: 10.3129/i08-001] [Medline: 18347620]
41. Becker BJ. Synthesizing standardized mean-change measures. Br J Math Stat Psychol 1988 Nov;41(2):257-278. [doi: 10.1111/j.2044-8317.1988.tb00901.x]
42. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Three-level meta-analysis of dependent effect sizes. Behav Res Methods 2013 Jun;45(2):576-594. [doi: 10.3758/s13428-012-0261-6] [Medline: 23055166]
43. Fernández-Castilla B, Declercq L, Jamshidi L, Beretvas SN, Onghena P, Van den Noortgate W. Visual representations of meta-analyses of multiple outcomes: extensions to forest plots, funnel plots, and caterpillar plots. Methodology 2020 Dec 22;16(4):299-315. [doi: 10.5964/meth.4013]
44. Amin AM, Tong X, Gromala D, Shaw CD. Cardboard mobile virtual reality as an approach for pain distraction in clinical settings: comparison, exploration and evaluation with oculus rift. 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, USA p. 2345-2351. [doi: 10.1145/3027063.3053234]
45. Fowler CA, Ballistrea LM, Mazzone KE, Martin AM, Kaplan H, Kip KE, et al. Virtual reality as a therapy adjunct for fear of movement in veterans with chronic pain: single-arm feasibility study. JMIR Form Res 2019 Oct 30;3(4):e11266 [FREE Full text] [doi: 10.2196/11266] [Medline: 31670696]
46. Garrett B, Taverner T, McDade P. Virtual reality as an adjunct home therapy in chronic pain management: an exploratory study. JMIR Med Inform 2017 May 11;5(2):e11 [FREE Full text] [doi: 10.2196/medinform.7271] [Medline: 28495661]

47. Gromala D, Tong X, Choo A, Karamnejad M, Shaw CD. The virtual meditative walk: virtual reality therapy for chronic pain management. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15; April 18-23, 2015; Seoul, Republic of Korea p. 521-524. [doi: 10.1145/2702123.2702344]
48. Guarino D, La Paglia F, Daino M, Maiorca V, Zichichi S, Guccione F, et al. Chronic pain treatment through virtual reality. Annu Rev CyberTherapy Telemed 2017 Dec;15:181-184.
49. 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;220:154-160. [Medline: 27046570]
50. Jones T, Moore T, Choo J. The impact of virtual reality on chronic pain. PLoS One 2016 Dec 20;11(12):e0167523 [FREE Full text] [doi: 10.1371/journal.pone.0167523] [Medline: 27997539]
51. Pamment J, Aspell JE. Putting pain out of mind with an 'out of body' illusion. Eur J Pain 2017 Feb;21(2):334-342. [doi: 10.1002/ejp.927] [Medline: 27509229]
52. Shahrbanian S, Simmonds MJ. Effects of different virtual reality environments on experimental pain rating in post-stroke individuals with and without pain in comparison to pain free healthy individuals. Annu Rev CyberTher Telemed 2008 Jul:653-659.
53. Wiederhold BK, Gao K, Kong L, Wiederhold MD. Mobile devices as adjunctive pain management tools. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):385-389 [FREE Full text] [doi: 10.1089/cyber.2014.0202] [Medline: 24892202]
54. Wiederhold BK, Gao K, Sulea C, Wiederhold MD. Virtual reality as a distraction technique in chronic pain patients. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):346-352 [FREE Full text] [doi: 10.1089/cyber.2014.0207] [Medline: 24892196]
55. Botella C, Garcia-Palacios A, Vizcaíno Y, Herrero R, Baños RM, Belmonte MA. Virtual reality in the treatment of fibromyalgia: a pilot study. Cyberpsychol Behav Soc Netw 2013 Mar;16(3):215-223. [doi: 10.1089/cyber.2012.1572] [Medline: 23496678]
56. Collado-Mateo D, Dominguez-Muñoz FJ, Adsuar JC, Merellano-Navarro E, Gusi N. Exergames for women with fibromyalgia: a randomised controlled trial to evaluate the effects on mobility skills, balance and fear of falling. PeerJ 2017 Apr 20;5:e3211 [FREE Full text] [doi: 10.7717/peerj.3211] [Medline: 28439471]
57. Garcia-Palacios A, Herrero R, Vizcaíno Y, Belmonte MA, Castilla D, Molinari G, et al. Integrating virtual reality with activity management for the treatment of fibromyalgia: acceptability and preliminary efficacy. Clin J Pain 2015 Jun;31(6):564-572. [doi: 10.1097/AJP.0000000000000196] [Medline: 25551475]
58. Herrero R, García-Palacios A, Castilla D, Molinari G, Botella C. Virtual reality for the induction of positive emotions in the treatment of fibromyalgia: a pilot study over acceptability, satisfaction, and the effect of virtual reality on mood. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):379-384. [doi: 10.1089/cyber.2014.0052] [Medline: 24892201]
59. Mortensen J, Kristensen LQ, Brooks EP, Brooks AL. Women with fibromyalgia's experience with three motion-controlled video game consoles and indicators of symptom severity and performance of activities of daily living. Disabil Rehabil Assist Technol 2015 Jan;10(1):61-66. [doi: 10.3109/17483107.2013.836687] [Medline: 24028282]
60. Villafaina S, Borrega-Mouquinho Y, Fuentes-García JP, Collado-Mateo D, Gusi N. Effect of exergame training and detraining on lower-body strength, agility, and cardiorespiratory fitness in women with fibromyalgia: single-blinded randomized controlled trial. Int J Environ Res Public Health 2019 Dec 24;17(1):161 [FREE Full text] [doi: 10.3390/ijerph17010161] [Medline: 31878319]
61. Villafaina S, Collado-Mateo D, Fuentes JP, Rohlfs-Domínguez P, Gusi N. Effects of exergames on brain dynamics in women with fibromyalgia: a randomized controlled trial. J Clin Med 2019 Jul 11;8(7):1015 [FREE Full text] [doi: 10.3390/jcm8071015] [Medline: 31336706]
62. Alemanno F, Houdayer E, Emedoli D, Locatelli M, Mortini P, Mandelli C, et al. Efficacy of virtual reality to reduce chronic low back pain: proof-of-concept of a non-pharmacological approach on pain, quality of life, neuropsychological and functional outcome. PLoS One 2019 May 23;14(5):e0216858 [FREE Full text] [doi: 10.1371/journal.pone.0216858] [Medline: 31120892]
63. Brown L, Chen ET, Binder DS. The use of virtual reality for Peri-procedural pain and anxiety at an outpatient spine clinic injection visit: an exploratory controlled randomized trial. Am J Transl Res 2020 Sep 15;12(9):5818-5826 [FREE Full text] [Medline: 33042461]
64. 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]
65. Monteiro-Junior RS, de Souza CP, Lattari E, Rocha NB, 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]
66. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. J Pain 2016 Dec;17(12):1302-1317 [FREE Full text] [doi: 10.1016/j.jpain.2016.08.011] [Medline: 27616607]
67. Trujillo MS, Alvarez AF, Nguyen L, Petros J. Embodiment in virtual reality for the treatment of chronic low back pain: a case series. J Pain Res 2020 Nov 25;13:3131-3137 [FREE Full text] [doi: 10.2147/JPR.S275312] [Medline: 33269003]

68. 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]
69. Rezaei I, Razeghi M, Ebrahimi S, Kayedi S, Rezaeian Zadeh A. A novel virtual reality technique (Cervigame®) compared to conventional proprioceptive training to treat neck pain: a randomized controlled trial. J Biomed Phys Eng 2018 Nov 18;9(3):355-366. [doi: 10.31661/jbpe.v0i0.556] [Medline: 31341881]
70. 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]
71. 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]
72. 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]
73. Chau B, Phelan I, Ta P, Chi B, Loyola K, Yeo E, et al. Immersive virtual reality for pain relief in upper limb complex regional pain syndrome: a pilot study. Innov Clin Neurosci 2020 Apr 01;17(4-6):47-52 [FREE Full text] [Medline: 32802594]
74. Matamala-Gomez M, Diaz Gonzalez AM, Slater M, Sanchez-Vives MV. Decreasing pain ratings in chronic arm pain through changing a virtual body: different strategies for different pain types. J Pain 2019 Jun;20(6):685-697. [doi: 10.1016/j.jpain.2018.12.001] [Medline: 30562584]
75. Sato K, Fukumori S, Matsusaki T, Maruo T, Ishikawa S, Nishie H, et al. Nonimmersive virtual reality mirror visual feedback therapy and its application for the treatment of complex regional pain syndrome: an open-label pilot study. Pain Med 2010 Apr;11(4):622-629. [doi: 10.1111/j.1526-4637.2010.00819.x] [Medline: 20202141]
76. Mouraux D, Brassinne E, Sobczak S, Nonclercq A, Warzée N, Sizer PS, et al. 3D augmented reality mirror visual feedback therapy applied to the treatment of persistent, unilateral upper extremity neuropathic pain: a preliminary study. J Man Manip Ther 2017 Jul;25(3):137-143 [FREE Full text] [doi: 10.1080/10669817.2016.1176726] [Medline: 28694676]
77. 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]
78. Rutledge T, Velez D, Depp C, McQuaid JR, Wong G, Jones RC, et al. A virtual reality intervention for the treatment of phantom limb pain: development and feasibility results. Pain Med 2019 Oct 01;20(10):2051-2059. [doi: 10.1093/pm/pnz121] [Medline: 31165893]
79. Tong X, Wang X, Cai Y, Gromala D, Williamson O, Fan B, et al. "I dreamed of my hands and arms moving again": a case series investigating the effect of immersive virtual reality on phantom limb pain alleviation. Front Neurol 2020;11:876 [FREE Full text] [doi: 10.3389/fneur.2020.00876] [Medline: 32982914]
80. Griffin A, Wilson L, Feinstein AB, Bortz A, Heirich MS, Gilkerson R, et al. Virtual reality in pain rehabilitation for youth with chronic pain: pilot feasibility study. JMIR Rehabil Assist Technol 2020 Nov 23;7(2):e22620 [FREE Full text] [doi: 10.2196/22620] [Medline: 33226346]
81. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-administered skills-based virtual reality intervention for chronic pain: randomized controlled pilot study. JMIR Form Res 2020 Jul 07;4(7):e17293 [FREE Full text] [doi: 10.2196/17293] [Medline: 32374272]
82. House G, Burdea G, Grampurohit N, Polistico K, Roll D, Damiani F, et al. A feasibility study to determine the benefits of upper extremity virtual rehabilitation therapy for coping with chronic pain post-cancer surgery. Br J Pain 2016 Nov;10(4):186-197 [FREE Full text] [doi: 10.1177/2049463716664370] [Medline: 27867508]
83. Phoon Nguyen AH, Balasubramaniam R, Bellan V, Newport RN, Stanton TR. The effect of multisensory illusions on pain and perceived burning sensations in patients with Burning Mouth Syndrome: a proof-of-concept study. J Oral Pathol Med 2020 Jul;49(6):505-513. [doi: 10.1111/jop.13065] [Medline: 32531809]
84. Solcà M, Krishna V, Young N, Deogaonkar M, Herbelin B, Orepic P, et al. Enhancing analgesic spinal cord stimulation for chronic pain with personalized immersive virtual reality. Pain 2021 Jun 01;162(6):1641-1649. [doi: 10.1097/j.pain.0000000000002160] [Medline: 33259460]
85. Stirling ER, Lewis TL, Ferran NA. Surgical skills simulation in trauma and orthopaedic training. J Orthop Surg Res 2014 Dec 19;9:126 [FREE Full text] [doi: 10.1186/s13018-014-0126-z] [Medline: 25523023]
86. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002 Oct;236(4):458-464. [doi: 10.1097/00000658-200210000-00008] [Medline: 12368674]
87. Kehlet H, Holte K. Effect of postoperative analgesia on surgical outcome. Br J Anaesth 2001 Jul;87(1):62-72 [FREE Full text] [doi: 10.1093/bja/87.1.62] [Medline: 11460814]

88. Das DA, Grimmer KA, Sparnon AL, McRae SE, Thomas BH. The efficacy of playing a virtual reality game in modulating pain for children with acute burn injuries: a randomized controlled trial [ISRCTN87413556]. BMC Pediatr 2005 Mar 03;5(1):1 [FREE Full text] [doi: 10.1186/1471-2431-5-1] [Medline: 15745448]
89. Patterson DR, Hoffman HG, Palacios AG, Jensen MJ. Analgesic effects of posthypnotic suggestions and virtual reality distraction on thermal pain. J Abnorm Psychol 2006 Nov;115(4):834-841. [doi: 10.1037/0021-843X.115.4.834] [Medline: 17100541]
90. Rousseaux F, Bicego A, Ledoux D, Massion P, Nyssen AS, Faymonville ME, et al. Hypnosis associated with 3D immersive virtual reality technology in the management of pain: a review of the literature. J Pain Res 2020 May 21;13:1129-1138 [FREE Full text] [doi: 10.2147/JPR.S231737] [Medline: 32547176]
91. Konstantatos AH, Angliss M, Costello V, Cleland H, Stafrace S. Predicting the effectiveness of virtual reality relaxation on pain and anxiety when added to PCA morphine in patients having burns dressings changes. Burns 2009 Jun;35(4):491-499. [doi: 10.1016/j.burns.2008.08.017] [Medline: 19111995]
92. Goudman L, De Smedt A, Linderoth B, Eldabe S, Witkam R, Henssen D, et al. Identifying goals in patients with chronic pain: a European survey. Eur J Pain 2021 Oct;25(9):1959-1970. [doi: 10.1002/ejp.1814] [Medline: 34048121]
93. Fehrmann E, Kotulla S, Fischer L, Kienbacher T, Tuechler K, Mair P, et al. The impact of age and gender on the ICF-based assessment of chronic low back pain. Disabil Rehabil 2019 May;41(10):1190-1199. [doi: 10.1080/09638288.2018.1424950] [Medline: 29327593]
94. The International Classification of Functioning, Disability and Health (ICF). World Health Organization. 2001. URL: http://www.who.int/classifications/icf/en/ [accessed 2021-09-29]
95. How to use the ICF: a practical manual for using the International Classification of Functioning, Disability and Health (ICF) - Exposure draft for comment. World Health Organization. 2013. URL: https://www.medbox.org/pdf/5e148832db60a2044c2d4d88 [accessed 2021-08-12]
96. 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]
97. Cerritelli F, Chiera M, Abbro M, Megale V, Esteves J, Gallace A, et al. The challenges and perspectives of the integration between virtual and augmented reality and manual therapies. Front Neurol 2021 Jun 30;12:700211 [FREE Full text] [doi: 10.3389/fneur.2021.700211] [Medline: 34276550]
98. Assink M, Wibbelink CJ. Fitting three-level meta-analytic models in R: a step-by-step tutorial. Quant Meth Psych 2016 Oct 01;12(3):154-174. [doi: 10.20982/tqmp.12.3.p154]

Abbreviations

ICF: International Classification of Functioning, Disability, and Health
PICO: Patient or Population, Intervention, Comparison, and Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VR: virtual reality

Edited by N Zary; submitted 21.10.21; peer-reviewed by J Gold, R Nieto; comments to author 30.12.21; revised version received 18.02.22; accepted 16.03.22; published 10.05.22.

Please cite as:
Goudman L, Jansen J, Billot M, Vets N, De Smedt A, Roulaud M, Rigoard P, Moens M
Virtual Reality Applications in Chronic Pain Management: Systematic Review and Meta-analysis
JMIR Serious Games 2022;10(2):e34402
URL: https://games.jmir.org/2022/2/e34402
doi:10.2196/34402
PMID:35536641

Review

Using Video Games to Improve the Sexual Health of Young People Aged 15 to 25 Years: Rapid Review

Ignacio Franco Vega[1], MSc; Anastasia Eleftheriou[2], PhD; Cynthia Graham[3], ClinPsy, PhD

[1]Department of Social & Policy Sciences, University of Bath, Bath, United Kingdom
[2]TRI Technologos Research and Innovation Services Ltd, Larnaka, Cyprus
[3]Department of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, United Kingdom

Corresponding Author:
Cynthia Graham, ClinPsy, PhD
Department of Psychology
Faculty of Environmental and Life Sciences
University of Southampton
University Road, Room 44/3073
Southampton, SO17 1BJ
United Kingdom
Phone: 44 023 8059 3091
Email: c.a.graham@soton.ac.uk

Abstract

Background: Sexually transmitted infections and unintended pregnancies among young people remain public health concerns in many countries. To date, interventions that address these concerns have had limited success. Serious games are increasingly being used as educational tools in health and professional public education. Although acknowledged as having great potential, few studies have evaluated the use of serious games in sexual health education among young people, and to date, there have been no published reviews of these studies.

Objective: This study aims to assess the effects of video game–based sexual health interventions for risky sexual behavior in young people aged between 15 and 25 years.

Methods: A rapid review of randomized controlled trials and quasi–randomized controlled trials was performed. The search included the following bibliographic databases: Cochrane Central Register of Controlled Trials, Embase, MEDLINE, PsycINFO, and Scopus. A total of 2 reviewers independently screened 50% (35/70) of the retrieved articles during the full-text screening phase.

Results: From a total of 459 identified citations, after removing duplicates, 327 (71.2%) articles were deemed eligible for title and abstract screening. Of the 327 articles, 70 (21.4%) full texts were screened, from which 10 (3.1%) articles (evaluating 11 different games) were included in the review. The findings highlighted the considerable diversity in video game–based interventions and assessed sexual health outcomes. Although there were some promising findings in outcome studies using game-based interventions, the results across studies were mixed.

Conclusions: Although game interventions for sexual health have been in existence for almost three decades, relatively few studies have evaluated them, and the results of previous outcome studies have been mixed. Moreover, there is little clarity regarding which specific elements of a game facilitate positive outcomes. We provide recommendations for future researchers developing video game–based interventions to improve sexual health in young people.

(JMIR Serious Games 2022;10(2):e33207) doi:10.2196/33207

KEYWORDS

sex education; serious games; sexually transmitted infections; rapid review; mobile phone

Introduction

Background

Although contraceptive access and sexual education are more widely available than ever, risky sexual behavior remains an issue for people of all ages, especially younger individuals. More than 1 million sexually transmitted infections (STIs) are acquired each day worldwide among individuals aged between 15 and 49 years [1,2]. The most common STIs are chlamydia, gonorrhea, and trichomoniasis, although the diagnosis of syphilis has also increased in recent years [1]. In many countries, young people aged between 15 and 24 years have the highest rates of STIs [3,4]. Despite global and national efforts to stop the spread of STIs, the World Health Organization recently reported a “concerning lack of progress” in achieving reductions [5].

The best method for preventing the spread of STIs is the correct use of condoms [6]. However, many young people [7] engage in risky sexual behaviors such as having sex with multiple partners without the use of condoms or incorrect or incomplete condom use [8]. Many interventions have been developed to encourage consistent condom use and safer sexual behaviors; however, most of these interventions have been shown to have limited effectiveness and/or are very resource intensive [9,10].

Adolescent pregnancies are also a global concern in high-, middle-, and low-income countries. Although the past 30 years have seen a global decline in unintended pregnancy rates, a recent review of 166 low- to middle-income countries reported that approximately half of all pregnancies are unintended [11]. Furthermore, although rates of unintended pregnancies in the United States and the United Kingdom have dropped in recent decades [12], adolescent pregnancy rates remain high in many middle- to high-income countries (particularly in the United States) [13].

Despite consistent evidence that comprehensive sex education can increase protective behaviors [14], there are still many gaps in knowledge. Furthermore, access to contraceptives and sexual health services for young people remains limited in many countries [4]. Sex education, often delivered in schools as part of the national curriculum, can be a highly contested area, reflecting political, moral, and cultural debates. In the United States, school-based sex education curricula have long been criticized for being sex negative, often focusing on abstinence and omitting any mention of nonheterosexual experiences [15,16]. In many countries, traditional gatekeepers such as religious and educational authorities still powerfully restrict access, content, and materials used for sex education [4]. Therefore, for many young people, obtaining reliable information about sex and relationships can be difficult.

The internet has been identified as a potentially valuable resource for comprehensive, interactive, web-based, and youth-friendly sex education [16]. Young people worldwide use the internet and social media to access information on sexual and reproductive health and rights [16,17]. Interventions delivered through digital media could particularly help reach marginalized groups such as young people in rural areas; lesbian, gay, bisexual, transgender, and intersex individuals; people with disabilities; and migrant populations [17].

It has been argued that education through games is more efficient and enjoyable than classroom teaching for several reasons [18]. First, it is predominantly the player who directs activity in games, whereas in school, it is predominantly the teacher who directs activity. This is why serious games use a learner-centered approach in which learners are involved in the process (learning through doing), in contrast to traditional education, which uses a teacher-centered approach in which learners are relatively passive.

Second, children and adolescents often find it difficult to properly engage in school exercises [19], in which the challenge level is not well adjusted to their skills. In a class, there are many students with different skills, making it difficult for teachers to equally engage all students in the class. In contrast, video games engage players naturally by gradually adjusting their difficulty level as they progress in the game [20]. Game developers understand that for a game to be successful, players of varying abilities need to feel a sense of reward or achievement, often enough to retain their engagement.

Third, students are sometimes discouraged by the school system as they are penalized for the mistakes they make (eg, they receive bad grades). However, in games, players are expected to make wrong decisions and do so without being discouraged (ideally, unless the game is poorly designed). In fact, games have the advantage of allowing users to train in real-life decision-making situations where the wrong choice may involve some risk without having to actually be at risk. For example, pilots often train using Microsoft Flight Simulator, whereas the military often uses battle simulators to train recruits. This allows players to make mistakes in a safe environment.

Finally, an important characteristic of educational games is the constant real-time feedback provided to the user. Players almost instantly know how well a certain move or strategy works toward the goal of the game. Feedback can take the form of points, lives, levels, scores, ranks, or progress bars. Real-time feedback ensures that users are motivated throughout the game by promising that a goal is achievable.

Some authors have argued that there is a strong case for integrating video games into sex education, whether by supplementing sex education classes with existing games that explore sex and sexuality or developing new games for the purpose of sex education [21]. Given the interactive nature of video games, their lack of real consequences, their capacity for privacy, and the familiarity that many adolescents already have with games, when used correctly, games could be very effective tools for students.

Aims of the Review

This review was conducted as part of a larger Erasmus+ funded project (Safe4Play) that aims to develop an innovative tool for sex and reproductive health education for young people using serious games with machine learning features. The aim of this review was to analyze the core elements and effects of video game–based interventions for improving the sexual health of young people. The findings informed the development of the intervention that was produced as part of the Safe4Play initiative.

Methods

We conducted a rapid review following the Cochrane Rapid Reviews Method Group guidelines [22,23]. A rapid review can be defined as a type of knowledge synthesis in which the usual procedures of a traditional systematic review are streamlined and accelerated such that the most crucial elements are still present, but the research time is considerably abridged [24].

Criteria for Study Selection

The criteria for selecting studies were based on the Population, Intervention, Comparison, Outcomes, and Study characteristics framework.

Population

This involved interventions aimed at working with youth (aged 15-25 years). Where studies included participants who fell both inside and outside of our target bracket (eg, aged 12-16 years), we tried, where possible, to select the appropriate results from the subset of the sample that met our age criteria; if that was not possible, we captured that specific limitation in the narrative form.

Intervention

This involved any video game–based sexual health intervention aimed at reducing risky sexual behavior. We considered a *video game–based intervention* as an educational intervention delivered through an electronic or digital medium that relied heavily on game mechanics, aesthetics, or game thinking (competition, cooperation, exploration, and storytelling) to engage, motivate action, promote learning, and solve problems [25].

Comparison

This criterion was not applicable.

Outcomes

As we were broadly interested in sexual health, we chose to include studies that assessed a broad range of knowledge, attitudinal, and behavioral variables. We defined primary outcomes as any of the following: decrease in unintended pregnancies and STIs, increase in contraceptive use, increase in intention to use contraceptives, acquisition of new knowledge regarding sexual health, change in the perception of risk of pregnancy, and change in the perception of risk of STIs. Secondary outcomes included changes in attitudes toward safe sex, self-efficacy toward sexual health, decrease in the number of sexual partners, increase in safe and consensual relationship practices, and increase in adherence to pre-exposure prophylaxis (PrEP).

Where studies reported >1 relevant outcome, each one was captured and reported in a narrative form. When outcomes were provided at multiple follow-up points, all outcomes were reported for each follow-up point.

Study Characteristics

We included randomized controlled trials (RCTs) and quasi-RCTs (studies in which participants were allocated to different arms of the study using a method of allocation that is not truly random). Publications in either English or Spanish were considered.

Search Strategy and Search Terms

The search strategy was validated by the Safe4Play research team and an information retrieval specialist from the University of Bath. It was piloted to analyze the quality and quantity of its results; only small changes were made based on the findings.

We used 5 databases to identify relevant studies: Cochrane Central Register of Controlled Trials, Embase, MEDLINE, PsycINFO, and Scopus. Searches were conducted on April 23, 2021. In addition, we hand-searched the reference lists of the included trials for referenced articles that were not retrieved in the original search. We also contacted experts in the field for additional recent publications that the original search might not have identified. For details of the search terms used for each of the databases, see Multimedia Appendix 1.

Study Selection

A total of 2 steps were undertaken to assess the eligibility of the studies: title and abstract screening and full-text screening. A total of 2 reviewers (IFV and CG) were involved in the process. Approximately 20% (51/257) of the abstracts were independently screened by both reviewers, which served as a pilot to identify any salient issues. The remaining 80% (206/257) of the abstracts were screened by IFV. Interrater reliability was found to be moderate (weighted κ=0.53) [26]. All cases of uncertainty or discrepancy were resolved through discussions between the 2 reviewers.

In the full-text screening stage, both reviewers independently screened half of the articles to confirm whether the studies identified during the title and abstract screening should be included. Reliability was found to be substantial (κ=0.71) [26]. The same procedure was used to resolve any discrepancies between reviewers. The remaining articles were screened solely by IFV.

Data extraction was performed by IFV. All pertinent data were extracted from the full text using a spreadsheet template. When an intervention was analyzed in multiple papers, data from all papers were considered during the extraction.

Results

Search Results

As shown in Figure 1, the search strategy produced 449 results, of which, after removing 132 (29.4%) duplicates, 317 (70.6%) articles remained (299/317, 94.3% of empirical papers, and 18/317, 5.7% of reviews). All systematic reviews were scanned to identify additional articles to screen; 10 additional articles were found through this process. A total of 327 abstracts were deemed appropriate for screening.

Overall, of the 327 articles found, 257 (78.6%) were screened at the title and abstract screening, leaving 70 (21.4%) articles for full-text screening. These 70 articles were downloaded and examined. After this final screening procedure of the 70 articles, 60 (86%) articles were excluded, leaving 10 (14%) articles with suitable games to analyze. Most articles described 1 game each, although one of the articles evaluated 2 games. Thus, the final search product was 11 games.

In some cases, to obtain the information required to conduct a proper analysis, additional supplementary materials had to be downloaded. Most of these were in the form of protocols for trials or articles that reported preliminary results. In the following sections, we briefly describe each of the identified games. Table 1 presents some of the key features (sample, location, and type of game) of each video game.

Figure 1. Flow diagram from the article selection process.

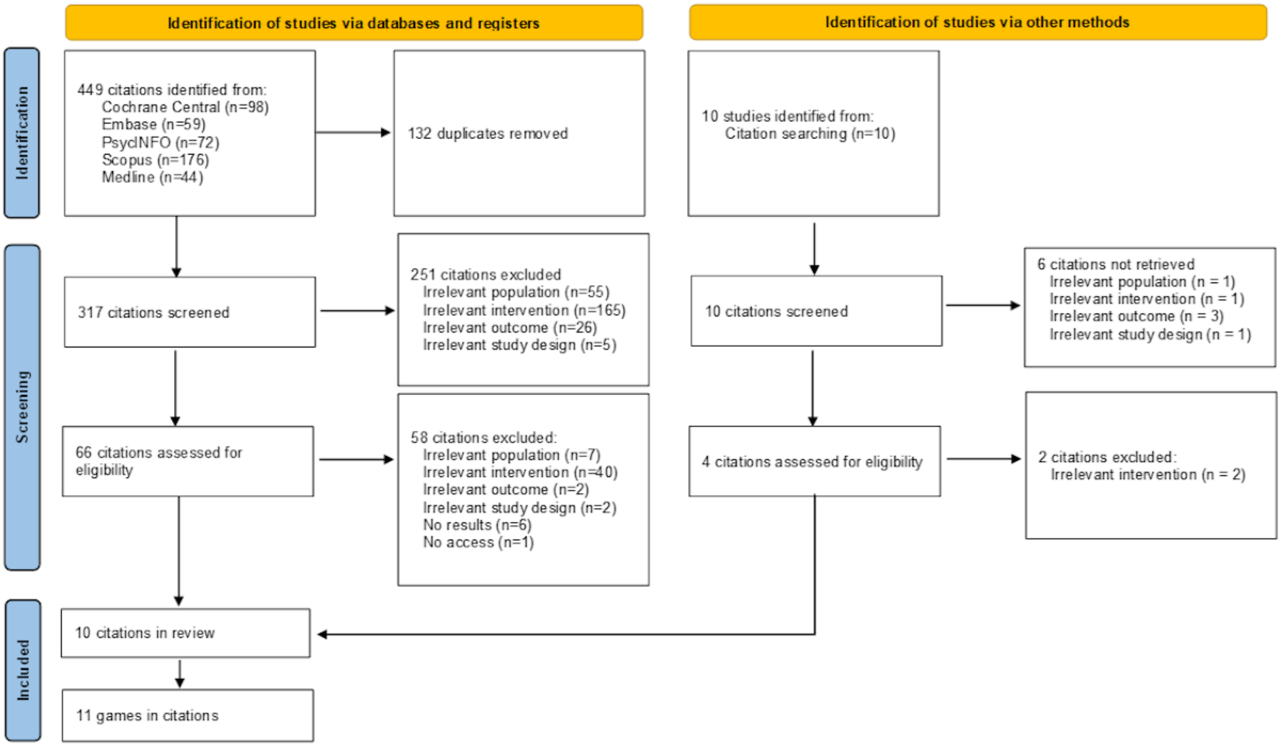


Table 1. Key features of included studies.

Game name	Target population	Age (years), range	Location and publication date	Console	Type of game
The Baby Game! [27]	High school students	13-18	Hawaii, United States, 1989	PC	Management simulator
Romance [27]	High school students	13-18	Hawaii, United States, 1989	PC	Text adventure
VODO [28]	High school students	15	New Orleans, United States, 1989	PC	Text adventure
Choose Your Own Adventure [29]	High school sophomores	15-16	Kentucky, United States, 2007	PC	Dating simulator
SOLVE[a] [30]	Men who have sex with men	18-24	United States, 2013	PC	3D dating simulator
HIV risk game [31]	Youth	15-19	Cape Town, South Africa, 2015	PC	Quizzes
Keep it up! [32-34]	YMSM[b]	18-29	Atlanta, Chicago, and Georgia, United States, 2017	PC	Dating simulator and minigames
BattleViro [35,36]	Young patients of antiretroviral therapy	14-26	Mississippi, United States, 2018	Smartphone (only IOS[c])	Twin-stick shooter and quizzes
Viral Combat [37]	YMSM	18-35	Mississippi, United States, 2021	Smartphone (only IOS)	Twin-stick shooter and quizzes
MyPEEPS [38-40]	Male youth sexually attracted to men	13-18	United States, 2019	PC and smartphone (any; usable as a web application)	Role-playing games
First-person scenario game[d] [41,42]	College students	17-27	Hong Kong—China, 2020	Smartphone (nonspecified)	First-person dating simulator

[a]SOLVE: Socially Optimized Learning in Virtual Environments.

[b]YMSM: young men who have sex with men.

[c]IOS: iPhone Operating System.

[d]See Multimedia Appendix 2.

Identified Games

The Baby Game!

The Baby Game [27] is a management simulator in which high school students are asked to simulate a budget and schedule based on different scenarios. The students' task is to establish how many hours they could devote to different activities (eg, chores, homework, sleep, recreation, and caring for their baby). They receive feedback, printed on a scorecard, based on how close their schedules are to a hidden correct time distribution.

The game aims to provide realistic information about the life changes that would occur if a student had a baby and how the newly added responsibility might affect their lives, with the assumption that this would enhance young people's intentions of delaying parenthood and using contraceptives.

Romance

Romance [27] is a text adventure in which high school students write down how they will deal with a set of scenarios of romantic and sexual nature. They then receive feedback in the form of a simulated outcome. At the end of their run, players obtain a final scorecard based on the adequacy of their decisions. The exercise aims to improve students' knowledge about sexuality and contraception, increase their skills for interaction, and serve as a practice for responsible sexual decision-making.

VODO

VODO [28] is a text adventure game in which high school students aged 15 years have to guide the main character through a series of scenarios. The game presents the player with a detailed written description of a situation; for example, "You are in your room. It is a sunny room full of things that are important to you. Tell the computer what you want to do?" The players then respond using simple English sentences; the game has an extensive vocabulary and is able to anticipate the responses typically provided by the students. Efforts were made so that although the player needs to make many choices, decisions are not presented overtly. This was done because the researchers wanted to convey the lesson that one has a choice, even when apparent conditions suggest otherwise.

An important aspect of this game is that it includes a roster of nonplayer characters (NPCs) with whom a player can interact and even form relationships that may or may not involve sex. Each of the NPCs has different names, personalities, and motives. In cases where the player chooses to have unprotected sex, the game creates a scenario in which the character has a child. The child randomly cries for different reasons and requires careful attention, creating tension between the character and their friends.

VODO was designed to improve participants' decision-making skills by providing a scenario in which they were able to rehearse and obtain feedback on their choices. The topics presented in the game were broad. Although they are focused on matters of sexual health (eg, contraceptive use, STIs, and the consequences of unwanted pregnancies), it also includes other issues that might affect teenagers (eg, drunk driving, drug use, and the ability to be alone without being lonely). Strategies such as complementary quizzes were meant to increase real-life communication about sex within the family.

Choose Your Own Adventure

Choose your own adventure [29] is the name that we have provided for 1 of the 6 modules that formed an unnamed intervention aimed at reducing rates of unintended pregnancy and STIs in adolescents from rural areas in the United States.

The game comprises half of one of the modules. Players are expected to play through a virtual date and make choices that could put them in a situation where their dates want to have sex, but they do not. The game finishes with different positive or negative outcomes and products of the in-game decisions that were taken. To make the game more engaging and increase its replay value, the developers built in some remarkable elements. For example, they included >150 images of various people, places, and STIs, which were randomly selected at various points in the game so that they would be different during each run. Furthermore, all in-game dialogs were recorded, and the NPCs actually spoke to the players. The other half of the module comprised submitting an original refusal line. The researchers reported that the entire module (game+refusal line submission) had a completion rate of 41%.

Socially Optimized Learning in Virtual Environments

SOLVE (Socially Optimized Learning in Virtual Environments) [30] is a 3D dating simulator aimed at men who have sex with men (MSM) aged 18 to 24 years who reported having engaged in recent unprotected anal intercourse (UAI). The settings are constructed around different scenarios that might be faced by young MSM involving some form of sexual decision (eg, meeting someone at a party and going to their apartment afterward). In each situation, the player encounters a series of *choice points* where they need to make self-regulatory decisions (eg, accepting or refusing alcohol or offers of casual sex). After choosing to engage (or not) in virtual sex, there is a customized recap sequence in which the player's virtual behavior is shown in sequence so that he can identify the different decisions that led to a particular outcome (Figure 2).

The idea was that through rehearsal and feedback, players could practice their decision-making skills. Throughout the process, they are guided by different NPCs (peers and one's virtual future self) who instruct them to follow a set of guidelines when faced with risky situations.

Figure 2. SOLVE (Socially Optimized Learning in Virtual Environments) avatars [30].

HIV Risk Game

This intervention [31], unnamed in the published article, is a relatively simple game in which youth are expected to identify who was more likely to have HIV between 2 randomly generated individuals. Each participant plays 10 rounds of the game. Instead of receiving a direct answer regarding whether they made the right choice, the participants receive information about HIV and risk and construct their own learning based on their experimentation.

Keep it Up!

Keep it up! [32-34] is a 7-module, multi-method intervention (one that includes the use of >1 method of data collection in a study) aimed at improving STI prevention strategies among young MSM in the United States. The main gaming component is called The Club Game. This game uses a real-life scenario (going to the club) to explore decision-making around using condoms; the steps to use condoms properly; and the effects of excessive alcohol consumption, drug use, and sexual arousal on decision-making. The player goes through 5 rooms and interacts with other patrons while completing the activities related to the abovementioned topics. The intervention uses diverse delivery methods (eg, videos, animation, and games) to improve HIV knowledge, motivate safer behaviors, teach skills, and increase self-efficacy for preventive behaviors.

BattleViro

BattleViro [35] is a twin-stick shooter mobile game aimed at improving antiretroviral treatment adherence among young MSM in the United States. During the game, players control an avatar that is shrunken down to fight viruses and other infections in 6 levels of increasing challenge. Each level is set on a specific organ ranging from the lungs to the brain. Throughout the different levels, the player shoots down threats to the host's body while picking up health points in the form of medicine (Figure 3). The character also receives messages from health care personnel, encouraging them to carry on and providing clues in challenging areas of the run. In addition, the player might answer quizzes from clinician avatars to earn additional points or powers. Wrong answers are corrected and explained. In addition to the game, participants with perfect adherence would receive congratulatory texts, whereas the other participants would receive motivational messages encouraging them to carry on.

Figure 3. BattleViro gameplay [35].

Viral Combat

Viral Combat [37] was developed by the same team that created *BattleViro*. However, instead of targeting people already with HIV, this game attempts to promote PrEP adherence. The levels are slightly different, as are the messages received by physicians and nurses; however, the main mechanics are similar. The game includes quizzes that go beyond PrEP adherence, including information on HIV and other STIs.

MyPEEPS

MyPEEPS [38-40] is a role-playing game in which young MSM with little to no sexual experience go through different scenarios guided by 4 characters (the *peeps*) who teach them about sexual health care. The game comprises 4 sequential modules

(PEEPScapades). The completion of the different modules is incentivized by in-app trophies.

First-person Scenario Game

First-person scenario game (*FPSG*) is the name we have provided for a multi-method intervention that aims to protect university students from the risks of using dating apps. The intervention comprises short informative videos in which students are taught about different risks, such as sexual abuse and scams. It includes a first-person simulation game in which the participant is presented with multiple choices when faced with real-life scenarios (Figure 4). The game was designed with various algorithms that resulted in positive or adverse outcomes, depending on the character choices.

Figure 4. First-person scenario game example screens [41] (translation available in Multimedia Appendix 2).

Results of Studies

In the following sections, we discuss the main results of our review, structured around the findings related to the target populations, interventions, outcomes, and study designs.

Target Population

Finding games that were targeted exclusively at youth aged between 15 and 25 years was not an easy task. Several identified interventions included some participants outside our selected age bracket, especially at the younger end of the age range. This was likely due, in part, to the physical location where the interventions were delivered. For example, several interventions were conducted in high schools [27-29], including students from both within and outside our age group. We did not find many studies that included participants outside the upper end of our age range. We found only 1 trial in which recruitment was done in colleges; however, even in this particular setting, the number of students aged >25 years (our upper age limit) was small [41].

Concerning gender and sexual orientation, many recent studies have focused on MSM and the prevention of STIs. A total of 7 studies had been conducted over the past 10 years; 5 of these focused exclusively on MSM. All but one of the games targeted urban youth; only 1 focused on rural populations [29].

On the basis of target population profiles, we divided the studies into 2 broad categories. First, there were those that were based in an educational institution (high school or university), included people of all genders and sexual orientations, and focused on contraception and STIs [27-29,41,42]. The second group was most commonly recruited from youth centers or sexual health clinics, focused exclusively on MSM, and had a clear focus on STI prevention and management [30,32,34,35,37,43,44].

Interventions

Overview

In the following sections, we describe the elements of the 11 games. However, first, it is important to note how little information about the games themselves was readily available in published articles. Authors often devoted little space to describing how the games looked and how they were played. Considering that playing games is a visual activity, the inclusion of images could have been a valuable way of providing this information. However, with some noteworthy exceptions [27,30,35,41], this was rarely performed; most articles did not include any form of visual aid to explain the game. Some articles included links to demos or webpages where the content was said to be available; however, in >1 case, the links were broken, or the page had already ceased to exist. Owing to the rapid nature of this review, no author and game developers were contacted during the process.

Game Age

Regarding the age of the games themselves, we identified a clear bimodal distribution; 3 games were created in the wake of the digital revolution ≥30 years ago, whereas the remaining group was developed more recently. Earlier games were graphically simpler but in no way less informative. We found it striking that from very early on, video games were seen as valuable tools to promote learning and attitude change.

Gaming Platform

The interventions used different platforms to host the games. Considering the time span across which the games were developed, it is not surprising that the most commonly used platform was that with the longest history, the PC. The changes in the games reflect the technological evolution that has affected this platform. Although popular, tablets were only mentioned in one of the games [33]. As might be expected, when mobile phones became widespread, the idea of using smartphones as platforms for serious gaming became an option. Of the 4 reviewed games for mobile phones, 2 (50%) were exclusively developed for the iPhone [35,37], whereas the other 2 (50%) did not specify which operating system they worked with [38,42]. These later games used mobile connectivity to increase participant engagement through push notifications and messages [35,36]. However, it is noteworthy that most interventions did not appear to use patches and updates to freshen their content, even when they were constantly connected to the internet.

Accessibility and Difficulty

A decision that is implicit in platform selection is accessibility. Most interventions aspire to be easily scalable. To achieve this, they must operate on an already popular and ubiquitous platform and use relatively little graphic processing power. We do not have access to the system requirements for any of the listed games. However, based on their description, it appears most operated on or below the considered average computing power at the time of their release.

With regard to game difficulty, all interventions can be placed on a continuum. On one end, there are games that provide an easy experience; that is, the game is seen as a vehicle through which information can be provided in an entertaining fashion. On the other end, we have games that provide a very demanding experience. The best example of this in the games we identified is *VODO*. The developers included a section of the game in which the player needed to answer 2 questions about sexual health. If either of these answers was incorrect, the game was over, and the player had to start all over again. This type of failed outcome was a very common scenario of text adventures at the time and motivated the player to replay the game several times, learning all its intricacies. An even more challenging aspect of these questions is that the answers were not provided in-game. The player was meant to search elsewhere for information; or even better, the topic should become a conversation starter for youth to discuss with their friends, parents, or teachers. This is an example of how a severely challenging task was used to *frustrate* the players into action.

As could be expected, these extremes in game difficulty were not without their issues. A nonchallenging game can be boring and can inhibit engagement. At the same time, a too-challenging game can tire a proportion of the players out of the intervention. In an effort to resolve these 2 positions, some game developers opted for an increasing level of difficulty; that is, instead of having a flat, low-level difficulty throughout the experience or a series of extreme spikes every so often, these games ramp up the difficulty with every level [35,37].

Expected Playthroughs and Playing Time

There are 2 related elements to consider here: expected playthroughs and expected contact time. In other words, how many times are players expected to play the game, and for how long? As we will discuss in the following sections, repetition, rehearsal, and feedback play an essential role in learning through games; therefore, it is expected that most games were designed to keep the player engaged in >1 run. Unfortunately, there is little information about these aspects in most of the included articles. Among the games that specified an expected playing time, we found periods of <1 hour of gaming. However, it was not always clear whether these times were for single or multiple playthroughs.

Type of Game and Game Setting

Game settings can be broadly divided into 2 types: realistic and science fiction. There was a clear preference for the latter in our sample of games. Most games were situated in locations and environments to which the player could directly relate, such as hanging out with friends after school or going to a house party. The idea behind this is that a greater similarity between the simulated situation and a plausible real-life event might make it more likely that players will relate to the content and act upon what they have learned. Game developers have gone into great efforts to create content grounded in reality, where situations that the players have directly experienced are portrayed both didactically and accurately. For example, *SOLVE* allowed players to personalize their avatars, and *The Baby Game* used actual prices when they calculated the costs of raising a child.

The types of games in the realistic group were highly diverse. Table 1 presents a list of the different styles of games that were included. We want to highlight 3 features. First, the variety is notable; as can be seen, games ranged from management simulators to role-playing games. The second feature is the relative preponderance of the dating simulators. Approximately 44% (4/9) of the nonfiction games were dating simulators. However, even within this specific setting, we found different styles (eg, text adventures, 3D, and chat simulators). Finally, the role of complementary activities in each game should be mentioned. Minigames are found in several games, particularly in the form of quizzes [29,34,36,37]. This seems to have been one of the main strategies through which game developers delivered specific sexual health knowledge.

However, not all games followed a realistic route. *BattleViro* and *Viral Combat* opted instead for science fiction in the action setting [35,37]. In both games, the characters are shrunk down to a microscopic scale and are meant to protect the human body from infections by shooting down viruses, bacteria, and vectors of disease. Here, the developers tried to create a power fantasy in which the player can take control of their actions and reach a desirable healthy state. The content and settings were still related to the topic of focus (antiretroviral treatment and PrEP adherence); however, the developers avoided making direct or explicit statements about them.

We can see that there is >1 valid strategy for promoting engagement with the material. One school of thought aims to create easy-to-relate experiences, whereas the other uses a fast-paced game to empower participants into action.

Single or Multiplayer Game (Private or Social Experiences)

An element mentioned throughout the different interventions is the contrast between creating a single or multiplayer experience. In other words, was the game designed to be played alone or with a group? Most of the games reviewed appear to have been designed for single-player use. However, it should be noted that all single-player experiences can be turned into multiplayer experiences by the players themselves. Researchers reported that, in several cases, games that were not meant to be social experiences were transformed into a group activity when a player spontaneously brought their friends or partner to play the game together and comment on it.

The decision to develop a single- or multiplayer game is affected by several factors. The first is the target population. Interventions focused on MSM were very keen on not *outing* their players involuntarily or having them openly disclose their health information; hence, single-player games might have been preferred. Second, the game's topic of focus is an important factor to consider. Some topics are easier to work with at the individual level than at the group level. For example, *SOLVE* was a game that tried to decrease the feelings of shame that gay or bisexual men might experience regarding their sexual preferences [35]. Considering that many players had strong feelings of shame and were reticent to disclose information about their sexual interests, they may not have been comfortable playing a game with others.

The third influence is logistical, technological, or economic restriction. When PCs were not ordinary household items but specialized pieces of hardware, they were not as commonly available as they are now. For this reason, older games tended to be a social experience; many people had to use the same computer to make it viable for enough players to play the game [27,28].

The final reason for choosing single- or multiplayer games relates to the learning strategy of choice. Some game developers opted to purposely promote out-of-game discussions of sexual health topics [28]. The aim was to make the game a topic for discussion with family members, teachers, and friends.

Outcome Change Mechanisms

There are several ways in which we could try to classify the underlying mechanisms used in video games to change specific behaviors in users. Here, we divided the mechanisms into 3 categories: those based on knowledge, those focused on enhancing skills and self-efficacy, and those that motivate change through emotions. These groups are not mutually exclusive; 1 intervention might have >1 underlying mechanism.

Some games aim to provide knowledge, expecting that it will generate behavior change. For example, some games share facts about contraceptives, their efficacy, and the risks involved in not using them. In such cases, one of the most critical elements is to provide a clear and easy-to-understand message. It has been noted that most interventions try to make the message grounded in a specific element or situation in the game. Many games focusing on providing knowledge prioritized the provision of immediate and clear feedback, specifying where and when an error was made and what its potential outcomes might be. The same applies to decisions that have a positive outcome. For example, in *SOLVE*, when a player chose to engage (or not) in in-game sex, they were offered a quick recap of all the previous decisions that drove them to their current state (decisions that were not always apparent at the time they were made).

The final element of knowledge is how it is constructed. We have previously stated that the message must be clearly stated. However, for some interventions, this did not necessarily mean that the message had to be explicitly delivered. For example, the *HIV risk game* had a clear message that needed to be delivered: older people were more likely to have HIV than younger people. Players played 10 rounds of the game in which they made a judgment about which character was more likely to have HIV. As feedback, the players did not receive the correct answer; they only knew whether they were right or wrong. This key message was supposed to be inferred (constructed) by the participants based on their in-game experience.

Two of the most frequently used strategies in games focused on increasing the participants' skills and self-efficacy, which was achieved by a mixture of relatability and rehearsal. By relatability, we mean all the different factors that can make the situation in a game similar to the ones players face or think they will face. The developers made great efforts to provide experiences grounded in those that the players have had or will experience. The assumption is that, in general, the closer a setting and its characters are to the real world, the easier it will be for the player to assimilate the lesson and put it into practice. This is one of the reasons why several games designed characters with different personalities and stories so that the player can easily associate 1 or several of them with their friends and acquaintances. Similarly, one of the reasons why some avatars were customizable was to make it easier for players to empathize with their in-game presence. The same can be said of the setting in which the interactions occur. In several cases, the setting was very similar to that currently experienced by the players. One of the clearest cases of relatability is in the FPSG game. The player learns about the risks of dating apps by playing a game that uses an instant messaging app as one of its primary interfaces.

Similarly, rehearsal and repetition also played a significant role in improving self-efficacy. The idea is that players will train themselves to make safe decisions in real life because they have made the same correct decisions in a virtual world before. The more times a player does something, the more likely it is for him or her to feel (and be) proficient in it.

The final strategy relies on using emotions to generate a reaction in the player. There are several methods in which this has been performed in different games. Some developers opted to generate negative emotions that frustrated or scared players into action. For example, in *Romance*, if the players initiated unprotected sex, they would have a baby that would cry randomly during the game, negatively affecting their relationships with their friends. Other games used positive emotions to inspire players to act. *BattleViro* and *Viral Combat* are good examples of this practice, having used fast action, increasingly challenging shooter-style games to empower their players to take control of their treatment. Finally, there were games that aimed to reduce the negative emotions that inhibit players' ability to do something. The best example of this practice is *SOLVE*, a game whose main aim was to reduce the feelings of shame that MSM might experience. Through a series of stories and vignettes, the intention was that the player might consciously acknowledge their desires as something normal, which carries no stigma.

Game Development

We cover 3 main topics in this section. We begin with a general description of the development process of the games. We then assess the involvement of stakeholders in the creation of the game: who was invited, when, and in what capacity.

Development Process

The published papers provided little information on the development of game mechanics. There was often no data regarding how long the game design lasted, how much its budget was, who and how many people were involved, and what program or programs and engine or engines they used to create it. From conversations with researchers, we know that in some cases, university-based groups were in charge of software development. However, apart from *FPSG*, very little additional information is readily available from these articles or other related publications on game development.

Some interventions adapted previous activities or interventions for the construction of new games. In some cases, existing materials and activities from previous interventions were adapted to a video game form. The details of these interventions were usually left nebulous; however, we know that in the case of *MyPEEPS*, *Keep it up!*, and *SOLVE*, a considerable part of the content of the games was taken from previous non–video game–based interventions. For example, *MyPEEPS* included 4 characters (the titular *peeps*), who were a composite of previously existing characters used during the formative phase of the intervention.

The development of other interventions was probably informed by existing games, although few articles provided much detail about this. The only exception to this trend was *Viral Combat*, heavily influenced by *BattleViro* [37]. The same team of researchers developed both games, and one might even say that the former is an improved version of the latter.

In summary, we found that interventions have either been developed entirely from scratch or based on a previous in-person intervention. Explicit references to previously existing games were unusual in the reviewed studies.

Stakeholder Participation

The teams in charge of designing the interventions frequently made considerable efforts to involve different stakeholders throughout the process. Among the stakeholders who participated in the design of games were end users [32,35,41], members of nongovernmental organizations concerned with sexual health or youth well-being [34], and unspecified community leaders [28]. There is little to no mention of the involvement of parents, teachers, or other authority figures. Focus groups [34] and in-depth interviews [32,35] were used to access stakeholders' views.

There were 3 main reasons for stakeholder involvement. The first reason was to conduct a needs assessment. This allowed the intervention designers to prioritize topics or behaviors that required specific attention. For example, in the development of the *FPSG* intervention, 4 focus groups were held by the developer to identify key risks that caused concern among young people using dating apps in Hong Kong. The second reason to involve stakeholders was to improve the quality of the game itself. For example, *Keep it up!* conducted interviews with stakeholders to ensure that the situations and languages they used in their club games were similar to those experienced by young men in their everyday interactions [33]. This allowed them to generate greater engagement with the final users by presenting situations comparable with those they had experienced previously. Finally, approval from the governing body is needed. By involving community leaders and local authorities, intervention designers could ensure that they would receive support for the subsequent stages of the process. For example, *VODO* involved people from 30 different local institutions to avoid the inclusion of content or situations that might have been perceived as unacceptable by the community [28].

Multi-Method Interventions

A final element to discuss is that although all interventions relied considerably on video games to achieve their goals, it was not necessarily the only method they used. Approximately 27% (3/11) of the games were meant to be played in conjunction with other activities.

For these specific interventions, the games seem to be one of the few activities in which the participants could take agency and act upon the knowledge they received. For example, when the participants are completing scales or watching videos, they are fairly passive, and the moments in which they play the games are the only times when they really take control, make decisions, and see their results. Although no intervention specified the playing time, or the time used in the other modules, it appeared that the games were the activities that comprised most of the participants' time.

Outcomes

Overview

In this section, we discuss the effectiveness of the interventions in achieving their goals, organized by the outcomes adopted in our search criteria. Multimedia Appendix 3 provides a summary of the results.

Decrease in Unintended Pregnancies

Rather unexpectedly, none of the studies assessed the number of pregnancies. There are 2 possible explanations for this. First, almost half of the chosen games were not marketed to women but to MSM. Second, the sample sizes were too small, and the follow-up periods were too short, for the relatively low occurrence of pregnancy to become a viable measure of the success of an intervention.

Decrease in STIs

Only *Keep it up!* used STI biomarkers to assess changes in STI incidence. The researchers tested for chlamydia and gonorrhea through self-collection of rectal swabs. Through matched odds ratios, the control group showed a 55% increase in STI incidence, whereas the treatment group showed a decrease of 51%. These results were significant; however, we must be mindful that this intervention had multiple components, and the video game was only one of them.

Increased Contraceptive Use

Approximately 27% (3/11) of interventions measured changes in reported contraceptive use. Unfortunately, they did so in very different ways, which limited our ability to compare them. *Choose Your Own Adventure* asked about condom use at the last intercourse and found no effect of the intervention. The other 2 studies assessed the frequency of UAI with nonprimary partners during the past 3 months. Although *SOLVE* was unable to show significant differences between the treatment and control groups, *Keep it up!* reported a significant decrease in the number of UAI events 1 year after the start of the intervention.

In summary, studies on the effects of video game interventions on contraceptive use have shown inconsistent results.

Acquisition of New Knowledge Regarding Sexual Health

The acquisition of new knowledge was one of the most commonly measured outcomes; however, the topics and measures varied significantly among the different interventions. Frequently, ad hoc questionnaires were created to assess differences between the treatment and control groups. *The Baby Game* quizzed participants on the costs (both time and money) involved in taking care of a baby. *Romance* used the same methodology but compared knowledge about the efficacy of different contraceptive methods. Studies on both of these interventions suggested improved knowledge in the treatment group compared with the control group. However, we should keep in mind that knowledge was assessed only immediately after the game ended and that no effect size measure was presented. Other interventions also relied on ad hoc tests; however, it was unclear exactly what topics they explored. *Choose Your Own Adventure* showed positive results (of medium effect size); however, *VODO* failed to do so. Both interventions followed a pretest-posttest design.

A comparable example is that of *BattleViro* and *Viral Combat.* They both tested their participants' HIV knowledge, and although they were very similar games, only *BattleViro* showed positive results. It should be noted that *BattleViro* tested their participants 16 weeks after the intervention started, whereas *Viral Combat* did so at weeks 12 and 24. The extended period between intervention and data collection in *Viral Combat*, compounded by attrition of 32% of the original sample, might have biased the intervention results. However, one would also expect that participants who stayed longer would be more engaged, would have clocked in more hours in the game, and would have a better overall performance.

In conclusion, we do not have enough evidence to clearly state that games have a significant effect on increasing knowledge of sexual health topics. The variety of topics assessed, the limited amount of information regarding the content of the tests, and the large variability in the time between intervention and postintervention assessment precludes our ability to establish a clear causal relationship between playing and learning.

Changes in Perceived Risk of Pregnancy

Only 2 games assessed this variable. *Romance* asked participants to assess the odds of becoming pregnant when having unprotected sex. Improvements in favor of the treatment group immediately after they finished playing the game were reported. Unexpectedly, a study on *Choose Your Own Adventure* found significant differences in favor of the control group. However, the researchers measured a construct called *susceptibility*, which merged the perceived risk of pregnancy with the perceived risk of STIs; thus, there might have been some cross-contamination in the assessment. The authors speculated that their results might be related to the fact that fewer people initiated sexual activities in the treatment group (ie, they were abstinent) than in the control group; hence, they did not feel at risk of any adverse outcomes related to having sex. They also considered that as the treatment group was more aware of the risks and the measures they could take against them, they felt better able to protect themselves.

Changes in Perceived Risk of STIs

Approximately 36% (4/11) of games addressed this topic. We have already discussed the findings of *Choose Your Own Adventure*. The entire intervention of the *HIV risk game* was centered on assessing the risk of someone having HIV based on their age and gender. Positive results were obtained for both men and women. The assessment was performed immediately after the intervention and 3 months later (the last time only for male participants).

BattleViro and *MyPEEPS* also assessed the perceived risk of STIs but in an indirect fashion by asking about STI testing. *BattleViro* measured the types of sexual behavior, frequency of sex, and number and gender of partners reported in the past 3 months. The authors found no differences between the control and treatment groups. *MyPEEPS* measured the frequency of STI testing and found that after the intervention, those in the treatment group were more likely to get tested than those in the control group.

Attitudinal Change Toward Safe Sex

Approximately 18% (2/11) of interventions targeted attitudinal changes toward safe sex. *VODO* measured participants' attitudes toward sex on 2 axes: liberal versus conservative and positive versus negative. Both the control and treatment groups shifted to a more liberal position; however, the change was greater in the treatment group. The change in this group was sufficient for it to move, on average, from a conservative perspective toward a liberal one. No significant changes were observed in the positive and negative axes.

Choose Your Own Adventure assessed the predisposition toward waiting to have sex and found a significant effect of their intervention among students aged between 15 and 16 years. After completing ≥1 of its modules, participants were more likely than those in the control group to postpone sexual initiation.

Self-efficacy Toward Sexual Health

Self-efficacy, be it general or specific to sexuality, was one of the most frequently chosen outcome variables across the different games. *Choose Your Own Adventure* considered 4 domains of sexual self-efficacy: condom negotiation self-efficacy, condom use self-efficacy, situational self-efficacy (the ability to control a situation that might be conducive to sex), and refusal self-efficacy (the ability to say no to sexual intercourse). However, the findings were mixed. The intervention increased participants' self-efficacy toward condom negotiation and situational self-efficacy; however, no effects were found regarding condom use or refusal. It is not surprising that condom use self-efficacy did not change considerably between groups as the intervention did not include any components that directly taught students how to apply and use a condom. However, it did have 1 activity specifically focused on improving refusal skills, which was tightly tied to the game itself. The fact that this activity failed to produce the desired results for this variable is noteworthy.

BattleViro and *Viral Combat* assessed a similar domain of self-efficacy: participants' belief that they would be able to adhere to a treatment regime. Both interventions found no significant change in either the short (12 weeks) or long-term (16 weeks and 24 weeks) assessments. *MyPEEPS* reported positive results when assessing HIV self-efficacy in the short term (3 months). Researchers have yet to publish their results for the long-term assessments (6 months).

We consider that the results on self-efficacy are mixed. Considering that the interventions are varied in methodology and topics and that they have worked on different domains of self-efficacy, this is not a particularly surprising result.

Decrease in the Number of Sexual Partners

Approximately 27% (3/11) of interventions aimed at reducing the number of sexual partners, all of them for MSM—*Keep it up!*, *Viral Combat*, and *MyPEEPS*—and none resulted in a reduction in the number of sexual partners in their samples.

Increase in Adherence to Prophylaxis or Treatment

Approximately 27% (3/11) of interventions focused on increasing adherence to either treatment or PrEP. *BattleViro* and *Viral Combat* used a mixture of bioindicators, self-reported behavior, and electronic device follow-ups to assess this outcome. A total of 2 bioindicators were used: HIV-1 viral load in *BattleViro* and 1ARV (activator protein 1) levels in *Viral Combat*. *BattleViro* produced equivalent decreases in HIV-1 viral load in both treatment and control groups. *Viral Combat* reported results that favored the treatment group at both 12 weeks and 24 weeks after the beginning of the intervention; however, these findings were not statistically significant. Treatment adherence, measured by self-report in *Viral Combat* and by self-report plus electronic device records in *BattleViro*, showed similar results. In this regard, no intervention showed better results than the usual treatment. *MyPEEPS* also measured PrEP and postexposure prophylaxis adherence using self-reported measures and found no significant differences between the treatment and control groups.

Study Designs

Bearing in mind that our search criteria only allowed for RCTs and quasi-RCTs, we identified 2 main study designs. Studies were either posttest-only trials [31,37,43] or pretest-posttest trials [27-30,33,35,36,42].

Depending on the study, the control group received different treatments. Waiting-list control was one of the most straightforward control designs. A more complex one was treatment as usual (TAU), where the usual or standard was given to a group of participants.

For example, for *The Baby Game*, researchers compared their game with a regular sexual health education class for that specific age group [28]. A similar option was TAU+. Here, the participants received TAU and an additional component that was functionally similar but thematically different from the experimental group. For example, as *BattleViro* provided smartphones so that participants could play the game, they also provided smartphones to the control group. However, these iPhones did not have the specific game installed but another non–HIV-related game [35]. Another form of control group provided more or less the same content as the game but in a delivery mode that had no ludic or interactive elements. For example, *Keep it up!* provided an internet-based experience with the same information as their intervention but using static slides instead of the more dynamic approach taken with the treatment group.

Finally, in reviewing whether the studies had adequate sample sizes, we found that the sample sizes were generally large enough to detect expected differences. Most studies, especially the more recent ones, determined their sample size based on a power analysis (although this analysis was usually constructed around educated guesses). Even if the sample size in the reviewed studies was usually large enough, one of the main threats to statistical power was a relatively large attrition rate, especially among studies with multiple or long follow-ups. The most extreme case was in the *HIV risk game* study, which reported an attrition rate of 66.8%. In the remaining studies, the attrition rate was approximately 30%.

Discussion

Principal Findings

The findings of our review yielded important conclusions and implications for future research and game development. First, the findings highlighted considerable diversity in video game–based interventions. Although all of them addressed similar topics, they did so in fairly distinct ways. The outcomes assessed in studies evaluating games were also very diverse and, even when similar, were measured differently in each study. Second, we found that game developers have made great efforts to elicit experiences tailored to the specific needs of the targeted population, most often achieving this through regular stakeholder participation activities throughout the game development process.

One of the most surprising findings was the age of the identified games. The fact that the games were developed over such an extended period suggests that even when the graphical complexity and the interface changed considerably, learning through gaming is and has been seen as a viable and successful strategy.

However, although game interventions for sexual health have been in existence for almost three decades, relatively few studies have evaluated them, and the results of previous studies have been mixed. Moreover, there is little clarity regarding which specific elements of a game facilitate a positive outcome. This is partly because of the diversity of the behavior change mechanisms underlying interventions, the variety of the games themselves, the populations they target, the outcomes measured, and how these are measured. All these differences make it challenging to identify a clear causal link between playing a game and improving an aspect of sexual health. However, although the impact on sexual health is not always clear, the fact that video game–based interventions are of interest to most young people is well-established.

Nevertheless, there are other less positive aspects of research in this area that we need to acknowledge. First, there is a lack of information available in published reports on different games, especially in the gameplay aspect. This ties to another unexpected finding of our review. Although video game–based interventions are meant to be easily scalable, there are no reports on any of these games being picked up for broad distribution. In fact, only *BattleViro* was readily available for download. Moreover, although the interventions were very varied, almost all of them targeted people living in urban areas of the United States. There were no interventions developed with populations from lower-income countries in mind, and almost no interventions were aimed at rural populations. Finally, very few interventions were informed by a behavioral theory or model.

Strengths and Limitations

Our review had some notable strengths. We followed the Cochrane Rapid Reviews Method Group guidelines [22,24] to conduct the review and searched several key literature databases. A second reviewer was involved in screening 20% (51/257) of the articles at the title and abstract screening stage and 50% (35/70) of the articles at the full-text screening stage.

Some limitations of our review should also be acknowledged. As this review was rapid, our search used 5 databases; thus, we may not have identified all the relevant literature. We restricted our age range to 15 to 25 years and, in the screening process, noticed that some interesting game-based interventions focused on younger adolescents and children. A final limitation was that we did not exclude studies based on quality.

Conclusions and Recommendations

In conclusion, we do not have enough evidence to clearly state that games have a significant effect on sexual health among young people. The interventions and how they were evaluated were too diverse to reach a clear conclusion. However, based on the original authors’ criteria for success, we have compiled a set of recommendations for developing game-based interventions to improve sexual health in young people (Textbox 1).

Textbox 1. Recommendations for developing game-based interventions to improve sexual health in young people.

Recommendations for developing game-based interventions to improve sexual health

- Stakeholders should be involved in different stages of the game development process. Most successful strategies used qualitative participatory methods involving multiple stakeholders.
- A pilot phase in the development of games is strongly encouraged. This enables specific elements that could otherwise jeopardize the success of the initiative to be identified and modified.
- One of the most crucial decisions during the game development process is whether intervention participants are expected to go through the game once or multiple times. This affects the length, difficulty setting, and the main mechanics of the game.
- There are several viable ways in which a game can try to change someone's behavior. None have proven to be markedly better than the others. However, some recommendations are as follows:
  - Knowledge-based interventions should aim to provide a clear message, and this message does not need to be explicit. In fact, some researchers recommend that the message is not explicitly stated but constructed by the players themselves.
  - Self-efficacy and skill-building interventions aim to provide easily relatable experiences and those that feel proximal to the player. The closer the player feels a game experience is to their own experience, the more likely it is that they will act upon it. The game serves as a rehearsal for the decisions they will make in real life.
- Disregarding the mechanisms chosen by developers, some common elements are shared by most strategies:
  - Feedback is better if it is clear, detailed, and immediate. When playing, it is encouraged that users recap their decisions and learn which actions drove them to their current stage (whether positive or negative).
  - Repetition (as long as it does not transform into tediousness) is usually favorable, especially for skill building and knowledge acquisition.
- There are 3 common threats that plague these interventions:
  - Lack of technical support, especially after the game development phase ends, is a common threat.
  - Another threat is the stagnation of the content; that is, no updates are provided, and no new content is delivered.
  - The games are not easily found when someone wants to use them in other contexts. The created game should be openly available on the web if possible.
- Game quality indicators (including playing time) were registered using self-reported measures. A suitable workaround using in-game data collection is recommended to bypass social desirability and recall issues that affect purely self-reported information.
- A plan for the implementation of the intervention should be made at the early stage of the project.

Acknowledgments

This study was funded by an Erasmus+ grant (project number 2020-2-UK01-KA205-079645).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[PDF File (Adobe PDF File), 117 KB - games_v10i2e33207_app1.pdf]

Multimedia Appendix 2
First-person scenario game example—English translation.
[PDF File (Adobe PDF File), 240 KB - games_v10i2e33207_app2.pdf]

Multimedia Appendix 3
Summary of results per intervention.
[DOCX File , 16 KB - games_v10i2e33207_app3.docx]

References

1. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. Bull World Health Organ 2019 Aug 01;97(8):548-62P [FREE Full text] [doi: 10.2471/BLT.18.228486] [Medline: 31384073]
2. Report on global sexually transmitted infection surveillance 2018. World Health Organization. 2018 Dec 19. URL: https://apps.who.int/iris/rest/bitstreams/1170799/retrieve [accessed 2022-05-13]
3. Sexually transmitted infections and screening for chlamydia in England, 2018: health protection report. Public Health England. 2019 Jun 4. URL: https://pcwhf.co.uk/wp-content/uploads/2019/06/hpr1919_stis-ncsp_ann18.pdf [accessed 2022-05-13]
4. Woog V, Singh S, Browne A, Philbin J. Adolescent women's need for and use of sexual and reproductive health services in developing countries. Guttmacher Institute. 2015 Aug. URL: https://www.guttmacher.org/report/adolescent-womens-need-and-use-sexual-and-reproductive-health-services-developing-countries [accessed 2022-05-13]
5. Mahase E. WHO reports "concerning lack of progress" in stopping STI spread. BMJ 2019 Jun 07;365:l4124. [doi: 10.1136/bmj.l4124] [Medline: 31175082]
6. Holmes KK, Levine R, Weaver M. Effectiveness of condoms in preventing sexually transmitted infections. Bull World Health Organ 2004 Jun;82(6):454-461 [FREE Full text] [Medline: 15356939]
7. Chanakira E, O'Cathain A, Goyder EC, Freeman JV. Factors perceived to influence risky sexual behaviours among university students in the United Kingdom: a qualitative telephone interview study. BMC Public Health 2014 Oct 09;14:1055 [FREE Full text] [doi: 10.1186/1471-2458-14-1055] [Medline: 25300195]
8. Sanders SA, Yarber WL, Kaufman EL, Crosby RA, Graham CA, Milhausen RR. Condom use errors and problems: a global view. Sex Health 2012 Mar;9(1):81-95. [doi: 10.1071/SH11095] [Medline: 22348636]
9. Free C, Roberts IG, Abramsky T, Fitzgerald M, Wensley F. A systematic review of randomised controlled trials of interventions promoting effective condom use. J Epidemiol Community Health 2011 Feb;65(2):100-110 [FREE Full text] [doi: 10.1136/jech.2008.085456] [Medline: 19822557]
10. Denford S, Abraham C, Campbell R, Busse H. A comprehensive review of reviews of school-based interventions to improve sexual-health. Health Psychol Rev 2017 Mar;11(1):33-52. [doi: 10.1080/17437199.2016.1240625] [Medline: 27677440]
11. Bearak J, Popinchalk A, Ganatra B, Moller A, Tunçalp Ö, Beavin C, et al. Unintended pregnancy and abortion by income, region, and the legal status of abortion: estimates from a comprehensive model for 1990-2019. Lancet Glob Health 2020 Sep;8(9):e1152-e1161 [FREE Full text] [doi: 10.1016/S2214-109X(20)30315-6] [Medline: 32710833]
12. Scott RH, Wellings K, Lindberg L. Adolescent sexual activity, contraceptive use, and pregnancy in Britain and the U.S.: a multidecade comparison. J Adolesc Health 2020 May;66(5):582-588 [FREE Full text] [doi: 10.1016/j.jadohealth.2019.11.310] [Medline: 32029341]
13. Sedgh G, Finer LB, Bankole A, Eilers MA, Singh S. Adolescent pregnancy, birth, and abortion rates across countries: levels and recent trends. J Adolesc Health 2015 Feb;56(2):223-230 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.09.007] [Medline: 25620306]
14. Kirby DB. The impact of abstinence and comprehensive sex and STD/HIV education programs on adolescent sexual behavior. Sex Res Soc Policy 2008 Sep;5(3):18-27. [doi: 10.1525/srsp.2008.5.3.18]
15. Santelli J, Ott MA, Lyon M, Rogers J, Summers D, Schleifer R. Abstinence and abstinence-only education: a review of U.S. policies and programs. J Adolesc Health 2006 Jan;38(1):72-81. [doi: 10.1016/j.jadohealth.2005.10.006] [Medline: 16387256]
16. Simon L, Daneback K. Adolescents' use of the Internet for sex education: a thematic and critical review of the literature. Int J Sex Health 2013 Oct;25(4):305-319. [doi: 10.1080/19317611.2013.823899]
17. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
18. Stapleton AJ. Serious games: serious opportunities. In: Australian Game Developers' Conference, Academic Summit. 2004 Presented at: AGDC '04; December 2-4, 2004; Melbourne, Australia.
19. Korteling H, Helsdingen A, Theunissen NC. Serious gaming @ work: learning job-related competencies using serious gaming. In: Derks D, Bakker A, editors. The Psychology of Digital Media at Work. London, UK: Psychology Press; 2012:129-150.
20. Dondlinger MJ. Educational video game design: a review of the literature. J Appl Educ Technol 2007;4(1):21-31.
21. Beck K. Video games should be used to teach sex education. Mashable. 2018 Sep 24. URL: https://mashable.com/article/video-games-sex-ed/?europe=true [accessed 2021-03-20]
22. Garritty C, Gartlehner G, Nussbaumer-Streit B, King VJ, Hamel C, Kamel C, et al. Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. J Clin Epidemiol 2021 Feb;130:13-22 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.10.007] [Medline: 33068715]
23. Garritty C, Gartlehner G, Kamel C, King VJ, Nussbaumer-Streit B, Stevens A, et al. Cochrane Rapid Reviews: Interim Guidance from the Cochrane Rapid Reviews Methods Group. Cochrane Methods Rapid Reviews. 2020 Mar. URL: http://methods.cochrane.org/sites/methods.cochrane.org.rapidreviews/files/uploads/cochrane_rr_-_guidance-23mar2020-final.pdf [accessed 2022-05-13]

24. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci 2010 Jul 19;5:56 [FREE Full text] [doi: 10.1186/1748-5908-5-56] [Medline: 20642853]
25. Kapp KM. The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education. San Francisco, CA, USA: Pfeiffer; 2012.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
27. Paperny DM, Starn JR. Adolescent pregnancy prevention by health education computer games: computer-assisted instruction of knowledge and attitudes. Pediatrics 1989 May;83(5):742-752. [Medline: 2654867]
28. Alemi F, Cherry F, Meffert G. Rehearsing decisions may help teenagers: an evaluation of a simulation game. Comput Biol Med 1989;19(4):283-290. [doi: 10.1016/0010-4825(89)90015-2] [Medline: 2680255]
29. Roberto AJ, Zimmerman RS, Carlyle KE, Abner EL. A computer-based approach to preventing pregnancy, STD, and HIV in rural adolescents. J Health Commun 2007;12(1):53-76. [doi: 10.1080/10810730601096622] [Medline: 17365349]
30. Christensen JL, Miller LC, Appleby PR, Corsbie-Massay C, Godoy CG, Marsella SC, et al. Reducing shame in a game that predicts HIV risk reduction for young adult MSM: a randomized trial delivered nationally over the Web. J Int AIDS Soc 2013 Nov 13;16(3 Suppl 2):18716 [FREE Full text] [doi: 10.7448/IAS.16.3.18716] [Medline: 24242264]
31. Datta S, Burns J, Maughan-Brown B, Darling M, Eyal K. Risking it all for love? Resetting beliefs about HIV risk among low-income South African teens. J Econ Behav Organ 2015 Oct;118:184-198. [doi: 10.1016/j.jebo.2015.02.020]
32. Motley DN, Hammond S, Mustanski B. Strategies chosen by YMSM during goal setting to reduce risk for HIV and other sexually transmitted infections: results from the Keep It Up! 2.0 prevention trial. AIDS Educ Prev 2017 Feb;29(1):1-13 [FREE Full text] [doi: 10.1521/aeap.2017.29.1.1] [Medline: 28195780]
33. Mustanski B, Madkins K, Greene GJ, Parsons JT, Johnson BA, Sullivan P, et al. Internet-based HIV prevention with at-home sexually transmitted infection testing for young men having sex with men: study protocol of a randomized controlled trial of Keep It Up! 2.0. JMIR Res Protoc 2017 Jan 07;6(1):e1 [FREE Full text] [doi: 10.2196/resprot.5740] [Medline: 28062389]
34. Mustanski B, Parsons JT, Sullivan PS, Madkins K, Rosenberg E, Swann G. Biomedical and behavioral outcomes of Keep It Up!: an eHealth HIV prevention program RCT. Am J Prev Med 2018 Aug;55(2):151-158 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.026] [Medline: 29937115]
35. Whiteley L, Brown LK, Mena L, Craker L, Arnold T. Enhancing health among youth living with HIV using an iPhone game. AIDS Care 2018;30(sup4):21-33 [FREE Full text] [doi: 10.1080/09540121.2018.1503224] [Medline: 30626196]
36. Whiteley L, Mena L, Craker LK, Healy MG, Brown LK. Creating a theoretically grounded gaming app to increase adherence to pre-exposure prophylaxis: lessons from the development of the viral combat mobile phone game. JMIR Serious Games 2019 Mar 27;7(1):e11861 [FREE Full text] [doi: 10.2196/11861] [Medline: 30916652]
37. Whiteley L, Craker L, Haubrick KK, Arnold T, Mena L, Olsen E, et al. The impact of a mobile gaming intervention to increase adherence to pre-exposure prophylaxis. AIDS Behav 2021 Jun;25(6):1884-1889. [doi: 10.1007/s10461-020-03118-3] [Medline: 33483897]
38. Kuhns LM, Garofalo R, Hidalgo M, Hirshfield S, Pearson C, Bruce J, et al. A randomized controlled efficacy trial of an mHealth HIV prevention intervention for sexual minority young men: MyPEEPS mobile study protocol. BMC Public Health 2020 Jan 15;20(1):65 [FREE Full text] [doi: 10.1186/s12889-020-8180-4] [Medline: 31941475]
39. Schnall R, Kuhns L, Pearson C, Bruce J, Batey DS, Radix A, et al. Preliminary results from a pragmatic clinical trial of MYPEEPS mobile to improve HIV prevention behaviors in young men. Stud Health Technol Inform 2020 Jun 16;270:1365-1366 [FREE Full text] [doi: 10.3233/SHTI200444] [Medline: 32570661]
40. Ignacio M, Garofalo R, Pearson C, Kuhns LM, Bruce J, Scott Batey DS, et al. Pilot feasibility trial of the MyPEEPS mobile app to reduce sexual risk among young men in 4 cities. JAMIA Open 2019 Jul;2(2):272-279 [FREE Full text] [doi: 10.1093/jamiaopen/ooz008] [Medline: 31294422]
41. Wong WC, Song L, See C, Lau ST, Sun WH, Choi KW, et al. Using crowdsourcing to develop a peer-led intervention for safer dating app use: pilot study. JMIR Form Res 2020 Apr 21;4(4):e12098 [FREE Full text] [doi: 10.2196/12098] [Medline: 32314975]
42. Wong WC, Sun WH, Chia SM, Tucker JD, Mak WP, Song L, et al. Effectiveness of a peer-led Web-based intervention to improve general self-efficacy in using dating apps among young adults: randomized clustered trial. J Med Internet Res 2020 Oct 30;22(10):e16378 [FREE Full text] [doi: 10.2196/16378] [Medline: 33124987]
43. Hidalgo MA, Kuhns LM, Hotton AL, Johnson AK, Mustanski B, Garofalo R. The MyPEEPS randomized controlled trial: a pilot of preliminary efficacy, feasibility, and acceptability of a group-level, HIV risk reduction intervention for young men who have sex with men. Arch Sex Behav 2015 Feb;44(2):475-485 [FREE Full text] [doi: 10.1007/s10508-014-0347-6] [Medline: 25135064]
44. Gordián-Arroyo A, Garofalo R, Kuhns LM, Pearson C, Bruce J, Batey DS, et al. Awareness, willingness, and perceived efficacy of pre-exposure prophylaxis among adolescent sexual minority males. J Urban Health 2020 Oct;97(5):749-757 [FREE Full text] [doi: 10.1007/s11524-020-00447-5] [Medline: 32789625]

Abbreviations

FPSG: first-person scenario game
MSM: men who have sex with men
NPC: nonplayer character
PrEP: pre-exposure prophylaxis
RCT: randomized controlled trial
SOLVE: Socially Optimized Learning in Virtual Environments
STI: sexually transmitted infection
TAU: treatment as usual
UAI: unprotected anal intercourse

Edited by N Zary; submitted 27.08.21; peer-reviewed by E Olsen, E Rahmani; comments to author 05.03.22; revised version received 19.03.22; accepted 07.04.22; published 19.05.22.

Please cite as:
Franco Vega I, Eleftheriou A, Graham C
Using Video Games to Improve the Sexual Health of Young People Aged 15 to 25 Years: Rapid Review
JMIR Serious Games 2022;10(2):e33207
URL: https://games.jmir.org/2022/2/e33207
doi:10.2196/33207
PMID:35587873

Review

Innovative Technology–Based Interventions to Reduce Stigma Toward People With Mental Illness: Systematic Review and Meta-analysis

Matías E Rodríguez-Rivas[1,2], BSN, RN; Adolfo J Cangas[2], PhD; Laura A Cariola[3], PhD; Jorge J Varela[1], PhD; Sara Valdebenito[4], PhD

[1]Facultad de Psicología, Universidad del Desarrollo, Santiago, Chile
[2]Department of Psychology, Health Research Centre, University of Almería, Almería, Spain
[3]Department of Clinical and Health Psychology, University of Edinburgh, Edinburgh, United Kingdom
[4]Violence Research Centre, Institute of Criminology, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:
Matías E Rodríguez-Rivas, BSN, RN
Facultad de Psicología
Universidad del Desarrollo
Avenida Plaza 680
Santiago
Chile
Phone: 56 223279110
Email: matrodriguezr@udd.cl

Abstract

Background: Stigma toward people with mental illness presents serious consequences for the impacted individuals, such as social exclusion and increased difficulties in the recovery process. Recently, several interventions have been developed to mitigate public stigma, based on the use of innovative technologies, such as virtual reality and video games.

Objective: This review aims to systematically review, synthesize, measure, and critically discuss experimental studies that measure the effect of technological interventions on stigmatization levels.

Methods: This systematic review and meta-analysis was based on PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines and included studies in English and Spanish published between 2016 and 2021. Searches were run in 5 different databases (ie, PubMed, PsycInfo, Scopus, Cochrane Library, and ScienceDirect). Only randomized controlled trials were included. Two independent reviewers determined the eligibility, extracted data, and rated methodological quality of the studies. Meta-analyses were performed using the Comprehensive Meta-Analysis software.

Results: Based on the 1158 articles screened, 72 articles were evaluated as full text, of which 9 were included in the qualitative and quantitative syntheses. A diversity of interventions was observed, including video games, audiovisual simulation of hallucinations, virtual reality, and electronic contact with mental health services users. The meta-analysis (n=1832 participants) demonstrated that these interventions had a consistent medium effect on reducing the level of public stigma (d=–0.64; 95% CI 0.31-0.96; P<.001).

Conclusions: Innovative interventions involving the use of technologies are an effective tool in stigma reduction, therefore new challenges are proposed and discussed for the demonstration of their adaptability to different contexts and countries, thus leading to their massification.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42021261935; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021261935

(JMIR Serious Games 2022;10(2):e35099) doi:10.2196/35099

KEYWORDS

stigma; mental illness; technology-based; serious games; virtual reality; e-contact; simulation intervention; internet intervention; meta-analysis

Introduction

Stigma Toward People With Mental Illnesses

Globally, stigma toward people with mental illness represents a serious public health problem and is considered the main barrier to social inclusion and participation of those impacted. It has a negative effect on their quality of life, worse therapeutic results, and even an increased risk of suicide and mortality [1-3].

Thus, discrimination, prejudice, and stereotypes present in society not only severely impact the recovery process, quality of life, and well-being of people with mental illnesses and their families, but also represent the main gap in accessing specialized mental health services by the general population [4,5]. The aforesaid has great relevance in the global context, where research has shown the presence of high levels of discrimination, stigma, and prejudice toward those impacted by mental health problems, especially schizophrenia and bipolar disorder [6,7].

Innovative Interventions Carried Out at International and National Levels

Faced with the high levels of stigmatization present in society, several initiatives and studies have been conducted that focused on its reduction. It has been demonstrated that direct contact with people with mental illnesses and educational initiatives are essential and effective interventions to reduce stigma [8,9]. Although direct face-to-face contact with people with mental illness has been shown to be a key component of successful stigma reduction programs, their implementation in virtual learning and innovative spaces is recent [10]. In the last years, several authors have shown that innovative technology-based interventions have had a great impact on the reduction of stigma toward people with mental illness, mainly due to its adaptability to different contexts and age ranges [11]. Among this type of interventions, the use of video games has been highlighted as an effective tool to reduce anxious and depressive symptomatology in patients, and has been useful to reduce misconceptions and stigmatization about severe mental illnesses, such as schizophrenia or bipolar disorder [12,13].

In addition, because of the growth of technologies, the application of virtual and immersive reality in mental health has become increasingly common [14]. For example, it has demonstrated its utility in the treatment of mental health problems, such as phobias and anxiety symptoms, among others [15], along with a reduction of the negative perceptions and attitudes toward people with mental illnesses [16]. This usefulness and effectiveness can be explained by the degree of immersion in a strongly educational environment, which promotes the change of attitudes and beliefs [17]. By contrast, the use of simulation platforms has showed controversial results, with some studies showing an increase of stigmatization when used through the simulation of hallucinatory symptomatology, which can promote negative beliefs and attitudes toward people with mental illness [18,19].

Finally, as an innovative intervention, electronic contact (e-contact)—defined as "computer-mediated real-time interactions where members of different groups interact online" through virtual media with mental health services users [20]—has been used as a strategy for promoting awareness and reducing prejudice among ideologically different groups [21]. However, e-contact implementation in the field of stigma reduction is new and innovative, and it has been demonstrated that e-contact using chat and synchronous videoconferencing can reduce anxiety, stigma, and stereotypes toward the population impacted by mental disorders, and promote an inclusive attitude [11,22].

Despite the considerable progress in this field, further research is needed on innovative technologies and their application in mental health care, such as advances in detection, treatment, and promotion of inclusion and well-being of people with mental health problems [23,24]. Regarding this growing field of application of new technologies and the need to synthesize, measure, and critically discuss the effects of the studies performed for the reduction of stigma, the objectives of this systematic review and meta-analysis are to assess the effectiveness of technology-based interventions to reduce stigma associated with people with mental health problems and to describe the experimental studies that use these types of interventions.

Methods

Data Sources and Search Strategy

The systematic review and meta-analysis protocol were registered in the Prospective Register of Systematic Reviews (PROSPERO) international database (registration ID: CRD42021261935) and was conducted according to the guidelines and recommendations of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) [25].

From January to July 2021, searches were conducted from the following 5 databases, including texts dated from March 5, 2016, to March 5, 2021: PubMed, PsycInfo, Scopus, Cochrane Library, and ScienceDirect, using the following string of search terms: [technology or technologies or simulation or virtual or digital or Internet or web or games or computer or app or online or electronic or social media] AND [stigma or discrimination or prejudice or negative attitude or stereotypes] AND [mental health or mental disorder or mental illness or schizophrenia or psychosis or bipolar disorder or depression].

In addition, a final manual trace back literature search was conducted in August 2021 to identify any recently published sources/literature.

Inclusion and Exclusion Criteria

Published articles that met the following criteria were included: (1) randomized controlled trials using innovative technologies (defined as software apps used with smartphone, videogames, e-contact, or virtual reality); (2) interventions aimed at reducing stigma toward people with mental illness (eg, schizophrenia, psychosis, or bipolar disorder), which included at least one relevant quantitative measure of public stigma (eg, attitudes, stereotypes, and social distance); (3) interventions relevant for any populations (eg, students and general population); (4) all age groups; and (5) articles written in English and Spanish, published in peer-reviewed journals.

Exclusion criteria were (1) reduction of stigma not related to mental health problems; (2) technology using only video or education (eg, films or presentations) not combined with any other innovative technology (eg, virtual reality, videogames, or e-contact); (3) reduction focused on self-stigma only; (4) interventions focused on stigma toward psychiatry or addictions; (5) research protocols; (6) measurement of stigma; and (7) studies that did not include a randomized control group as a comparison.

Study Selection and Data Extraction

After excluding duplicates using Endnote reference manager software, 2 researchers (MER-R and AJC) independently selected articles for inclusion. A third investigator (LAC) examined all the included articles to review this selection and resolve discrepancies. In addition, to check and ensure consistency and clarity at the screening and coding stages across studies, we calculated the interrater reliability using the Cohen κ coefficient [26]. We obtained a Cohen κ of 0.75 (SE 0.11), demonstrating a medium to high degree of agreement among coders. Following this calculation, coders (MER-R and AJC) reviewed the articles in which they found disagreements, and any discrepancies were resolved by a consensus discussion with a third investigator (LAC) who was not involved in the searches.

The eligibility of search results was examined in 2 stages: first by title and abstract, and then by full text. Reasons for exclusion were recorded for each document excluded.

Data were extracted using a standardized table format, which was then reviewed by a third author (LAC). Characteristics of each article included the study design, sample size, setting, type of new technology applied, control or comparison intervention, as well as the main outcomes and effect sizes of the interventions. In the cases where the data for the calculation of the effect size were incomplete, we contacted the principal investigator to request the necessary additional information.

Quality and Risk-of-Bias Assessment

To assess the quality of the selected articles, 2 researchers (MER-R and AJC) independently assessed the risk of bias (ROB) of each study using the Cochrane ROB-2 assessment tool, which assesses 7 study design quality criteria (ie, random sequence generation, allocation concealment, blinding of participant and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other bias), graded as high, medium, or low risk [27]. Discrepancies were resolved by further discussions and consensus among the authors. Figure 2 summarizes the assessment of the ROB, which was performed using the robvis visualization tool.

The use of the funnel plot for the evaluation of publication bias was not incorporated, because it has been demonstrated that its use is not reliable when the number of studies pooled in direct comparison is less than 10 [27].

Statistical Analysis

The meta-analysis was conducted by researchers (SV and MER-R) using the Comprehensive Meta-analysis software (version 2) [28]. Standardized mean difference and the inverse variance method with a 95% CI were used for continuous and normally distributed data, respectively. The I^2 and Q-statistic were used to explore heterogeneity of effect sizes [29]. Random effects models were used due to the heterogeneity in the type of intervention in the studies [30].

As 3 of the studies included more than 1 scale that assessed levels of stigma [11,22,31], we conducted an analysis that combined them into a single effect size. In those cases, we followed the methodology suggestion for complex data structures [30]. As a consequence, we computed a summary effect including the multiple measures; this synthetic effect size was then included in the meta-analysis.

In addition, as 2 of the studies [32,33] reported their results through standardized regression, the β coefficients were entered into the comprehensive meta-analysis software (CMA) as correlation coefficient, according to the recommendations of Peterson and Brown [34].

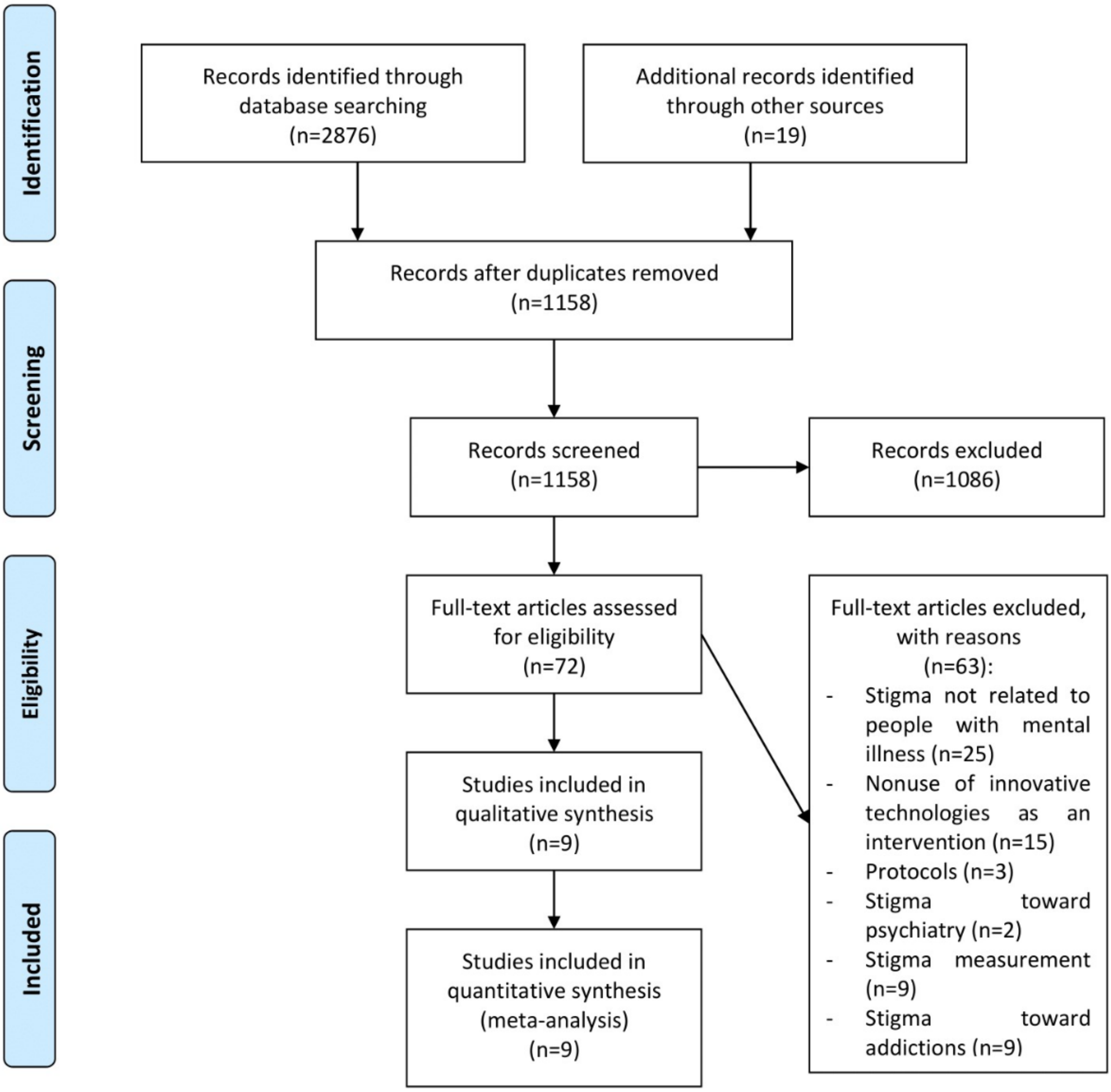
In addition, we used statistical procedures to quantify the effect of publication bias, by Duval and Tweedie's trim-and-fill analysis and Rosenthal's fail-safe N test [35].

Results

Output of Searches

Of the 2876 studies retrieved from the selected databases, 1718 duplicates were removed and 1158 were screened. Upon the screening of the titles and abstracts, 1086 studies were removed. For the remaining 72 studies, their full texts were checked, among which an additional 63 were removed due to specific reasons (Figure 1). Only 9 articles presented enough statistical data for meta-analysis.

Figure 1. Flowchart of the systematic review process.



Characteristic of Included Studies

The characteristics of the included studies are presented in Table 1. A total of 9 randomized control trial studies were included, which utilized a variety of technology-based interventions to reduce stigma, including interventions using video games (n=4), audiovisual simulation of hallucinations (n=1), virtual reality (n=2), and the use of e-contact with mental health services users through videoconferencing and online chats (n=2). Most of these studies were conducted with undergraduate students (8/9, 89%). As many as 4 studies (44%) were conducted in Europe, 2 (22%) in North America, and only 1 each in Asia (11%), Australia (11%), and Latin America (11%). Most of the participants were female. The proportion of female participants ranged from 50.3% to 81.1%, and the mean age of the participants ranged from 15.7 to 24 years. Public stigma in the studies was measured through different scales, with the most commonly used being the Attribution Questionnaire and the Questionnaire on Student Attitudes Toward Schizophrenia.

Table 1. Characteristics of selected studies (n=9).

Experimental intervention	Control group	Sample size	Population	Mean age (SD)	Female, %	Public stigma scale	Outcome public stigma scale (EG[a] and CG[b])	Study
e-Contact with a person diagnosed with schizophrenia through an online chat.	Group without intervention.	133	UGS[c]	18.8 (1.5)	64.7	• AQ-ER[d]	• EG: Decreasing stigma in the factors Fear (*d*=–0.87; *P*<.001) and Anger (*d*=–0.65; *P*<.001) compared with the control condition. Pity did not have a significant difference compared with the CG (*d*=–0.25; *P*=.229).	[22]
						• SS-8[e]	• EG: Decreasing stigma in the factor Stereotypes (*d*=–0.70; *P*<.001) compared with the CG.	
						• PDD[f]	• EG: No statistically significant difference in the social distance among EG and CG (*d*=–0.34; *P*>.05).	
e-Contact with a person diagnosed with schizophrenia by videoconferencing.	Cardiovascular health-related activity.	40	UGS	20.6 (1.3)	80.0	• AQ-E[g]	• EG: Decreasing total stigma level (*d*=–2.33; *P*<.001), and the factors Dangerous-Fear (*d*=–1.73; *P*<.001), Avoidance (*d*=–2.32; *P*<.001), Coercion (*d*=–1.77; *P*<.001), and Lack of Solidarity (*d*=–0.83; *P*=.002). No statistically significant difference in the CG (*P*=.38).	[11]
						• QSAS[h]	• EG: Decreasing total stigma level (*d*=–1.11; *P*=.004) and the factors Dangerousness (*d*=–1.02; *P*=.007) and Stereotypes (*d*=–.83; *P*=.02). No statistically significant differences among the CG (*P*=.66).	
Audiovisual simulation of hallucination symptoms.	Group without intervention.	244	UGS	18.62 (1.0)	62.7	• PDD	• EG: No immediate significant change and 1 week later was documented for any of 2 stigma factors evaluated (*P*>.05). No significant differences in the pre-post intervention and 1 week later among the CG (*P*>.05).	[31]
A serious videogame called Stigma-Stop.	A video game unrelated to mental health.	552	HSS[i]	15.78 (2.65)	50.0	• QSAS	• EG: Decreasing total stigma level (*d*=–0.39; *P*<.001), and the factors Dangerousness (*d*=–0.66; *P*<.001) and Stereotypes (*d*=–1.36; *P*=.001). No statistically significant differences among the CG (*P*=.44).	[36]
A serious videogame called Stigma-Stop.	A video game unrelated to mental health.	530	UGS and HSS	18.51 (4.34)	61.5	• QSAS	• EG: University students had decreasing total stigma level (*d*=–0.37; *P*<.001) and the factor Social Distance (*d*=–0.65; *P*<.001), but not the factor Stereotypes (*P*=.64).	[37]
							• EG: High-school students had reductions in total stigma level (*d*=–0.50; *P*<.001), and the factors Social Distance (*d*=–0.72; *P*<.001) and Stereotypes (*d*=–0.22; *P*<.001). No statistically significant difference among the CG (*P*=.95).	

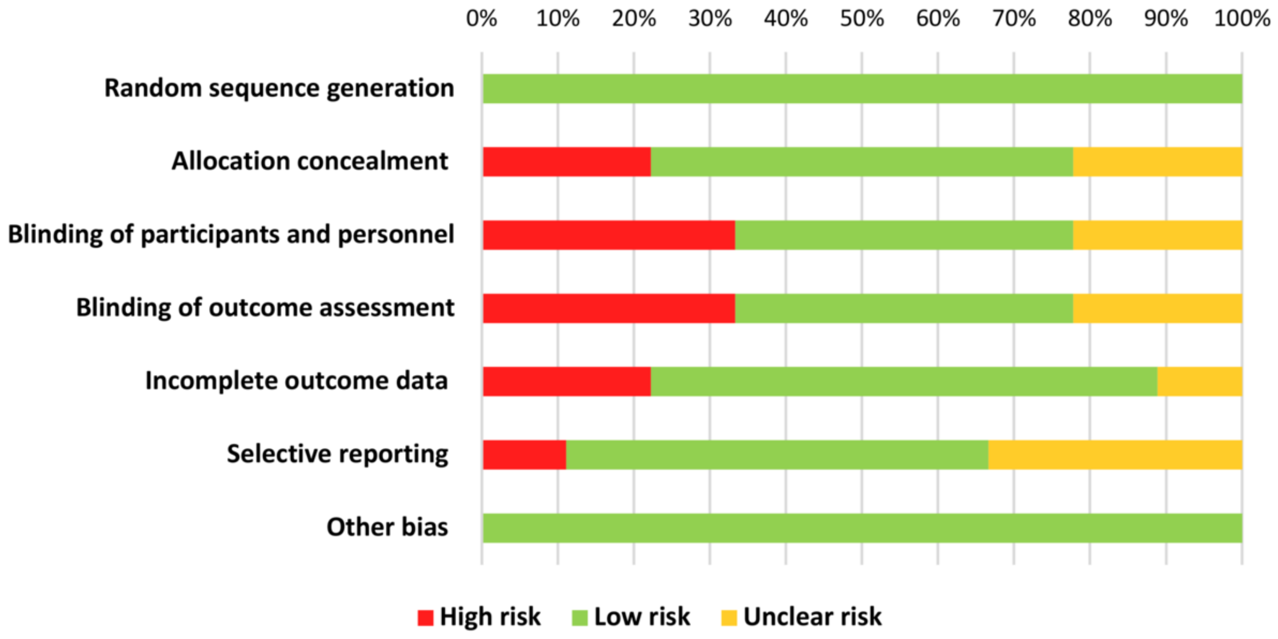
Experimental intervention	Control group	Sample size	Population	Mean age (SD)	Female, %	Public stigma scale	Outcome public stigma scale (EG[a] and CG[b])	Study
A videogame that uses avatars with mental illnesses.	Group watched gameplay footage.	207	UGS	20.42 (not reported)	66.7	• SDS-6[j]	• EG: Structural equation model analysis for both measures show a decrease in Social Distance (B=–0.21; *P*<.05) in the participants that played the videogame compared with the CG.	[32]
						• MIS-9[k]	• EG: The structural model did not show a significance difference in the reduction of Stereotypes (B=–.09; *P*>.05) in participants that played the videogames compared with the CG.	
A serious videogame called Stigma-Stop.	Group attendance at routine class.	118	UGS	21.17 (5.8)	81.1	• AQ-E	• EG: Videogame intervention had a significant effect on decreasing the level of Anger (*d*=–0.95; *P*<.001), Dangerousness (*d*=–1.01; *P*<.05), Fear (*d*=–0.94; *P*<.001), Segregation (*d*=–0.87; *P*<.05), Coercion (*d*=–0.39; *P*<.05), and Avoidance (*d*=–1.03; *P*<.05), and increased the level of Help (*d*=–0.54; *P*<.05). The CG had no statistically significant differences for any of the factors (*P*>.05), except for Anger (*d*=–0.11; *P*<.05).	[13]
Virtual reality implemented by video recording.	Group without intervention.	114	UGS	24.0 (6.6)	58	• SPS-6[l]	• Structural equation modeling analyses that included all participants who positively evaluated the protagonist of the video game showed significantly increased social proximity (B=0.41; *P*=.002).	[33]
Virtual reality by immersive animated story.	Virtual reality exoplanet video.	206	UGS	21.76 (5.04)	55.3	• PSA-21[m]	• EG: Decreased total stigma for the virtual reality immersive intervention (*d*=–0.44; *P*=.003) and even for 1-week follow-up (*d*=–0.32; *P*=.02) compared with the CG. In the mediation model in the virtual reality immersive intervention a decrease in the level of stigma was reported (B=–0.42; *P*<.001).	[16]

[a]EG: experimental group.
[b]CG: control group.
[c]UGS: undergraduate student.
[d]AQ-ER: Attribution Questionnaire, Emotional Response factors.
[e]SS-8: 8-item Stigmatization Scale.
[f]PDD: Perceived Devaluation and Discrimination Scale.
[g]AQ-E: Attribution Questionnaire, Spanish version.
[h]QSAS: Questionnaire on Student Attitudes toward Schizophrenia.
[i]HSS: high-school student.
[j]SDS-6: 6-item Social Distance Scale.
[k]MIS-9: 9-item Mental Illness Stereotypes.
[l]SPS-6: Social Proximity to persons with schizophrenia Scale.
[m]PSA-21: 21-item Public Stigma and Acceptance Scale.

Risk of Bias of the Included Studies

As shown in Figure 2, most of the studies included in this review were considered as having low ROB in terms of their methodological quality. However, 3 studies [31,32,37] showed a high ROB in the blinding of the participants and blinding of outcome assessment. Besides, 2 other studies [11,32] showed a high ROB related to the allocation concealment. Regarding data, 2 studies [22,37] presented incomplete data. Finally, only 1 study [31] had a high ROB in selective reporting, and no study showed a high risk of other bias or random sequence generation.

Figure 2. Risk of bias graph: review of authors' judgments about each risk-of-bias item presented as percentages across all included studies.



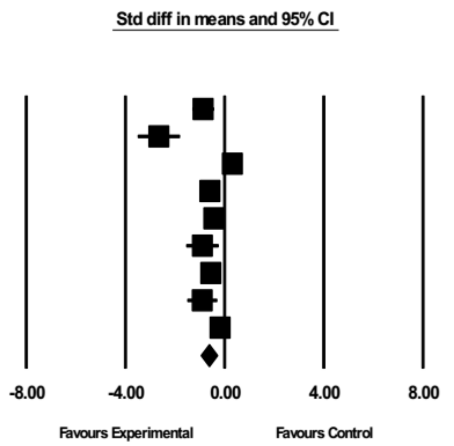
Study and Quantitative Synthesis Outcomes: Public Stigma

A total of 9 articles were included in the meta-analysis, with a total sample of 1832 participants. As shown in Figure 3, the technology-based interventions had medium effects on reducing the level of public stigma (*d*=–0.64; 95% CI 0.31–0.96; *P*<.001) compared with the control group. Only 1 study [31] that used audiovisual simulation symptoms showed an increase in the level of stigma (*P*=.036; *d*=0.32), and another that used videogame with avatars [32] did not show any statistical effect in the level of stigma (*P*=.14; *d*=–0.18). High heterogeneity was observed among the included studies (*P*<.001; I^2=87.6%; *Q*=64.96), which was expected due to the variety of interventions.

Figure 3. Forest plot comparison of the effect on public stigma (standardized means difference) for the innovation-based intervention and control groups.

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Maunder et al (2019) [22]	-0.865	0.221	0.049	-1.297	-0.432	-3.920	0.000
Rodríguez-Rivas et al (2021) [11]	-2.650	0.433	0.188	-3.499	-1.800	-6.115	0.000
Brown (2020) [31]	0.316	0.151	0.023	0.020	0.612	2.092	0.036
Cangas et al (2017) [36]	-0.593	0.131	0.017	-0.849	-0.337	-4.535	0.000
Cangas et al (2019) [37]	-0.432	0.105	0.011	-0.638	-0.226	-4.106	0.000
Mullor et al (2019) [13]	-0.896	0.332	0.110	-1.546	-0.245	-2.700	0.007
Yuen et al (2019) [16]	-0.556	0.172	0.029	-0.892	-0.219	-3.236	0.001
Stelzmann et al (2021) [33]	-0.899	0.307	0.094	-1.501	-0.297	-2.928	0.003
Ferchaud et al (2020) [32]	-0.181	0.123	0.015	-0.422	0.060	-1.471	0.141
	-0.636	0.168	0.028	-0.964	-0.307	-3.791	0.000

Statistics for each study



Publication Bias Analysis

We ran statistical analyses for publication bias [35], where Duval and Tweedie's trim-and-fill analysis identified the differences in effect sizes that could potentially be attributed to bias; the technique imputes effect sizes until the error distribution gets close to normality. In this way, the test offers the best estimate of the unbiased effect [30]. Results suggest that there were no differences in effect sizes attributable to bias.

Under a random effect model, the point estimate for the combined studies did not differ when comparing the original and the adjusted estimate (in both cases, standardized mean difference 0.63; 95% CI 0.94-0.30). Based on the parameter of Duval and Tweedie's trim-and-fill analysis, it seems that no studies are missing.

Additionally, Rosenthal's fail-safe N test is a technique for computing the number of missing studies that would be necessary to nullify the found effect [38]. Small numbers of missing studies would reveal the likelihood of biased effects. Test was equal to 180, suggesting that it would be necessary to allocate and include 180 missing studies with no effects for every observed study to achieve the combined 2-tailed *P* value exceeding .05. Therefore, it is highly unlikely that missing studies could alter the substantive conclusion.

Discussion

Principal Findings

The results of this meta-analysis support the use of new and innovative technology–based interventions to reduce stigma toward people with mental illness.

This study shows a medium effect on stigma reduction, demonstrating a positive impact and effectiveness of these interventions. Its findings are better compared with a previous meta-analysis [39], which reported only a small effect of contact interventions for people with mental illness, and for the educative intervention, both classical and common-type interventions. This is probably due to the increasing technological development, where realism, immersion, and technological interaction are greater, becoming an increasingly natural means of communication and daily application [40].

In this regard, a recent meta-analysis showed that antistigma interventions involving contact had an immediate small-to-medium effect, and it was equivalent with diverse types of contact mediums, such as videos and presentation [41].

Another finding of our study is that the intervention that used simulation of symptoms of hallucination [31] had an increase in the level of stigma, which is concordant with a previous meta-analysis that demonstrated that it can increase social distance and negative attitudes [42]. These negative results may be explained by the focus on symptoms rather than on the recovery process, which may increase stereotypes and prejudice, especially toward people with schizophrenia. Thus, it has been suggested that it should be used with caution and ideally in combination with educational or contact interventions [43]. In this sense, it is important to consider that several research studies show that the information provided is essential to reduce stigma, where, for example, it has been helpful to refer to biographical aspects (eg, related to difficulties, personal adversities), because it allows people to empathize, understand, and generate a change in their attitudes, knowledge, and stigmatizing behaviors [44]. Therefore, the aim is not only to show “symptoms,” but also to promote an understanding of these experiences and the social consequences for the people who experience them [37].

As an additional result, one of the included studies, which used videogames through avatar identification [32], despite showing no significant effect on stigma in the combined effect analysis, reported a significance effect (B=–0.21; P<.05) for the reduction of social distance, making it a tool that can be considered for future research.

Our study demonstrated the usefulness of innovative interventions in stigma reduction and summarized its latest advances, in accordance with the growing interest and need for the application of new technologies in the field of mental health in the contemporary world. These types of interventions have a variety of advantages and offer innovative solutions to everyday problems, due to their adaptability to different contexts and lower associated cost, along with the possibility of privacy in a protected environment, which allows strategies to be focused and adapted to different key population.

Furthermore, it is important to mention that while stigma is currently one of the most important problems in mental health, there are also a number of socioeconomic, cultural, and structural barriers present in society that limit access to treatment and should be considered as an integral part of reducing discrimination and inequalities in mental health.

Limitations

This study has some limitations. First, we did not include gray literature, which can increase the risk of publication bias [45]. Despite this, our statistical analysis of publication bias seems to show no missing studies, and that it is also very unlikely that a possible missing study could alter our results. Second, the small number of studies and high heterogeneity prevented us from estimating which technology-based intervention was the most effective, a limitation that may be considered in future studies as the number of publications in this emerging area increases, allowing future subgroup analyses of each type of intervention and addressing the high heterogeneity. Third, all the studies found and included focused on the population of high-school and university students, which may limit the generalizability of the results to the general population. In addition, most of the included studies did not include follow-up studies, which prevented the analysis of long-term stigma reduction.

Future Research

As stigma is a complex social phenomenon transversally present in the society [1-6], innovative interventions involving use of technologies can be an effective tool for its reduction. However, one challenge in the implementation of this type of intervention is its adaptability to different contexts and countries, and so future interventions should include cross-cultural comparison. In addition, all experimental studies involving technology-based interventions in stigma reduction focus on the young population, which represents a major challenge for future studies focused on the development, implementation, and evaluation of these types of interventions for different ages, such as adults and the elderly [46,47].

Because of the extensive evidence supporting key strategies in the development of antistigma programs, such as the educative and contact components, future research should consider and adapt them to different innovation-based interventions. Its adaptation had demonstrated a great utility, for example, in e-contact with people with mental illness, as they offer the possibility of wider dissemination and even the possibility of reaching remote areas [11,22]. New lines of work should generate greater access and development of low-cost tools with the use of new technologies that allow their use and integration in workplaces, health systems, and educational communities as a daily support tool in mental health, for example, through the development and use of free apps for smartphones and freeware for computers.

Finally, it is essential to generate integrated technological systems not only to reduce stigma, but also to consider preventive interventions in mental health, together with timely referrals to specialized health services and access to treatment. In this regard, we highlight the recently developed app Help Club [48], which provides the possibility of access to communities of mutual support in mental health in virtual spaces, through the use of virtual reality, demonstrating the potential and growing impact of metaverse as a space for social interaction and an increasingly used tool.

Conclusions

Our meta-analysis showed that innovative interventions involving the use of virtual reality and communication technologies are effective tools for stigma reduction toward people with mental illness and can be an alternative and a complement for the traditional methods on stigma reduction. As this field is growing and emerging, future studies present several challenges in their adaptation and dissemination in different populations and countries.

Acknowledgments

This publication is part of the project UAL2020-SEJ-D1912 financed by FEDER "A way to make Europe," and by the National Agency for Research and Development (ANID)/Scholarship Program/MAGISTER NACIONAL/2020–22200859. We thank the researchers who provided us with additional data on their studies.

Authors' Contributions

MER-R conceived the study design. MER-R and AJC independently performed screening, selection of the included studies, and data extraction. LAC reviewed this information and resolved discrepancies. MER-R and AJC assessed the risk of bias of each study. SV, MER-R, and JJV conducted the meta-analysis. All authors contributed to the writing of the article and approved the final version.

Conflicts of Interest

None declared.

References

1. Corrigan PW, Druss BG, Perlick DA. The Impact of Mental Illness Stigma on Seeking and Participating in Mental Health Care. Psychol Sci Public Interest 2014 Oct;15(2):37-70. [doi: 10.1177/1529100614531398] [Medline: 26171956]
2. Michaels PJ, López M, Rüsch N, Corrigan PW. Constructs and concepts comprising the stigma of mental illness. PSYE 2017 Apr 27;4(2):183. [doi: 10.25115/psye.v4i2.490]
3. Rüsch N, Zlati A, Black G, Thornicroft G. Does the stigma of mental illness contribute to suicidality? Br J Psychiatry 2014 Oct 02;205(4):257-259. [doi: 10.1192/bjp.bp.114.145755] [Medline: 25274313]
4. Thornicroft G. Stigma and discrimination limit access to mental health care. Epidemiol Psichiatr Soc 2008 May 18;17(1):14-19. [doi: 10.1017/s1121189x00002621] [Medline: 18444452]
5. Link BG, Struening EL, Neese-Todd S, Asmussen S, Phelan JC. Stigma as a barrier to recovery: The consequences of stigma for the self-esteem of people with mental illnesses. Psychiatr Serv 2001 Dec;52(12):1621-1626. [doi: 10.1176/appi.ps.52.12.1621] [Medline: 11726753]
6. Rodríguez-Rivas ME. Attitudes towards people with mental illness among medical students: A qualitative research. PSYE 2021 Sep 01;13(2):21-33. [doi: 10.25115/psye.v13i2.3663]
7. Parcesepe AM, Cabassa LJ. Public stigma of mental illness in the United States: a systematic literature review. Adm Policy Ment Health 2013 Sep;40(5):384-399 [FREE Full text] [doi: 10.1007/s10488-012-0430-z] [Medline: 22833051]
8. Patten SB, Remillard A, Phillips L, Modgill G, Szeto AC, Kassam A, et al. Effectiveness of contact-based education for reducing mental illness-related stigma in pharmacy students. BMC Med Educ 2012 Dec 05;12(1):120-129 [FREE Full text] [doi: 10.1186/1472-6920-12-120] [Medline: 23216787]
9. Jorm AF. Effect of Contact-Based Interventions on Stigma and Discrimination: A Critical Examination of the Evidence. Psychiatr Serv 2020 Jul 01;71(7):735-737. [doi: 10.1176/appi.ps.201900587] [Medline: 32188364]
10. Fonseca A, Osma J. Using Information and Communication Technologies (ICT) for Mental Health Prevention and Treatment. Int J Environ Res Public Health 2021 Jan 08;18(2):461 [FREE Full text] [doi: 10.3390/ijerph18020461] [Medline: 33430057]
11. Rodríguez-Rivas ME, Cangas AJ, Fuentes-Olavarría D. Controlled Study of the Impact of a Virtual Program to Reduce Stigma Among University Students Toward People With Mental Disorders. Front Psychiatry 2021 Feb 9;12:632252-632259 [FREE Full text] [doi: 10.3389/fpsyt.2021.632252] [Medline: 33633613]
12. Lau HM, Smit JH, Fleming TM, Riper H. Serious Games for Mental Health: Are They Accessible, Feasible, and Effective? A Systematic Review and Meta-analysis. Front Psychiatry 2016;7:209 [FREE Full text] [doi: 10.3389/fpsyt.2016.00209] [Medline: 28149281]

13. Mullor D, Sayans-Jiménez P, Cangas AJ, Navarro N. Effect of a Serious Game (Stigma-Stop) on Reducing Stigma Among Psychology Students: A Controlled Study. Cyberpsychol Behav Soc Netw 2019 Mar;22(3):205-211. [doi: 10.1089/cyber.2018.0172] [Medline: 30855993]
14. Cieślik B, Mazurek J, Rutkowski S, Kiper P, Turolla A, Szczepańska-Gieracha J. Virtual reality in psychiatric disorders: A systematic review of reviews. Complement Ther Med 2020 Aug;52:102480 [FREE Full text] [doi: 10.1016/j.ctim.2020.102480] [Medline: 32951730]
15. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review. Curr Psychiatry Rep 2017 Jul;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]
16. Yuen ASY, Mak WWS. The Effects of Immersive Virtual Reality in Reducing Public Stigma of Mental Illness in the University Population of Hong Kong: Randomized Controlled Trial. J Med Internet Res 2021 Jul 14;23(7):e23683 [FREE Full text] [doi: 10.2196/23683] [Medline: 34259636]
17. Makransky G, Petersen GB. The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. Educ Psychol Rev 2021 Jan 06;33(3):937-958. [doi: 10.1007/s10648-020-09586-2]
18. Rodríguez-Rivas ME, Rodríguez Ferrer JM, Alonso López ID, Fernández Campoy JM, Molina Alonso L, Manzano León A, et al. Efectos Del Contacto Electrónico Vía Videoconferencia En La Reducción Del Estigma Hacia Las Personas Con Enfermedades Mentales. Interv en Context Clínicos y la Salud Un enfoque práctico desde la Investig 2021:53-62. [doi: 10.2307/j.ctv2gz3w4s.7]
19. Silva RDC, Albuquerque S, Muniz ADV, Filho PPR, Ribeiro S, Pinheiro PR, et al. Reducing the Schizophrenia Stigma: A New Approach Based on Augmented Reality. Comput Intell Neurosci 2017;2017:2721846 [FREE Full text] [doi: 10.1155/2017/2721846] [Medline: 29317860]
20. White FA, Abu-Rayya HM. A dual identity-electronic contact (DIEC) experiment promoting short- and long-term intergroup harmony. Journal of Experimental Social Psychology 2012 May;48(3):597-608. [doi: 10.1016/j.jesp.2012.01.007]
21. White FA, Verrelli S, Maunder RD, Kervinen A. Using Electronic Contact to Reduce Homonegative Attitudes, Emotions, and Behavioral Intentions Among Heterosexual Women and Men: A Contemporary Extension of the Contact Hypothesis. J Sex Res 2019 Jul 18;56(9):1179-1191. [doi: 10.1080/00224499.2018.1491943] [Medline: 30019950]
22. Maunder RD, White FA, Verrelli S. Modern avenues for intergroup contact: Using E-contact and intergroup emotions to reduce stereotyping and social distancing against people with schizophrenia. Group Processes & Intergroup Relations 2018 Sep 19;22(7):947-963. [doi: 10.1177/1368430218794873]
23. 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 [FREE Full text] [doi: 10.2196/25847] [Medline: 33913817]
24. Shek AC, Biondi A, Ballard D, Wykes T, Simblett SK. Technology-based interventions for mental health support after stroke: A systematic review of their acceptability and feasibility. Neuropsychol Rehabil 2021 Apr 13;31(3):432-452. [doi: 10.1080/09602011.2019.1701501] [Medline: 31833819]
25. 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]
26. Belur J, Tompson L, Thornton A, Simon M. Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making. Sociological Methods & Research 2018 Sep 24;50(2):837-865. [doi: 10.1177/0049124118799372]
27. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928-d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
28. Pierce CA. Organizational Research Methods 2007 Jul 23;11(1):188-191. [doi: 10.1177/1094428106296641]
29. Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol 2008 Oct 02;37(5):1158-1160. [doi: 10.1093/ije/dyn204] [Medline: 18832388]
30. Pigott T, Polanin J. Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review. Review of Educational Research 2019 Sep 25;90(1):24-46. [doi: 10.3102/0034654319877153]
31. Brown S. The Effectiveness of Two Potential Mass Media Interventions on Stigma: Video-Recorded Social Contact and Audio/Visual Simulations. Community Ment Health J 2020 Apr 04;56(3):471-477. [doi: 10.1007/s10597-019-00503-8] [Medline: 31686302]
32. Ferchaud A, Seibert J, Sellers N, Escobar Salazar N. Reducing Mental Health Stigma Through Identification With Video Game Avatars With Mental Illness. Front Psychol 2020;11:2240 [FREE Full text] [doi: 10.3389/fpsyg.2020.02240] [Medline: 33013585]
33. Stelzmann D, Toth R, Schieferdecker D. Can Intergroup Contact in Virtual Reality (VR) Reduce Stigmatization Against People with Schizophrenia? J Clin Med 2021 Jun 30;10(13):2961 [FREE Full text] [doi: 10.3390/jcm10132961] [Medline: 34209466]

34. Peterson RA, Brown SP. On the use of beta coefficients in meta-analysis. J Appl Psychol 2005 Jan;90(1):175-181. [doi: 10.1037/0021-9010.90.1.175] [Medline: 15641898]
35. Shi L, Lin L. The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses. Medicine (Baltimore) 2019 Jun;98(23):e15987 [FREE Full text] [doi: 10.1097/MD.0000000000015987] [Medline: 31169736]
36. Cangas AJ, Navarro N, Parra JMA, Ojeda JJ, Cangas D, Piedra JA, et al. Stigma-Stop: A Serious Game against the Stigma toward Mental Health in Educational Settings. Front Psychol 2017 Aug 21;8:1385-1310 [FREE Full text] [doi: 10.3389/fpsyg.2017.01385] [Medline: 28878702]
37. Cangas, Navarro, Aguilar-Parra, Trigueros, Gallego, Zárate R, Gregg. Analysis of the Usefulness of a Serious Game to Raise Awareness about Mental Health Problems in a Sample of High School and University Students: Relationship with Familiarity and Time Spent Playing Video Games. J Clin Med 2019 Sep 20;8(10):1504 [FREE Full text] [doi: 10.3390/jcm8101504] [Medline: 31547017]
38. Fragkos K, Tsagris M, Frangos C. Publication Bias in Meta-Analysis: Confidence Intervals for Rosenthal's Fail-Safe Number. Int Sch Res Notices 2014;2014:825383 [FREE Full text] [doi: 10.1155/2014/825383] [Medline: 27437470]
39. Corrigan PW, Morris SB, Michaels PJ, Rafacz JD, Rüsch N. Challenging the public stigma of mental illness: a meta-analysis of outcome studies. Psychiatr Serv 2012 Oct;63(10):963-973. [doi: 10.1176/appi.ps.201100529] [Medline: 23032675]
40. Snodgrass JG, Bagwell A, Patry JM, Dengah HF, Smarr-Foster C, Van Oostenburg M, et al. The partial truths of compensatory and poor-get-poorer internet use theories: More highly involved videogame players experience greater psychosocial benefits. Computers in Human Behavior 2018 Jan;78:10-25. [doi: 10.1016/j.chb.2017.09.020]
41. Maunder RD, White FA. Intergroup contact and mental health stigma: A comparative effectiveness meta-analysis. Clin Psychol Rev 2019 Aug;72:101749. [doi: 10.1016/j.cpr.2019.101749] [Medline: 31254936]
42. Morgan AJ, Reavley NJ, Ross A, Too LS, Jorm AF. Interventions to reduce stigma towards people with severe mental illness: Systematic review and meta-analysis. J Psychiatr Res 2018 Aug;103:120-133. [doi: 10.1016/j.jpsychires.2018.05.017] [Medline: 29843003]
43. Ando S, Clement S, Barley EA, Thornicroft G. The simulation of hallucinations to reduce the stigma of schizophrenia: a systematic review. Schizophr Res 2011 Dec;133(1-3):8-16. [doi: 10.1016/j.schres.2011.09.011] [Medline: 22005017]
44. Longdon E, Read J. 'People with Problems, Not Patients with Illnesses': Using Psychosocial Frameworks to Reduce the Stigma of Psychosis. Isr J Psychiatry Relat Sci 2017;54(1):24-28 [FREE Full text] [Medline: 28857755]
45. Paez A. Gray literature: An important resource in systematic reviews. J Evid Based Med 2017 Aug;10(3):233-240. [doi: 10.1111/jebm.12266] [Medline: 28857505]
46. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. “These devices have not been made for older people's needs” – Older adults' perceptions of digital technologies in Finland and Ireland. Technology in Society 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
47. Vaportzis E, Clausen MG, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. Front Psychol 2017 Oct 04;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
48. Help Club app: Mental Health Help for the Metaverse. Very Real Help. URL: https://veryrealhelp.com [accessed 2022-03-18]

Abbreviations

AQ-E: Attribution Questionnaire, Spanish version
AQ-ER: Attribution Questionnaire, Emotional Response factors
CG: control group
e-contact: electronic contact
EG: experimental group
HSS: high-school student
MIS-9: 9-item Mental Illness Stereotypes
PDD: Perceived Devaluation and Discrimination Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses
PROSPERO: Prospective Register of Systematic Reviews
PSA-21: 21-item Public Stigma and Acceptance Scale
QSAS: Questionnaire on Student Attitudes toward Schizophrenia
ROB: risk of bias
SDS-6: 6-item Social Distance Scale
SPS-6: 6-item Social Proximity to persons with schizophrenia Scale
SS-8: 8-item Stigmatization Scale
UGS: undergraduate student

Edited by N Zary; submitted 24.11.21; peer-reviewed by J Wu, M Ahmed Kamal, A Kazdin; comments to author 16.03.22; revised version received 21.03.22; accepted 14.04.22; published 30.05.22.

Please cite as:
Rodríguez-Rivas ME, Cangas AJ, Cariola LA, Varela JJ, Valdebenito S
Innovative Technology–Based Interventions to Reduce Stigma Toward People With Mental Illness: Systematic Review and Meta-analysis
JMIR Serious Games 2022;10(2):e35099
URL: https://games.jmir.org/2022/2/e35099
doi:10.2196/35099
PMID:35635744

Original Paper

A Web-Based Escape Room to Raise Awareness About Severe Mental Illness Among University Students: Randomized Controlled Trial

Jose M Rodriguez-Ferrer[1*], MSc; Ana Manzano-León[1*], PhD; Adolfo J Cangas[1*], PhD; Jose M Aguilar-Parra[1*], PhD

Faculty of Psychology, University of Almeria, Spain

[*]all authors contributed equally

Corresponding Author:
Ana Manzano-León, PhD
Faculty of Psychology
University of Almeria
Carretera Sacramento s/n
La Cañada de San Urbano
Almería, 04120
Spain
Phone: 34 950 01 55 50
Email: aml570@ual.es

Abstract

Background: People with severe mental illness (SMI) face discriminatory situations because of prejudice toward them, even among health care personnel. Escape rooms can be a novel educational strategy for learning about and empathizing with SMI, thus reducing stigma among health care students.

Objective: This study aimed to examine the effect of the Without Memories escape room on nursing students' stigma against SMI.

Methods: A pre- and postintervention study was conducted with a control group and an experimental group. A total of 306 students from 2 Andalusian universities participated in the study. Data were collected through a pre-post study questionnaire, consisting of an adapted version of the Attributional Style Questionnaire and a questionnaire on motivation for cooperative playful learning strategies. The control group carried out an escape room scenario without sensitizing content, whereas the experimental group carried out an escape room scenario on SMI, with both escape rooms being carried out in a 1-hour session of subjects related to mental health. To answer the research questions, a 2-way analysis of variance with repeated measures, a linear regression, and a 2-way analysis of variance were performed.

Results: After the intervention, a significant reduction ($P<.001$) was observed in the experimental group in stigmatizing attitudes compared with the control group, in which no statistically significant changes ($P>.05$) were observed. In contrast, the linear regression ($t_{195}=-22.15$; $P<.001$) showed that there was an inverse relationship between flow and the level of reduced stigma. When controlling for having or not having a close relative, the intervention was also shown to be effective ($P<.001$) in reducing the stigma displayed, both for people with affected and unaffected relatives.

Conclusions: Our findings suggest that the Without Memories escape room can be used as an effective tool to educate and raise awareness about stigmatizing attitudes toward SMI in university students studying health care. Future testing of the effectiveness of educational escape rooms should be designed with new programs through playful strategies of longer duration to evaluate whether they can achieve a greater impact on motivation, acquisition of knowledge, and awareness. In addition, the feasibility of implementing the Without Memories escape room in other careers related to health and community should be investigated.

(JMIR Serious Games 2022;10(2):e34222) doi:10.2196/34222

KEYWORDS

escape room; severe mental disorder; higher education; nursing education; mental health; mental disorder; serious games

Introduction

Background

Severe mental illness (SMI) refers to different nosological entities with certain common severity criteria, lasting more than 2 years, and is associated with a loss of functionality [1]. Formally, it includes diagnoses such as schizophrenia, bipolar disorder, or severe personality disorders. The main difficulty faced by this group is stigma. People who are affected from stigma face discriminatory situations that prevent them from fully participating in society. Stigma is a multifactorial phenomenon that thrives on stereotypes and prejudices toward people with mental disorders. The most common stereotypes toward people with SMI are, for example, that they can be dangerous and commit more crimes than the general population [2]; they can be violent, unpredictable, and dependent; and they are responsible for their illness [2]. Stigma has severe consequences on the quality of life and self-esteem of people with mental disorders, hindering their social inclusion [3], which, in turn, has a negative effect on finding work, being economically independent, and having a solid circle of close friends and supporters [4].

To reduce stigma, different proposals for socioeducational programs have been made, including the dissemination of information, leisure activities such as sports or art, and contact with people with mental disorders. For example, in a meta-analysis study [5], through the analysis of 72 articles, it was observed that educational and social programs had positive effects in reducing stigma in adults and adolescents with mental illness.

With the arrival of serious games and educational video games, the possibilities of electronic resources in socioeducational intervention are also beginning to be studied. For example, in the video game Stigma-Stop, which aimed at reducing stigma toward people with mental disorders, its characters had various mental disorders, with whom participants could empathize. The results of 552 students aged between 14 and 18 years showed a statistically significant decrease in the levels of stigma toward people with schizophrenia [6].

Playful strategies, such as video games, have the potential to be widely motivating for students, making them pivotal to the acquisition of social skills among young people with mental disorders [7] and to the reduction of stigma [8].

In recent years, new forms of games have emerged and established themselves as a leisure alternative, including escape rooms [9]. Escape rooms consist of being locked in a room, solving a series of puzzles, unlocking locks, and finding hidden clues to escape from that room. Using a variety of settings and challenges, escape rooms create an experience that manages to be both motivating and educational for participants [10].

The results of recent studies [11] show that games and the use of escape rooms have been effective in engaging students in the learning process and helping them retain information. In higher education, specifically in the health care field, games have been used to evaluate the theoretical and practical knowledge of nursing students and study their motivation, where it can be observed that the students perceive that it helps them to learn, consider that these types of activities are fun, and believe that they should be included in their training curricula [12]. Owing to the situation caused by the COVID-19 pandemic, the daily face-to-face use of escape rooms has had to be transformed to apply it during web-based teaching. Previous studies [13-15] show that playful strategies have broad benefits in motivation and academic performance in massive web-based open courses.

Objectives

University students in health care and social work studies are especially concerned that they are empathic with the population affected by mental disorders because they are one of the vulnerable groups in their professional future that they will have to attend to. The objective of this study was to evaluate the impact of a web-based escape room on raising awareness of SMI among university students in the health care branch during the 2020-2021 academic year in web-based mode. The specific objectives of this study were (1) to determine whether through a web-based escape room, it is possible to modify the stigmatizing attitudes of university students taking a nursing degree regarding SMI; (2) to study the influence of the immersion (flow) experienced during the escape room experience and on the awareness of the participants toward stigma; and (3) to determine whether the degree of stigmatizing attitudes is different between students with family members with severe mental disorders and students without family members with SMI and to analyze whether the program has the same effects on these people.

Methods

Participants

All the participants were first-semester nursing students. The participants were chosen based on the purpose of the research, consideration of the resources available when performing the intervention with web-based escape rooms, and the willingness of the university teachers to carry out the intervention in their class schedule.

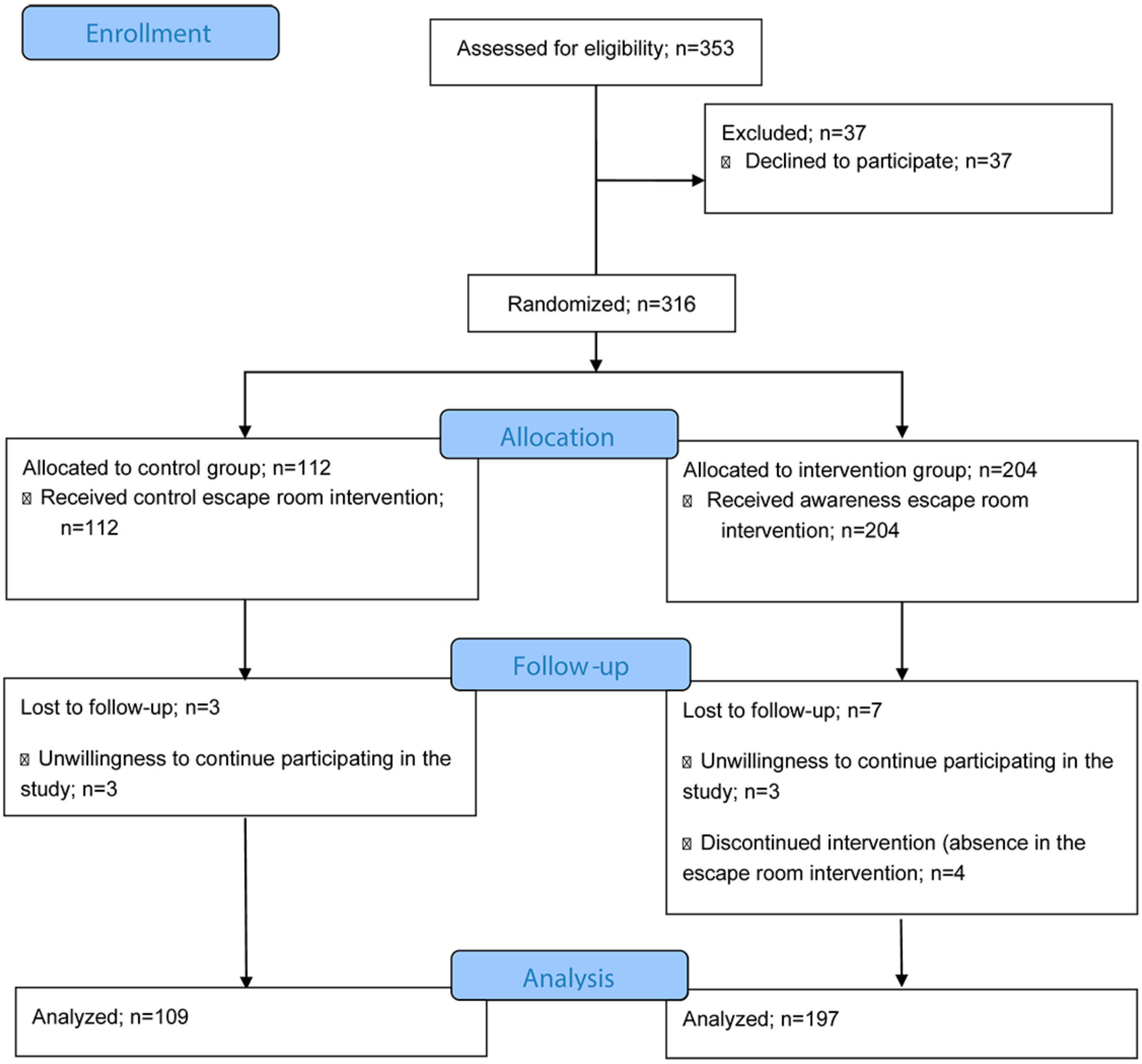
Both the escape room of the control group and the escape room of the experimental group were used during class time in subjects that had some content related to mental health. Students enrolled in the first semester of their nursing degree could participate in the research if they gave their written consent. The escape room did not score on the average of the course, and no financial reward or any other type of incentive was offered (Figure 1; Multimedia Appendix 1).

A total of 306 nursing students participated in the study, with a mean age of 23.34 (SD 8.37) years, comprising 57 (18.6%) men and 249 (81.4%) women. They were selected through cluster sampling, where a particular batch of students were selected as participants, dividing them by participating classrooms. They were randomly allocated between the control group and the experimental group. The experimental group consisted of 197 students with a mean age of 22.93 (SD 7.93) years, comprising 37 (18.8%) men and 160 (81.2%) women; the control group consisted of 109 students with a mean age of

24.1 (SD 9.10) years, comprising 20 (18.3%) men and 89 (81.7%) women.

To fulfill the third objective of this study—to determine whether the degree of stigmatizing attitudes is different between students with family members with severe mental disorders and students without family members with SMI and to analyze whether the program has the same effects on these people—demographic questions were asked of the students before beginning the intervention about whether they had relatives who had SMI. The data showed that 67 students had relatives with SMI, with a mean age of 22.78 (SD 4.27) years. Of these 67 students, 23 (34.3%) belonged to the control group, 19 (82.6%) women and 4 (17.4%) men, with a mean age of 23 (SD 4.62) years, and 44 (65.7%) belonged to the experimental group, 38 (86.4%) women and 6 (13.6%) men, with a mean age of 22.7 (SD 4.07) years.

Figure 1. Flowchart followed for the selection of participants.



Ethics Approval

This study was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent. Ethics approval was obtained from the Research Ethics Committee of the University of Almería (Ref. UALBIO 2021/01).

Instrument

Attributional Questionnaire

The 14-item abbreviated version of public stigma in mental health (Spanish) was used, adapted from the Attributional Questionnaire (AQ) 27 [16], with 14-point Likert-type questions from 1 to 9. It has a Cronbach α of .87. A higher score on this questionnaire indicates a greater number of stigmatizing attitudes toward people with SMI. It measures the following four factors:

1. Dangerousness—measures whether people with SMI pose a threat or create feelings of fear. An example question is, "Do you feel that people with SMI are dangerous?"
2. Solidarity—measures the willingness to help a person with SMI; for example, "If I owned an apartment, I would rent it to people with SMI."
3. Coercion—people with SMI are required to participate in treatment. An example of an item for this factor is "I think it is in the best interest of the community where a person with SMI lives to be placed in a psychiatric facility."

4. Avoidance—indicates the willingness to live or work near someone with SMI. For example, “I will share my car every day with a person with SMI.”

Motivation Questionnaire for Cooperative Playful Learning Strategies

The Motivation Questionnaire for Cooperative Playful Learning Strategies [17] was used to evaluate the learning process, degree of immersion, and motivation toward gamified activities. It is a Likert scale, ranging from 1 to 7, with 1 being totally disagree and 7 being totally agree. The flow factor has been used in this questionnaire, with a Cronbach α of .83. An example of the items contained in the factor is, “While playing, I was not aware of what was going on around me.”

For this study, an escape room has been designed based on SMI, namely, the *Without Memories* escape room. The narrative is that the main character must remember his or her identity and leave the apartment where he or she is to be on time for an appointment. In the escape room, the player wakes up with no memory of who or where he or she is and, through solving tests and exploration, discovers that he or she is a young man with a mental disorder. Through different elements, such as newspapers, mobile phones, social networks, and computers, a normalized life is gradually presented together with the different barriers that people with SMI face because of the stigma and discrimination they face; for example, difficulties in health services or difficulty finding work.

The escape room features linear mechanics: one test allows the next to be solved and so on until the final task, which requires the resolution of 2 tasks to obtain the code that ultimately allows the player to exit. The escape room is web-based and is designed within the *Genialy* digital platform, designed as a graphic adventure (Click and Point). The difficulty of the escape room is low because it is designed for a nonexpert audience in these activities, and its duration is approximately 1 hour.

Procedure

To investigate whether web-based escape rooms could modify stigmatizing attitudes toward SMI, 2 web-based escape rooms were designed and evaluated according to the intervention group (Table 1).

Owing to sanitary measures arising from the COVID-19 situation, both escape rooms were designed on the web. The escape rooms were prepared using the *Genially* digital platform, allowing the creation of *Click and Point* spaces, which is why it favors creating an interactive experience that can favor student immersion. To create the scenarios proposed in the escape room, images and illustrations from the FreePik resource bank were used, owing to a premium license for free use (Figures 2 and 3).

The main advantage of placing the escape room on the web is the ease of implementation in different universities through a link. The steps to carry out both escape rooms were as follows:

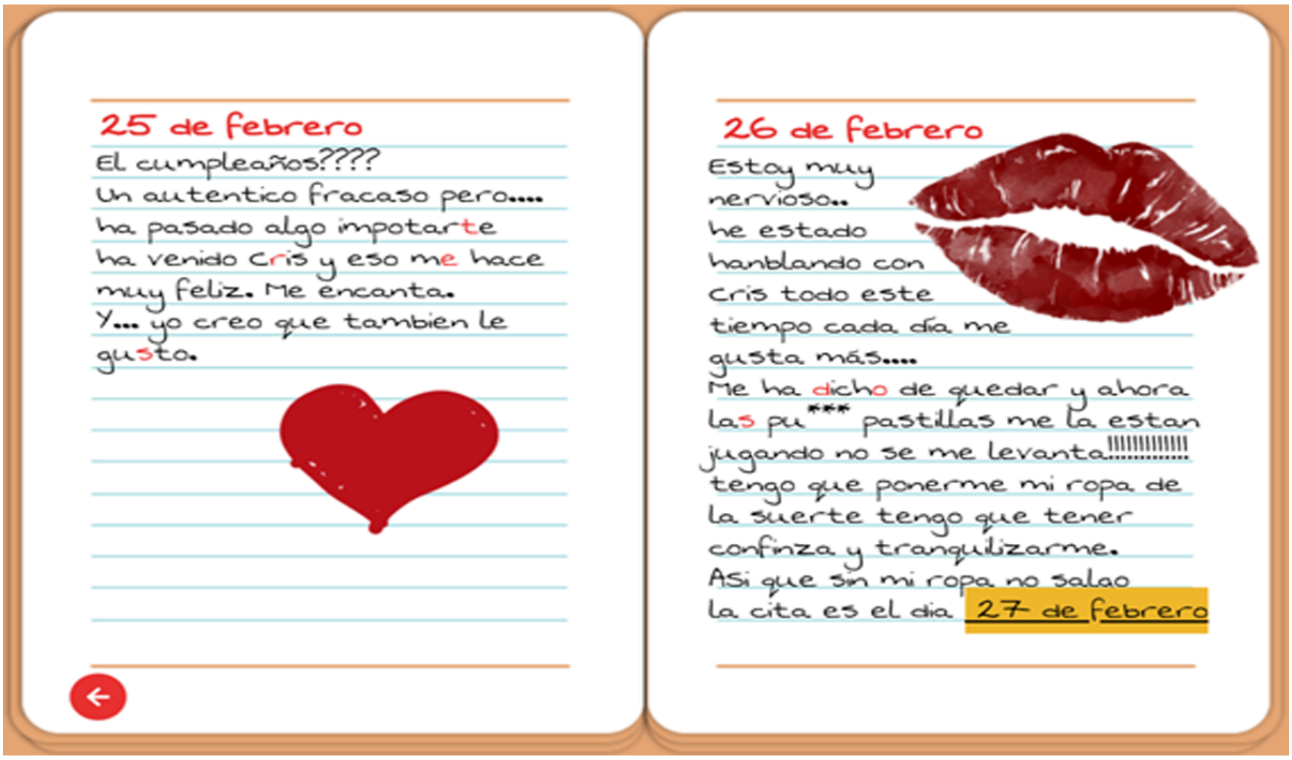
1. On the official platform of the university (Blackboard web-based learning platform), the class began, and the game master explained in a general way what an escape room was, how the Genially platform was used, and that this escape room was cooperative. Once any questions among the students had been resolved, the escape rooms were distributed to the students in teams of 4 people.
2. Each team was housed in a web-based workroom so that they could share ideas and opinions about solving the escape room. They had 2 options: either one of them shared a screen and they solved it together through that screen or each person solved the escape room simultaneously on their computer, maintaining the relationship between their partners in both options.
3. The game master was in the general room of the Blackboard platform during the entire escape room implementation time, in case a team contacted to request clues. There was no limit to the number of clues, as it was important that they felt they could receive feedback and were able to get out of the escape room even if they had never played before.
4. Once the team passed the last test and managed to leave the escape room, they completed the posttest questionnaire individually.

Table 1. Investigation procedure.

	Pretest	Intervention	Posttest
Control group	Attributional Questionnaire [14] and Motivation Questionnaire for Cooperative Playful Learning Strategies [16]	Escape room *Locked In*	Attributional Questionnaire [14] and Motivation Questionnaire for Cooperative Playful Learning Strategies [16]
Experimental group	Attributional Questionnaire [14] and Motivation Questionnaire for Cooperative Playful Learning Strategies [16]	Escape room *Without Memories*	Attributional Questionnaire [14] and Motivation Questionnaire for Cooperative Playful Learning Strategies [16]

Figure 2. Example of elements where you must search for clues.

Figure 3. Example of awareness elements mixed with clues for other puzzles.



To design the escape room *Without Memories*, interviews were conducted with people with SMI and 2 professionals from a local mental health association with the aim of understanding the difficulties people with SMI deal with daily and the reality they face as told from their own experiences. The game design for an escape room has a linear structure. Solving one clue will provide the object necessary to solve the next clue, and so on and so forth until the students escape. Tasks in a linear room must be solved in a particular order. The experiment was conducted on the web and synchronously. The average duration of the escape room was 60 minutes.

The escape room narrative involves waking up in a room without any memories of who or where you are. You must overcome different challenges until you know that your name is Enrique, you have an SMI, and you have a normal life.

The different tasks of the escape room are presented in the following formats inside the game:

- Computer: in this game, there are elements hidden around the room (click and point) that allow the student to know the password of a computer. On the computer, Enrique's browsing history can then be accessed, where there are 2 videos of awareness campaigns on mental health and mental disorders, and Enrique's social networks can be accessed, where 2 conversations with friends and a conversation with his family physician are observed, in which he mentions difficulties related to medication, such as the fear of not being able to have erections and to relate intimately with the woman he likes. Between these conversations, the key for the next test is observed.
- Diary and mobile phone: more conversations with Enrique's family and friends are reflected on the mobile phone. These conversations reflect Enrique's concerns and limited relationships. Some responses to his messages can be interpreted as condescension, childishness, and other behaviors that people with SMI reported in previous interviews. It is also narrated how these people may have a perception of the world that is different from the norm and may display different concerns and hobbies.
- Safe and door: the safe is opened by means of a phrase that is encoded in the computer. To achieve this, it is necessary to find the decoding key located in one of the rooms. This coded information allows the safe to be opened using information contained in the mobile phone and in the bedroom and allows you to remember that you are Enrique, you are an adult, you have an SMI, and you have met the girl you like for a coffee. The box finally gives you the key to open the door. When you click on the door again, there is a concluding message. This message provides statistical information about SMI and the possibility that the player may have an SMI or meet someone close with an SMI in the future.

For the control group, an escape room called *Locked In* was prepared. This escape room is a copy of the previous one in which all the elements that had been prepared to raise awareness, such as narrative texts and correspondence on social networks, were removed, replacing them with others without any burden of awareness, as they are logical tests, puzzles, and numerical locks. The narrative of this escape room assumes that you wake up in bed with no memories and must leave the apartment you are in. The steps for resolution are the same as those carried out by the experimental group; that is, they follow a linear logic in which one test leads to another until the final resolution is reached.

Data Analysis

Initially, a sample size calculation was performed using the G*Power 3.1 program [18]. The parameters used for the calculation were α=.05 and statistical power (1−β)=0.80. For the effect size reference, the scientific literature was reviewed [6,14,19]. The effect size was moderate (d>0.20). Using these data, the minimum sample size needed for ANOVA testing was 277 students. Therefore, the sample size for this study was considered sufficient.

Before starting the analyses to answer the research questions, the direct scores of the administered tests were calculated using the corresponding descriptive statistics.

To answer the first research question, an ANOVA (2×2) with repeated measures was performed for each factor of the AQ14, as well as another test for the total result of the battery. The Bonferroni correction was used as the adjustment method for the comparison of post hoc tests.

To determine the relationship between flow and the change in stigmatizing attitudes, a linear regression was performed, in which the differential scores were calculated. In other words, pretest scores were subtracted from posttest measurements. With the differential scores and the score obtained in the flow, the linear regression was performed with the score of the flow as an independent variable.

To address the last research objective, a 3-way ANOVA was carried out for each of the factors and the total score of the AQ14.

The analyses were carried out using the R programing language with the R Studio development environment. The libraries used were ggplot2, tidyverse, emmeans, and rstatix.

Results

The means and SDs are presented in Table 2. To better understand these results, it should be clarified that higher scores on the questionnaire indicate a greater stigma expressed by participants.

Regarding the first research question, ANOVAs (2×2) were performed with repeated measurements. The results of the tests between the groups are reported in Table 3, which shows the results of the ANOVAs along with the effect size using generalized eta squared. Post hoc tests are reported in this table, indicating the group in which there are statistically significant differences together with significance. This table shows that all statistical differences were found in favor of the experimental group in the posttest measurements. Table 4 indicates that the starting groups had equivalent scores, and statistically significant differences were found after applying the program. Regarding the effect sizes found in the variables, it can be indicated that in all the variables there is a large effect size (η^2_g>0.14), except in coercion, where the size is medium, and avoidance, where the size is small (η^2_g>0.01).

As for the second research question, as already mentioned, a linear regression was carried out between the flow variables (mean 18.4, SD 3.16), such as the independent variable, and the differential scores (mean −16.7, SD 18.4), such as the dependent variable. The results obtained in the analysis indicated that the predictor variables explained 84.5% of the total variance (r^2=−0.845) and a moderate slope of the curve (β=−0.145; t_{195}=−22.15; P<.001); that is, the higher the score obtained for flow, the more the stigma displayed was reduced.

Table 2. Mean and SD (pre- and posttreatment) of the control and experimental groups.

	Control, mean (SD)		Experimental, mean (SD)	
	Pretest	Posttest	Pretest	Posttest
Dangerousness	17.93 (6.78)	17.91 (6.76)	18.03 (7.57)	11.01 (6.17)
Solidarity	11.14 (5.5)	10.99 (5.39)	10.2 (5.48)	6.04 (3.58)
Coercion	8.63 (4.51)	8.87 (4.5)	8.18 (4.72)	5.65 (3.82)
Avoidance	11.86 (4.51)	11.78 (4.69)	11.16 (4.99)	8.13 (5.05)
Total	49.56 (16.03)	49.55 (16.02)	47.57 (16.7)	30.83 (14.79)

Table 3. ANOVA tests comparing between groups with the variables AQ14.

	Pretest			Posttest			Post hoc
	F test (*df*)	*P* value (adjusted)[a]	η^2_g	*F* test (*df*)	*P* value (adjusted)	η^2_g	
Dangerousness	0.01 (1,304)	.91	0	81.81 (1,304)	<.001	0.212	Exp[b]-Cont[c,d]
Solidarity	2.04 (1,304)	.16	0.007	92.61 (1,304)	<.001	0.234	Exp-Cont[d]
Coercion	0.67 (1,304)	.41	0.002	43.81 (1,304)	<.001	0.126	Exp-Cont[d]
Avoidance	1.52 (1,304)	.22	0	38.60 (1,304)	<.001	0.011	Exp-Cont[d]
Total battery	1.03 (1,304)	.62	0.003	106.16 (1,304)	<.001	0.258	Exp-Cont[d]

[a]Bonferroni adjusted.
[b]Exp: experimental.
[c]Cont: control.
[d]*P*<.001.

Table 4. ANOVA tests comparing within groups with the variables AQ14.

	Control			Experimental			Post hoc
	F test (*df*)	*P* value (adjusted)[a]	η^2_g	*F* test (*df*)	*P* value (adjusted)	η^2_g	
Dangerous-ness	0.01 (1,108)	.90	0	196.13 (1,196)	<.001	0.206	Pre-post[b]
Solidarity	2.12 (1,108)	.15	0	104.81 (1,196)	<.001	0.169	Pre-post[b]
Coercion	3.68 (1,108)	.58	0	55.42 (1,196)	<.001	0.08	Pre-post[b]
Avoidance	0.47 (1,108)	.49	0	64.61 (1,196)	<.001	0.084	Pre-post[b]
Total battery	0.01 (1,108)	.98	0	163.19 (1,196)	<.001	0.221	Pre-post[b]

[a]Bonferroni adjusted.
[b]*P*<.001.

The means and SDs before and after the intervention are displayed in Table 5, differentiating those who reported having a close relative and those who did not, both for the control group and the experimental group.

The results of the ANOVA indicated that having a close relative influences the level of stigma of the participants in all study variables (dangerousness: $F_{1.604}$=6.33, *P*=.01, η^2_g=0.01; solidarity: $F_{1.604}$=30.8, *P*<.001, η^2_g=0.04; coercion: $F_{1.604}$=25, *P*<.001, η^2_g=0.04; avoidance: $F_{1.604}$=49.5, *P*<.001, η^2_g=0.08; and total battery: $F_{1.604}$=42.2, *P*<.001, η^2_g=0.06), with statistically significant differences in favor of people who have a close relative. In other words, people with a close relative with SMI display a lower level of stigma than those without a close relative with SMI.

With the intention of knowing how the program affects the fact of having or not having family members with SMI, the 3-way ANOVA was continued, and the results of this are reported for the posttest measures. The ANOVA results are presented in Table 6. In addition, the post hoc tests are reported, indicating only the statistically significant results.

As can be seen, statistically significant differences were found between the control and experimental groups, especially in the case of not having a family member affected by SMI. In this case, statistically significant differences were observed in all study variables. However, when someone had a family member affected by SMI, no statistically significant changes were found for the variables solidarity, coercion, and avoidance, whereas for the variables dangerousness and total battery, there were statistically significant changes between the people in the control and experimental groups who had family members affected by SMI.

Table 5. Means and SDs of participants with and without relatives with severe mental illness in the control and experimental group.

	Control-pre, mean (SD)		Experimental-pre, mean (SD)		Control-post, mean (SD)		Experimental-post, mean (SD)	
	Family	No family	Family	No family	Family	No family	Family	No family
Dangerousness	15.43 (5.68)	18.59 (6.92)	16.48 (7.59)	18.46 (7.53)	15.95 (6.03)	18.43 (9.92)	11.48 (5.95)	10.87 (6.24)
Solidarity	7.91 (4.36)	12 (5.47)	7.48 (4.52)	10.96 (5.49)	7.82 (4.03)	11.83 (5.40)	5.58 (3.43)	6.16 (3.62)
Coercion	6.39 (4.08)	9.23 (4.45)	5.81 (3.58)	8.83 (4.79)	6.69 (3.94)	9.45 (4.48)	4.60 (2.87)	5.94 (3.99)
Avoidance	8.30 (4.34)	12.81 (4.08)	7.6 (4.65)	12.14 (4.62)	8.34 (4.93)	12.69 (4.19)	6.90 (4.85)	8.46 (5.06)
Total	38.04 (15.23)	52.63 (14.86)	37.39 (16.70)	50.40 (15.61)	38.82 (15.90)	52.41 (14.86)	28.58 (12.56)	31.45 (15.33)

Table 6. Three-way ANOVA with the interaction between the control and experimental group and whether they have declared a close relative or not in posttest measures.

	Family			No family			Post hoc
	F test (*df*)	*P* value (adjusted)[a]	η^2_g	*F* test (*df*)	*P* value (adjusted)	η^2_g	
Dangerousness	6.41 (1,65)	<.001	0.011	67.512 (1,238)	<.001	0.101	Exp[b]-cont[c] (no family)[d]/exp-cont (family)[e]
Solidarity	3.317 (1,65)	.280	0.005	77.812 (1,238)	<.001	0.114	Exp-cont (no family)[d]
Coercion	3.61 (1,65)	.232	0.006	37.40 (1,238)	<.001	0.058	Exp-cont (no family)[e]
Avoidance	1.45 (1,65)	.916	0	46.028 (1,238)	<.001	0.071	Exp-cont (no family)[d]
Total battery	6.79 (1,65)	.036	0.011	105.88 (1,238)	<.001	0.148	Exp-cont (no family)[d]/exp-cont (family)[f]

[a]Bonferroni adjusted.

[b]Exp: experimental.

[c]Cont: control.

[d]*P*<.001.

[e]Persons without family members affected by severe mental illness.

[f]Persons with family members affected by severe mental illness.

Discussion

Principal Findings

This study sought to provide new methodologies to raise awareness among the young adult population, especially university students who in their professional future could work with people with severe mental disorders. There are very few studies that have used a web-based methodology [20], and at present, we do not know about any other study that has been based on an escape room, despite being a methodology that is being used more and more in the educational field. A recent systematic review on this topic did not identify any studies that applied this tool [21]. Applying them in a web-based format is also novel and, of course, was brought on by the effects of the pandemic, which resulted in a rapid adaptation to these methods [22].

However, as stated earlier, escape rooms are an increasingly used tool in the educational field [23,24] because of their playful contribution and proximity to the current language of young people.

In this study, we investigated the effectiveness of an escape room as an awareness tool based on empathizing or putting oneself in the role of a leading character with an SMI. The intervention significantly reduced the stigmatizing attitudes of university students in the experimental group compared with the control group. It is worth highlighting the size of the large effect observed in the factors of danger, perceived fear, and in solidarity, variables in which the greatest reduction was obtained.

Dangerousness is the most common stigma dimension in the general population [25], and it scored the highest in the pretreatment evaluation in this study. In contrast, because of the characteristics of the sample, the intervention also had a great effect on solidarity. More specifically, providing information about mental health problems in a participatory way, such as in the escape room, favored the strengthening of this variable.

The implementation of the escape room was carried out with university students from the health sector, which is important considering that some of these professionals will work directly with people affected by SMI, and their expectations and beliefs will affect the recovery of their patients [26]. Another aspect that should be highlighted is that professionals who work in the field of mental health continue to have negative beliefs, paternalistic attitudes, and even restrictive and directive attitudes on occasions [27]. These can have repercussions not only on recovery but also on the patients' own perception of themselves, causing them to label themselves in a stigmatizing way, which is known as self-stigma [28].

In contrast, it has been proven how the narrative of the story and immersion in the activity encourage the re-elaboration of value judgments that alter the beliefs and attitudes of people with SMI. Escape rooms have already been tested as active methodological tools that promote motivation and student commitment toward learning [10,29], obtaining good results when they have been applied. Similar studies [30-32] have found that it is possible to reduce stigma through other methodologies, and the interesting aspect of this study is that it is able to raise awareness through a playful activity that is easy to implement and replicate.

As for other findings, a variable found to influence the results is whether the person has a family member with mental health problems. When a person has a family member with this type of difficulty, we can observe that the stigma is lower and possibly owes to the greater knowledge and direct contact that an individual has around these problems, which is fundamental in the reduction of stigma [33].

Limitations

This study has several limitations. First, the sample is relatively small, it comes from 2 universities, and it is not representative of the general population. Furthermore, the sample of university students included only individuals from nursing schools. This may undermine the generalizability of the results and limit the interpretation of the effect size. Second, the specific characteristics of the participants and the possible confounding factors that may influence the results, such as their sociocultural origin, their level of knowledge, and their desire to work with patients with SMI, were not evaluated; the only exception was whether they had family members with SMI, in which case they were evaluated. Finally, another limitation was not conducting a reevaluation months after the intervention to determine whether the changes achieved in the experimental group could be perpetuated over time.

Conclusions

Mental health awareness is a very important subject for the proper professional development of health care students. It is the responsibility of university teachers to prepare students to learn, understand, and know how to work with their patients. The lack of theoretical knowledge and stigma toward people with SMI can contribute to discrimination toward this group, which, in turn, can influence their self-esteem and quality of life. Therefore, it is necessary to implement new strategies in university education to improve the knowledge and awareness of mental disorders. The integration of playful strategies such as escape rooms can be of great interest because of their immersive and motivating capacity.

Without Memories, a web-based escape room where the narrative is about a character with SMI, is an effective tool to promote awareness around mental health because it favors learning by discovery and playful challenges through attractive dynamics and mechanics, which can encourage students to learn about SMI and be more empathetic toward these people.

Future studies could use this escape room not only in nursing degree studies but also in other degrees related to health and community education to promote awareness among other professionals who also have direct contact with people with SMI. Furthermore, future research could assess whether the results on the impact of playful strategies for mental health awareness vary because of other sociodemographic factors. Finally, the duration of the escape room *Without Memories* is very limited, and although this may have benefits for its replication, there is a need to plan future awareness programs through playful strategies of longer duration to evaluate whether they can achieve a greater impact on motivation, acquisition of knowledge, and awareness.

Conflicts of Interest

None declared.

Editorial Notice

This randomized study was not registered, explained by authors on the basis that the subjects were nursing students. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, given that the subjects were undergraduate students. 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.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 381 KB - games_v10i2e34222_app1.pdf]

References

1. Conejo Cerón S, Moreno Peral P, Morales Asencio JM, Alot Montes A, García-Herrera JM, González López MJ, et al. Opiniones de los profesionales del ámbito sanitario acerca de la definición de trastorno mental grave. Un estudio cualitativo. An Sist Sanit Navar 2014;37(2):223-233 [FREE Full text] [doi: 10.4321/s1137-66272014000200005] [Medline: 25189980]
2. Zhang Y, Kuhn SK, Jobson L, Haque S. A review of autobiographical memory studies on patients with schizophrenia spectrum disorders. BMC Psychiatry 2019 Nov 14;19(1):361 [FREE Full text] [doi: 10.1186/s12888-019-2346-6] [Medline: 31727046]
3. Galán Casado D, Castillo Algarra J, García Tardón B. Deporte e inclusión social en personas con Trastorno Mental Grave (TMG). Psychol Soc Educ 2020;12(1):71-83. [doi: 10.25115/psye.v0i0.2343]
4. Macías FJ. Cómo encontrar un lugar en el mundo: explorando experiencias de recuperación de personas con trastornos mentales graves. Hist Cienc Saude Manguinhos 2011;18(1):121-139 [FREE Full text] [doi: 10.1590/s0104-59702011000100008] [Medline: 21552693]
5. Corrigan PW, Morris SB, Michaels PJ, Rafacz JD, Rüsch N. Challenging the public stigma of mental illness: a meta-analysis of outcome studies. Psychiatr Serv 2012;63(10):963-973. [doi: 10.1176/appi.ps.201100529] [Medline: 23032675]
6. Cangas AJ, Navarro N, Parra JM, Ojeda JJ, Cangas D, Piedra JA, et al. Stigma-Stop: a serious game against the stigma toward mental health in educational settings. Front Psychol 2017;8:1385 [FREE Full text] [doi: 10.3389/fpsyg.2017.01385] [Medline: 28878702]
7. Sanchez R, Brown E, Kocher K, DeRosier M. Improving children's mental health with a digital social skills development game: a randomized controlled efficacy trial of adventures aboard the S.S. GRIN. Games Health J 2017 Feb;6(1):19-27. [doi: 10.1089/g4h.2015.0108] [Medline: 28051877]
8. Ferchaud A, Seibert J, Sellers N, Escobar Salazar N. Reducing mental health stigma through identification with video game avatars with mental illness. Front Psychol 2020;11:2240 [FREE Full text] [doi: 10.3389/fpsyg.2020.02240] [Medline: 33013585]
9. Stasiak A. Escape rooms: a new offer in the recreation sector in Poland. Tour 2016;26(1):31-47. [doi: 10.1515/tour-2016-0003]
10. Hermanns M, Deal B, Campbell AM, Hillhouse S, Opella JB, Faigle C, et al. Using an "Escape Room" toolbox approach to enhance pharmacology education. J Nurs Educ Pract 2017;8(4):89-95. [doi: 10.5430/jnep.v8n4p89]
11. Veach CC. Breaking out to break through: re-imagining first-year orientations. Ref Serv Rev 2019;47(4):556-569. [doi: 10.1108/rsr-06-2019-0039]
12. Gómez-Urquiza JL, Gómez-Salgado J, Albendín-García L, Correa-Rodríguez M, González-Jiménez E, Cañadas-De la Fuente GA. The impact on nursing students' opinions and motivation of using a "Nursing Escape Room" as a teaching game: a descriptive study. Nurse Educ Today 2019 Jan;72:73-76. [doi: 10.1016/j.nedt.2018.10.018] [Medline: 30453202]
13. Aparicio-García ME, Díaz-Ramiro EM, Rubio-Valdehita S, López-Núñez MI, García-Nieto I. Health and well-being of cisgender, transgender and non-binary young people. Int J Environ Res Public Health 2018 Sep 28;15(10):2133 [FREE Full text] [doi: 10.3390/ijerph15102133] [Medline: 30274141]
14. Borrás-Gené O, Martínez-Núñez M, Martín-Fernández L. Enhancing fun through gamification to improve engagement in MOOC. Informatics 2019;6(3):28. [doi: 10.3390/informatics6030028]
15. Antonaci A, Klemke R, Kreijns K, Specht M. Get gamification of MOOC right!. Int J Serious Games 2018;5(3):61-78. [doi: 10.17083/ijsg.v5i3.255]
16. Corrigan PW, Powell KJ, Michaels PJ. Brief battery for measurement of stigmatizing versus affirming attitudes about mental illness. Psychiatry Res 2014;215(2):466-470 [FREE Full text] [doi: 10.1016/j.psychres.2013.12.006] [Medline: 24388505]
17. Manzano-León A, Camacho-Lazarraga P, Guerrero-Puerta MA, Guerrero-Puerta L, Alias A, Aguilar-Parra JM, et al. Development and validation of a questionnaire on motivation for cooperative playful learning strategies. Int J Environ Res Public Health 2021 Jan 22;18(3):960 [FREE Full text] [doi: 10.3390/ijerph18030960] [Medline: 33499224]
18. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
19. Saavedra J, Arias-Sánchez S, Corrigan P, López M. Assessing the factorial structure of the mental illness public stigma in Spain. Disabil Rehabil 2021 Sep;43(18):2656-2662. [doi: 10.1080/09638288.2019.1710769] [Medline: 31906732]
20. Rodríguez-Rivas ME, Cangas AJ, Fuentes-Olavarría D. Controlled study of the impact of a virtual program to reduce stigma among university students toward people with mental disorders. Front Psychiatry 2021;12:632252 [FREE Full text] [doi: 10.3389/fpsyt.2021.632252] [Medline: 33633613]
21. Rodríguez-Rivas ME, Cangas AJ, Cariola LA, Varela JJ, Valdebenito S. Innovative technology-based interventions to reduce stigma toward people with mental illness: a systematic review and meta-analysis. JMIR Serious Games (forthcoming) 2021. [doi: 10.2196/preprints.35099]
22. Javed B, Sarwer A, Soto EB, Mashwani ZU. The coronavirus (COVID-19) pandemic's impact on mental health. Int J Health Plann Manage 2020 Sep;35(5):993-996 [FREE Full text] [doi: 10.1002/hpm.3008] [Medline: 32567725]
23. Manzano-León A, Rodríguez-Ferrer JM, Aguilar-Parra JM, Martínez Martínez AM, Luque de la Rosa A, Salguero García D, et al. Escape rooms as a learning strategy for special education Master's degree students. Int J Environ Res Public Health 2021 Jul 08;18(14):7304 [FREE Full text] [doi: 10.3390/ijerph18147304] [Medline: 34299753]

24. Morrell B, Eukel HN. Shocking escape: a cardiac escape room for undergraduate nursing students. Simul Gaming 2020;52(1):72-78. [doi: 10.1177/1046878120958734]
25. Angermeyer MC, Dietrich S. Public beliefs about and attitudes towards people with mental illness: a review of population studies. Acta Psychiatr Scand 2006 Mar;113(3):163-179. [doi: 10.1111/j.1600-0447.2005.00699.x] [Medline: 16466402]
26. Wagner LC, Geidel AR, Torres-González F, King MB. Cuidado en salud mental: percepción de personas con esquizofrenia y sus cuidadores. Cien Saude Colet 2011 Apr;16(4):2077-2087 [FREE Full text] [doi: 10.1590/s1413-81232011000400007] [Medline: 21584449]
27. Bengochea-Seco R, Arrieta-Rodríguez M, Fernández-Modamio M, Santacoloma-Cabero I, Gómez de Tojeiro-Roce J, García-Polavieja B, et al. Adaptación al español de la escala Internalized Stigma of Mental Illness para valorar el estigma personal. Rev Psiquiatr Salud Ment (Engl Ed) 2018;11(4):244-254. [doi: 10.1016/j.rpsm.2016.01.007] [Medline: 26971249]
28. Wang K, Link BG, Corrigan PW, Davidson L, Flanagan E. Perceived provider stigma as a predictor of mental health service users' internalized stigma and disempowerment. Psychiatry Res 2018 Jan;259:526-531 [FREE Full text] [doi: 10.1016/j.psychres.2017.11.036] [Medline: 29156425]
29. Morrell BL, Ball HM. Can you escape nursing school? Educational escape room in nursing education. Nurs Educ Perspect 2020;41(3):197-198. [doi: 10.1097/01.NEP.0000000000000441] [Medline: 30614962]
30. Knaak S, Patten S. A grounded theory model for reducing stigma in health professionals in Canada. Acta Psychiatr Scand 2016 Aug;134 Suppl 446:53-62. [doi: 10.1111/acps.12612] [Medline: 27426646]
31. McNicholas F, O'Connor C, O'Hara L, McNamara N. Stigma and treatment of eating disorders in Ireland: healthcare professionals' knowledge and attitudes. Ir J Psychol Med 2016 Mar;33(1):21-31. [doi: 10.1017/ipm.2015.24] [Medline: 30115174]
32. Pascucci M, Ventriglio A, Stella E, Di Sabatino D, La Montagna M, Nicastro R, et al. Empathy and attitudes towards mental illness among Italian medical students. Int J Cult Ment Health 2017;10(2):174-184. [doi: 10.1080/17542863.2016.1276947]
33. Corrigan PW, Nieweglowski K. How does familiarity impact the stigma of mental illness? Clin Psychol Rev 2019 Jun;70:40-50. [doi: 10.1016/j.cpr.2019.02.001] [Medline: 30908990]

Abbreviations

AQ: Attributional Questionnaire
SMI: severe mental illness

Edited by N Zary; submitted 12.10.21; peer-reviewed by N Gronewold, A Teles, X Wang; comments to author 23.12.21; revised version received 09.02.22; accepted 17.02.22; published 05.05.22.

Please cite as:
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;10(2):e34222
URL: https://games.jmir.org/2022/2/e34222
doi:10.2196/34222
PMID:35511232

Original Paper

Modification in the Motor Skills of Seniors in Care Homes Using Serious Games and the Impact of COVID-19: Field Study

Jana Marina Kleschnitzki[1], MSc; Inga Grossmann[2], PhD; Reinhard Beyer[1], Prof Dr; Luzi Beyer[3], Prof Dr

[1]Institute of Psychology, Faculty of Human Sciences, Humboldt-University of Berlin, Berlin, Germany

[2]Department of Science, All About Me GmbH, Berlin, Germany

[3]Department of Quantitative Research Methods, Alice-Salomon University of Berlin, Berlin, Germany

Corresponding Author:
Jana Marina Kleschnitzki, MSc
Institute of Psychology
Faculty of Human Sciences
Humboldt-University of Berlin
Rudower Chaussee 18
Berlin, 12489
Germany
Phone: 49 15231964606
Email: janakleschnitzki@posteo.de

Abstract

Background: The pandemic has highlighted the importance of low-threshold opportunities for exercise and physical activity. At the beginning of 2020, the COVID-19 pandemic led to many restrictions, which affected seniors in care facilities in the form of severe isolation. The isolation led, among other things, to a lack of exercise, which has led to a multitude of negative effects for this target group. Serious games can potentially help by being used anywhere at any time to strengthen skills with few resources.

Objective: The aim of this study is to evaluate the effectiveness of a serious game to strengthen motor skills (study 1) and the influence of pandemic restrictions (study 2) on seniors in care facilities.

Methods: The data on motor skills (measured by the Tinetti test) originated from an intervention study with repeated measurements that was interrupted by the pandemic conditions. Data were collected 4 times every 3 months with an intervention group (IG, training 3 times for 1 hour per week) and a control group (CG, no intervention). There were 2 substudies. The first considered the first 6 months until the pandemic restrictions, while the second considered the influence of the restrictions on motor skills.

Results: The sample size was 70. The IG comprised 31 (44%) participants, with 22 (71%) female and 9 (29%) male seniors with an average age of 85 years. The CG comprised 39 (56%) participants, with 31 (79%) female and 8 (21%) male seniors with an average age of 87 years. In study 1, mixed-design ANOVA showed no significant interaction between measurement times and group membership for the first measurements ($F_{2,136}$=1.414, P<.25, partial η2=.044), but there was a significant difference between the CG (mean 16.23, SD 1.1) and the IG (mean 19.81, SD 1.2) at the third time of measurement (P=.02). In study 2 the mixed-design ANOVA (used to investigate motor skills before and after the pandemic conditions between the 2 groups) couldn't reveal any significant interaction between measurement times and group membership: $F_{1,67}$=2.997, P<.09, partial η2=.043. However, there was a significant main effect of the time of measurement: $F_{1,67}$=5.44, P<.02, partial η²=.075.

Conclusions: During the first 6 months, the IG showed increased motor skills, whereas the motor skills of the CG slightly deteriorated and showed a statistically significant difference after 6 months. The pandemic restrictions leveled the difference and showed a significant negative effect on motor skills over 3 months. As our results show, digital games have the potential to break down access barriers and promote necessary maintenance for important skills. The pandemic has highlighted the importance of low-threshold opportunities for exercise and physical activity. This potentially great benefit for the challenges of tomorrow shows the relevance of the topic and demonstrates the urgent need for action and research.

Trial Registration: Deutsches Register klinischer Studien DRKS00016633; https://tinyurl.com/yckmj4px

(JMIR Serious Games 2022;10(2):e36768) doi:10.2196/36768

KEYWORDS

serious games; motor skills; motor; long-term care; exercise; movement; coronavirus effects; eHealth; seniors; older adult; elder; senior population; aged; care home; intervention effects analysis; COVID-19; pandemic; digital game; digital health; physical activity

Introduction

Background

By the spring of 2020, the rapid spread of coronavirus led several German states to implement drastic measures in order to protect their population. Many of these measures included restrictions on movement and social contact, which particularly impacted groups that were already affected by multimorbidity, restricted functionality, and the need for care [1]. For instance, the high mortality rate, in particular among the older, more vulnerable population (6.5% of the population 65+ years old and 14% of those 75+ years old) [2,3], put a considerable burden on the health care system. This tragically resulted in an increase in restrictions stipulated and implemented by local authorities. Especially in nursing homes, caution and strict adherence to regulations were essential [4].

Physical Consequences of the Pandemic-Related Isolation

Compared to the results from a survey study conducted in 2017, the pandemic-related measures led to a sharp reduction in physical activity [5-7] in 40% of individuals between the ages of 46 and 90 years, a 29% increase in sitting time, and a widespread reduction in leisure activities [8-11]. By now, a number of negative effects of the measures implemented during the pandemic are known and documented. The literature, for instance, provides evidence of an exacerbation of psychological issues with possibly persistent, long-term effects [12-14], which are also associated with increased depressive and anxiety symptoms [15-19]. At the same time, due to physical inactivity, more negative physical symptoms have been found [20,21]. These, together with other mentioned effects, lead to increased mortality [22] and reduced quality of life [23], which can be additional triggers for withdrawal and physical inactivity (a vicious circle) [4,24]. An effective strategy to counteract these harmful effects is to keep up regular physical activity. It can promote resilience (self-efficacy and optimism) and mental health [25-27], reduce depressive symptoms [28], and balance an individual's overall psychomotor performance [1]. During the pandemic [29], even light physical activity can mitigate the negative effects [8]. Both the global recommendations for physical activities published by the World Health Organization (WHO) and the German National Health survey (*Bundes-Gesundheitssurvey*) issued by the Robert Koch Institute [30] point out that physical activity is key for seniors to maintain not only their health and mobility but also their independence and self-reliance [30,31], especially in times of a pandemic [32]. To be sustainable, the latter should be integrated into daily routines and carried out collectively [30].

Social Consequences of the Pandemic

Even before the outbreak of the COVID-19 pandemic, the German health care system was faced with major challenges related to demographic change. Among those were the growing number of aging individuals in need of care (22% of 60-80-year-olds) [33] and, as a result, the increased utilization of health care services, which burdened the social security system and the health care structure [34]. Moreover, the rapidly growing need for support caused by the pandemic has placed excessive demands [35] on practitioners, creating an enormous amount of stress, which has negatively impacted the quality of care and, by extension, also the psychological well-being of seniors [36]. The forecasted shortage of half a million skilled workers in the German nursing sector by 2035 is also alarming, given that it may further exacerbate the issue [37]. Therefore, it is imperative to focus on health care promotion for seniors, which concentrate on health care resources that promise long-term autonomy and independence [38].

Effectiveness of Digital Serious Games That Promote Physical Activity

The positive effect of physical activities [39] combined with the challenges outlined before highlight the need for low-threshold, inexpensive, nationwide solutions that promote physical activity [40]. Digital technologies can play a key role in this effort [41]. These new digital technologies in the health care sector that primarily focus on prevention and health promotion via gameful experiences can be summarized under the umbrella term "serious games for health" [42].

The overall positive effect of serious games with a special focus on promoting motor movement sequences in older individuals has already been investigated and reported in several meta-analyses [43-47] that highlighted that the impact resulting from the use of serious games can range from moderate-to-strong training effects, especially in participants with initially low motor capability [48]. Vaziri et al [49] specifically point to a significant reduction in the risk of falling, especially in participants with low initial motor capability [50], while Stanmore [50] and Brox [51] were even able to report the first positive follow-ups in this regard. However, it needs to be emphasized that long-term effectiveness of physical activity has not been sufficiently investigated yet [52-54].

Initial studies conducted during the pandemic show the usefulness and positive effects of digital exercise on physical well-being [55]. For example, Parker et al [56] reported that participants who used digital exercise platforms were more likely to remain physically active during the pandemic. Thus, serious games can be part of a rehabilitative treatment plan that might prevent the functional deterioration of seniors suffering from COVID-19 [57]. In conclusion, it can be said that most of the findings indicate promising potential for using game-based interventions with senior citizens, thus highlighting the importance of establishing such opportunities, especially during the pandemic [58]. However, there is a clear need for research into the long-term effects [59], the study conditions [45], and the intensity and choice of suitable game modules relative to different physical conditions and symptoms. To exploit the potential of serious games in the psychomotor domain, additional variables related to game engagement, such as usefulness, purpose, and user orientation, must be clearly defined and implemented in the context of empirical studies [60].

This Study

Despite the positive effects of physical activity and the preliminary evidence that serious games have a positive impact on health and physical well-being, new digital technologies [61] should be researched more to become a key measure in the fight against the social- and health care–related challenges, negative effects, and consequences of the pandemic. This intervention study does so by analyzing the effects of the restrictive measures implemented during the COVID 19 pandemic in interaction with exergaming a social and physical activity on seniors' motor skills over the course of a long period. The study is unique insofar as the participants began the intervention before the onset of COVID-19 and carried on during the pandemic (with its restrictive measures). Key questions were, “Can the positive effects reported in the literature be replicated?” and “Can serious games cushion the far-reaching, undoubtedly negative impact of the pandemic measures on the motor skills of seniors?”

Methods

MemoreBox

In 2014, MemoreBox was developed in response to the need for promoting prevention and health relative to the cognitive, motor, and psychosocial skills of senior citizens and, in particular, nursing home residents. MemoreBox is a health game that includes a gesture-controlled game console, which records the participants' movement data by means of a Kinect camera (Microsoft Corporation) and an individual quick response (QR) code. Games can be played when sitting or standing, individually and in groups, and the therapeutic training program can be used preventively and independently of any indications. There are currently 6 games (motorcycling, bowling [Figure 1], table tennis, singing, postman, and dancing), which are based on everyday activities. Depending on the game, the player’s balance, memory, ability to react, hand-eye coordination, motor skills, stance, or mobility are trained.

Figure 1. Exemplary Bowling module in the serious game MemoreBox.

Study Design

As part of a large, applied research study, the data used here for the analysis came from an intervention study that was set up as a quasi-experimental design with repeated measurements. The overarching goal of this intervention study was to test the effectiveness of MemoreBox. To evaluate the game’s effectiveness, we designed a large-scale study with 100 German nursing homes. Over the course of 1 year, we recorded and examined a total of 1000 seniors in an intervention group (IG) and a control group (CG). The intervention, training by playing, was carried out 3 times per week for 1 hour in a group using a fixed training plan that was developed in advance by occupational therapists. Data were collected over the course of 1 year during which participants were asked to complete standardized questionnaires 8 times. The standardized questionnaires on cognition, motor skills, and psychosocial health were administered every 3 months and thus 4-5 times in total. In this paper, we focus on the participants' motor skills, which were recorded at 4 points over the course of 9 months using the Tinetti test [62].

COVID-19 Situation

Due to the global COVID-19 pandemic, which has also caused major changes in Germany, at least since March 2020, the research could not be carried out as originally planned. Protection of and care for the target population, that is, vulnerable senior citizens who are dependent on care due to physical or psychological frailties, is the top priority. For a large majority of the participating nursing homes, this situation led to an inevitable interruption of the study and thus of the planned data collection (also due to a lack of appropriate facilities) starting at the beginning of March 2020. However, the situation and the respective restrictions caused by COVID-19 differed greatly from 1 nursing home to the next. Given that such a large number of facilities were originally recruited throughout Germany, a few nursing homes emerged that were able to continue with the study and data collection due to their particular local setups and conditions within the respective facilities. These 11 nursing homes were able to continue the training plan unaltered, which ultimately led to a sample size of 70 participants (IG n=31, 44%, vs CG n=39, 56%).

Participants

The sample consisted of an IG that played regularly and a CG that did not play. All participants are residents of nursing homes in Germany. We started with 10 participants per nursing home (IG n=5 and CG n=5, 50% each), that is, N=1000. The assignment of groups was voluntary for practical and ethical reasons, which required a final parallelization of the data. Exclusion criteria for both groups were severe mental or neurological illnesses and age below 60 years. In addition, the state of health, comorbidities, and medications of both groups were surveyed.

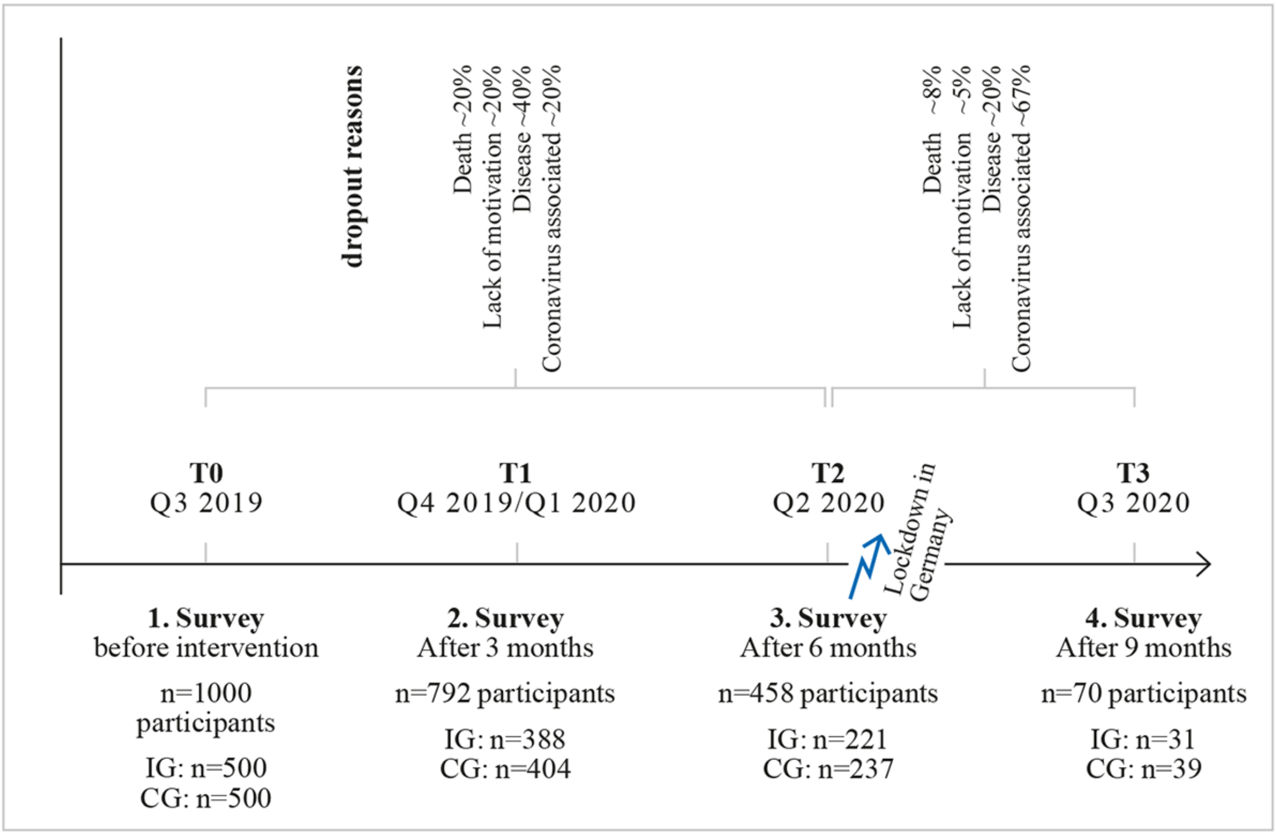
Data Collection Process and Instruments

All data collection was carried out by scientifically trained nursing staff. The data collection instruments were designed to suit the target population, and age-related restrictions in the process were considered. This paper deals with the results of the Tinetti test [62], 1 of the standardized questionnaires used to investigate the effectiveness of MemoreBox on the motor performance of seniors in relation to the effects of the pandemic. The Tinetti test is used to evaluate a person's static and dynamic balance skills. The test is divided into 2 components for separate evaluation of gait and balance skills. Both measurement units consist of 8 questions, with the answer options for the walking test being rated from 0 to 2 points and those for the balance test being rated from 0 to 4 points [63]. The maximum total of the Tinetti test is 28 points, made up of 15 points for the balance test and 13 points for the walking test [63].

Dropouts

The enormously high dropout rate was largely a result of the outbreak of the COVID-19 pandemic (see description before). Given the length of the study and the target population examined, a high dropout rate was generally assumed. Illnesses and deaths, as well as lack of motivation, were the other main reasons for dropping out. The weightings shown in Figure 2 roughly correspond to those reported in the literature. Over the entire data collection period, the following reasons led to attrition: 409 (44%) seniors dropped out due to COVID-19-associated reasons, 279 (30%) dropped out due to illness, 130 (13%) passed away, and 116 (13%) dropped out due to a lack of motivation.

Figure 2. Survey times, periods of operationalization, and dropout reasons. CG: control group; IG: intervention group.



Statistical Analysis

At the beginning of the data analysis, the 2 study groups were compared with regard to their characteristics in the dependent variables at baseline (T0). The reason for this was the nonrandomized group assignment, which made *t* tests for independent samples between the groups necessary in order to statistically demonstrate the comparability of the groups. Then, descriptive statistics regarding the groups and relevant sociodemographic variables were carried out. For psychometric calculations, the statistical significance was set at a level of $\alpha=.05$. To highlight the relevance or practical significance of the results, we calculated the effect size partial η^2 for mixed-design ANOVA ($\eta^2<.01$, small effect; $\eta^2<.06$, medium effect; $\eta^2<.14$, large effect) [63]. To examine the effects on the IG in comparison to the CG, a 2×3 and a 2×2 mixed-design ANOVA were used in 2 study sections. The dependent variable was motor ability; the factor time, with study segment 1 having 3 levels (T0, T1, T2,) and study segment 2 having 2 levels (T2, T3), acted as an inner-subject factor and group membership as a between-subject factor (2 factor levels). Corresponding Bonferroni corrections within and between subjects were also calculated. The assumptions for the interval scale level of the dependent variable, independence, and the nominal scale level of the between- and within-subject factors can be assumed. Outliers were removed based on the IQR. The assumptions of the normal distribution of the dependent variable residuals, the homogeneity of variance, the equality of the covariances, and the given sphericity can, according to Bortz and Schuster [64], be neglected due to the robustness of the variance analyses if the sample size is almost the same and n>10.

Ethics Approval

This study was approved by the Ethics Committee of Charite Berlin (Ethikausschuss am Campus Benjamin Franklin; review number: EA4/035/19).

Results

Baseline Comparison

Due to the lack of randomization of the study groups, independent sample *t* tests were used for the demographic variables to reveal possible distorting differences between groups. The groups did not differ with regard to the main variables at the beginning of the study (Table 1).

Table 1. Mean values, SDs, and statistical differences of the examined variables: IG[a] and CG[b] at measurement time T0.

Variables	IG (N=31), mean (SD)	CG (N=39), mean (SD)	Statistics		95% CI
			t (*df*)	*P* value	
Age (years)	85.45 (4.99)	86.66 (8.76)	.718 (60.5)	.48	–0.311 to 0.639
Level of care needed[c]	2.34 (0.974)	2.47 (0.71)	.592 (61)	.56	–0.347 to 0.645
State of health[d]	2.33 (1.07)	3.16 (1.22)	–.623 (66)	.54	–0.631 to 0.328
Health behavior[e]	2.42 (0.720)	2.54 (0.97)	.570 (68)	.57	–0.335 to 0.609
Health self-assessment[f]	2.84 (0.735)	2.92 (0.81)	.452 (68)	.65	–0.364 to 0.580
Tinetti mean	1.15 (0.37)	1.05 (0.41)	–.974 (68)	.33	–0.707 to 0.240
Tinetti total sum	18.16 (5.88)	16.67 (6.58)	–.989 (68)	.33	–0.710 to 0.236

[a]IG: intervention group (playing regularly).

[b]CG: control group.

[c]0=no need for care to 5=most severe impairment.

[d]0= healthy to 5=chronically ill.

[e]0=does not take care of health to 5=strongly focuses on health.

[f]0="I rate my health as very bad" to 5="I rate my health as very good."

Descriptive Statistics

The analysis sample was part of the total sample and included those participants who were present at all measurement times and throughout the entire intervention. Therefore, this study included 70 participants. The IG (participants who played regularly) comprised 31 (44%) participants, of which 22 (71%) were female and 9 (29%) male. Participants in the IG ranged were aged from 74 to 97 (mean 85.45, SD 4.99) years. The CG included 31 (79%) females and 8 (21%) males, who were aged from 61 to 102 (mean 86.66, SD 8.76) years. The distribution of age (mean 86.06, SD 8.2 years) and gender (female n=53, 76%, male n=17, 24%) in the total sample roughly corresponded to the findings on the need for care in Germany [65].

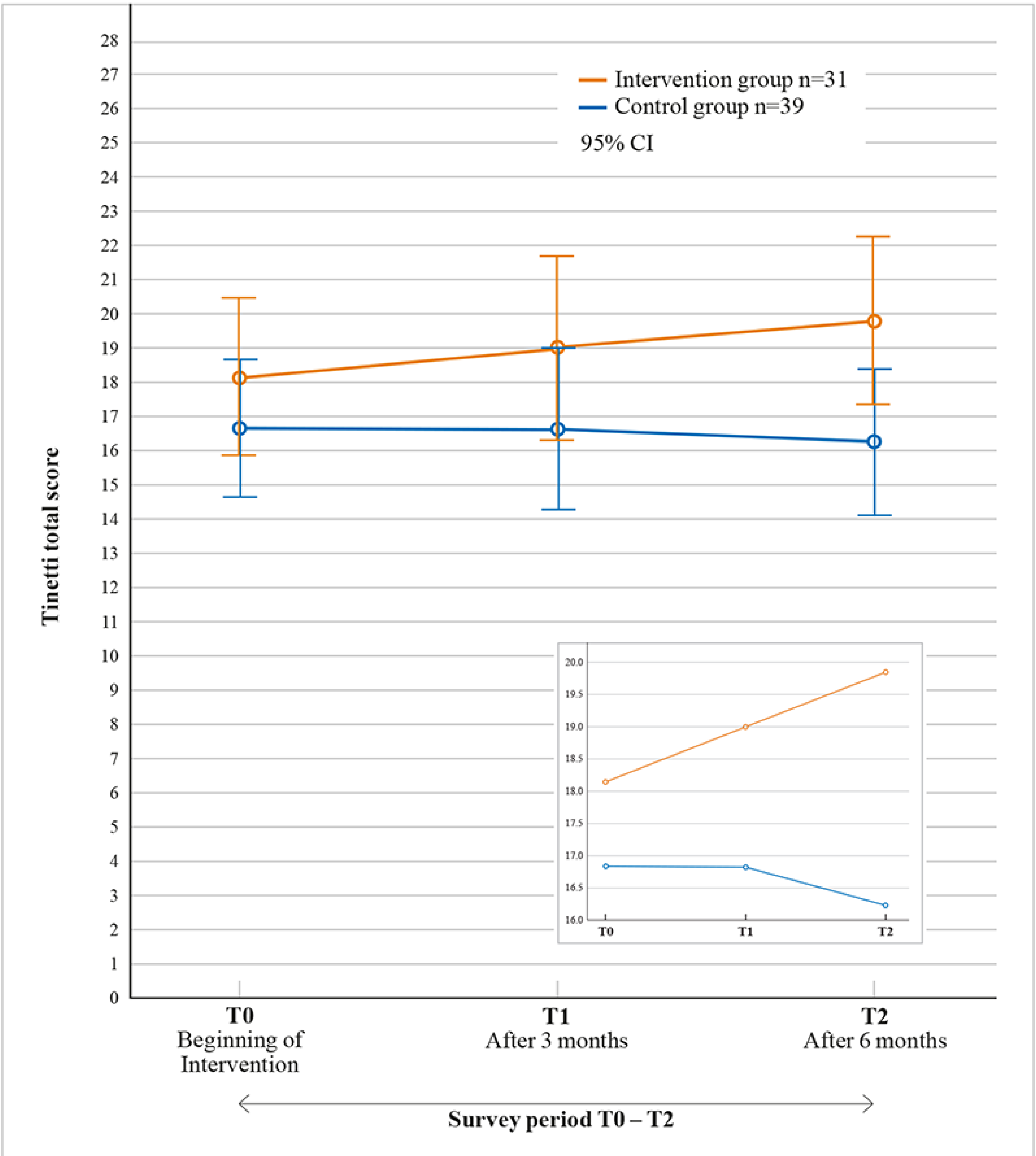
Outcome

Next, the results of the Tinetti data collected from the 70 participants over the course of 1 year were analyzed separately in 2 study sections. Study 1 looked at the first 6 months of the intervention, which analyzed the situation before the outbreak of the pandemic. The second study then dealt with the analysis of the data collected during the pandemic with its severe restrictions for the participants. Both studies looked at the CG in comparison to the IG, which carried out the study plan consistently despite massive, pandemic-related restrictions.

Study 1

Mixed-design ANOVA showed no statistically significant interaction between measurement times and group membership: $F_{2.136}$=1.414, P<.25, partial η^2=.044. Both main effects (time factor: $F_{2.136}$=0.489, P=.62, partial η^2=.016; group membership factor: $F_{1.68}$=2.792, P=.10, partial η^2=.044) showed no statistical significance. However, the post hoc tests carried out within and between subjects indicated a statistically significant difference between the CG (mean 16.23, SD 1.1) and the IG (mean 19.81, SD 1.2) at the third time of measurement (P=.02). This finding shows that the regular players differed statistically significantly in their motor skills from the nonplayers after 6 months of playing (Figure 3).

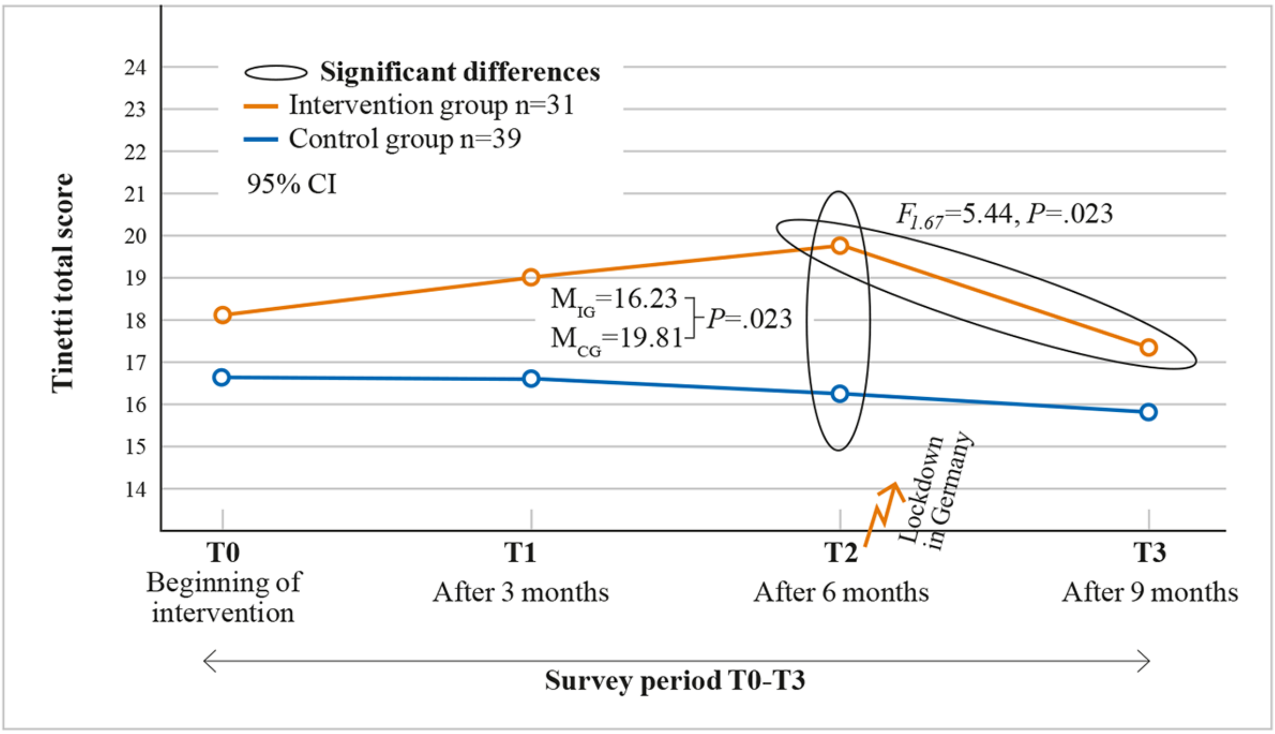
Figure 3. Results of the Tinetti total score for the IG and CG over 6 months of intervention. T0: IG (mean 18.16, SD 5.88), CG (mean 16.67, SD 6.58); T1: IG (mean 19, SD 6.77), CG (mean 16.62, SD 7.93); T2: IG (mean 19.81, SD 6.34), CG (mean 16.23, SD 7.14); CG: control group; IG: intervention group.



Study 2

Mixed-design ANOVA used to investigate motor skills before and after the outbreak of the pandemic between the groups did not reveal any statistically significant interaction between measurement times and group membership: $F_{1,67}$=2.997, P=.09, partial η²=.043. However, there was a statistically significant main effect of the time of measurement, which indicates a significant difference between the measurement times: $F_{1,67}$=5.44, P=.02, partial η²=.043. However, there was a statistically significant main effect of the time of measurement, which indicates a significant difference between the measurement times: $F_{1,67}$=5.44, P<.02, partial η²=.075. Considering the descriptive statistics (Figure 4), one can assume a significant drop in estimates between the 2 measurement times. A statistically significant main effect of group membership ($F_{1,67}$=2.34, P=.12, partial η²=.043) was not found.

Figure 4. Significant differences after 9 months of intervention between IG and CG and between measurement times. CG: control group; IG: intervention group.



Discussion

Principal Findings

Senior citizens were particularly affected by the COVID-19 pandemic, which led to far-reaching and extensive measures of isolation. These resulted in a significant reduction in physical activity, which is counterproductive for an independent life and disease prevention. The WHO states that high physical activity is essential for an independent life and disease prevention. Physical inactivity not only leads to deterioration and dependency in physical, psychological, and social areas, it also accelerates the need for nursing and health care services [66].

The goal of this paper was to investigate whether the use of the serious game MemoreBox has a positive impact on the motor skills of seniors in nursing home facilities. Additionally, the data collected provide insights into the effects of COVID-19 restrictions (especially isolating measures) on the development of motor skills of the participants.

Discussion Study 1

The data analysis of study I clearly shows that during the first 6 months of the study, the IG had increased motor skills, whereas the motor skills of the CG slightly deteriorated. After 6 months, there was a statistically significant difference in motor skills between the 2 groups (Figure 3). These findings are in line with the predominantly promising potential of video game–based interventions (serious games) for the motor skill development of seniors [42-45]. Under known conditions and in the context of a regular everyday life routine, motor stability can be built up, therefore reducing the risk of falling in old age [49]. The results of this study also support the known positive tendencies toward improved cognition in the IG. However, the missing main effects need to be addressed and discussed. Due to the clear trend of the data, it seems conclusive that the time frame for changes in particular is not sufficient. It should also be noted that there are many motor impairments in seniors that cannot be reversed or influenced by training. This might have a major impact on the test results, which do not measure these limitations. It seems imperative to address these points in future research.

Discussion Study 2

After 6 months of intervention, the outbreak of the COVID-19 pandemic led to severe restrictions in nursing home facilities, which essentially led to isolation and separation, resulting in an overall strong reduction in physical activity. The effects can be clearly seen in the collected data, which show a significant decrease in the motor skills previously gained, resulting in the IG almost falling back to the initial level (Figure 4) and a continued decrease in the CG motor skills as well. These data clearly show the serious effects of isolation measures on motor skills. In line with previous research, the data confirm a deterioration in motor skills, associated with a higher risk of falling [67]. The continued use of serious games does not seem to be able to stop this deterioration in isolation, a finding that may be due to the intensity of the changes. It can be assumed that mobility is negatively affected by the failure of conventional therapeutic, individually tailored, and medically necessary offers such as physical and occupational therapy. Opportunities to exercise without a special therapeutic approach, such as walks outside, excursions, or other therapy-related activities, have also ceased to exist during the pandemic-induced isolation. Isolation reduces social interaction and affects mental health, ultimately leading to reduced physical performance. Regardless of the use of MemoreBox, the lack of social contact in connection with far-reaching changes in daily routines had a considerable negative effect on the overall physical health of the participants.

Implications

Based on these findings, we conclude that serious games can have a positive influence on the motor skills of seniors. However, despite continued use, other severe motor restrictions (in this case, isolation and separation) can counteract the gain in motor skills. Moreover, regardless of the examined intervention, the data clearly show that COVID-19-related restrictions had a significantly negative influence on the motor skills of the participating seniors. Despite the subsequent limitations, the longitudinal data underlying this study offered a valuable and rare opportunity to examine the direct effects the isolating measures implemented during the COVID-19 pandemic have on senior citizens' health and mobility. Regarding the reported results on the effectiveness of MemoreBox, this intervention certainly provides a complementary measure to already established prevention programs.

Limitations

The data collection instrument designed by the study and the examined data have clear limitations. The design limitations, which were analyzed in detail in Kleschnitzki et al [68], are primarily due to the voluntary assignment of the participants to the groups. The lack of randomization is an artifact of the practical implementation.

The most important limitation of this study, however, is the sample size, which implies that the results cannot necessarily be generalized. However, there were at least 30 participants in each study group, which allowed for statistical methods. Furthermore, this population is difficult to reach. This stems from limitations (motor, cognitive, social) with regard to their location. Additionally, the longitudinal design of this study (most studies examine a period of 3 months) also posed great challenges for the participants and led to high dropout rates due to illnesses and deaths. Furthermore, there were large variances within the groups and thus among the participants. This variability, which had a similar range at all measurement times, also posed challenges to making possible significant differences visible. However, the examination of smaller, more homogeneous groups in regard to identifying potential statistically significant differences relative to certain characteristics did not provide any conclusive evidence. The G*Power analysis revealed an a priori minimum sample size of >260 (study 1) and >224 (study 2).

The outbreak of the COVID-19 pandemic undoubtedly had the greatest negative impact on the sample size. It led to high dropout rates among participants and created new challenges for data collection. Nevertheless, it also allowed us to collect data on COVID-19-related changes and associated limitations.

Future Studies

These limitations and analyzed consequences, in terms of their sustainability, cannot be fully foreseen at the moment and will certainly need to be examined more. The considerable variability in study designs and foci (eg, game applications, data collection instruments, target populations) makes it challenging to compare results from different studies that have examined the effectiveness of serious games [69,70]. Hence, there is a clear need for high-quality, large-scale studies that further investigate the use of serious games by senior citizens. Following a good scientific standard, these studies should focus on the effects that serious games may have on motor skills and fall prevention among seniors. Importantly, studies should consider the limitations of our research and thus include a larger sample size, a longitudinal study design, a randomized trial, and optimized training conditions. It should also be noted that the pandemic provides interesting areas of inquiry for different research domains. The insights gained from pandemic-related research could ultimately help to better understand and counterbalance the negative effects of the COVID-19 pandemic. Finally, the importance and feasibility of using serious games relative to the various issues that the health care system is facing and to the rising number of people in need of care should be investigated further.

Conclusion

Our results show that digital games can be deployed to promote health in a variety of contexts. They have the potential to break down access barriers and promote social engagement and interaction. The pandemic has highlighted the importance of low-threshold opportunities for exercise and physical activity, especially when conventional recreational programs are either greatly reduced or temporarily not available.

The results of this study help to further establish this new research area by (1) identifying that serious games can have a positive effect on motor skills of senior citizens and (2) revealing critical insights into the effects of COVID-19 on the motor development of seniors in isolation (physical inactivity, increased risk of falling, etc). Additionally, the results are discussed relative to the critical need for action and further research in that area. Lastly, our study contributes to identifying future global health-related challenges as well as potential preventive measures that could be developed and implemented in order to enable seniors, a consistently increasing population, to lead a healthier, independent, and more active lives.

Acknowledgments

We acknowledge the tremendous support from the RetroBrain and the BARMER team. They not only developed the product with a lot of dedication but also played a key role in the organization, communication, and implementation of the rollout. We also acknowledge Janik Traunspurger, who actively supported us with the data material. Another thanks goes to Kara Grote for graphical support.

A big thank-you also goes to all the nursing staff and the seniors who, under difficult conditions (at least since the pandemic), fulfilled the project with their effort and commitment.

Conflicts of Interest

None declared.

References

1. Witt K. Ausgewählte Themen der Sportmotorik für das weiterführende Studium (Band 2). Switzerland: Springer Nature; 2018.
2. Centre for Evidence-Based Medicine. Global COVID-19 Case Fatality Rates. URL: https://www.cebm.net/covid-19/global-covid-19-case-fatality-rates/ [accessed 2022-05-03]
3. Tiemann M, Mohokum M. Demografischer Wandel, Krankheitspanorama, Multimorbidität und Mortalität in Deutschland. Springer Reference Pflege - Therapie - Gesundheit 2021:3-11. [doi: 10.1007/978-3-662-62426-5_1]
4. Lebrasseur A, Fortin-Bédard N, Lettre J, Raymond E, Bussières EL, Lapierre N, et al. Impact of the COVID-19 pandemic on older adults: rapid review. JMIR Aging 2021 Apr 12;4(2):e26474 [FREE Full text] [doi: 10.2196/26474] [Medline: 33720839]
5. Visser M, Schaap LA, Wijnhoven HAH. Self-reported impact of the COVID-19 pandemic on nutrition and physical activity behaviour in Dutch older adults living independently. Nutrients 2020 Nov 30;12(12):3708 [FREE Full text] [doi: 10.3390/nu12123708] [Medline: 33266217]
6. Wang Y, Zhang Y, Bennell K, White DK, Wei J, Wu Z, et al. Physical distancing measures and walking activity in middle-aged and older residents in Changsha, China, during the COVID-19 epidemic period: longitudinal observational study. J Med Internet Res 2020 Oct 26;22(10):e21632 [FREE Full text] [doi: 10.2196/21632] [Medline: 33027035]
7. Yamada M, Kimura Y, Ishiyama D, Otobe Y, Suzuki M, Koyama S, et al. Effect of the COVID-19 epidemic on physical activity in community-dwelling older adults in Japan: a cross-sectional online survey. J Nutr Health Aging 2020 Jun 23;24(9):948-950. [doi: 10.1007/s12603-020-1501-6]
8. Ammar A, Brach M, Trabelsi K, Chtourou H, Boukhris O, Masmoudi L, et al. Effects of COVID-19 home confinement on eating behaviour and physical activity: results of the ECLB-COVID19 International Online Survey. Nutrients 2020 May 28;12(6):1583 [FREE Full text] [doi: 10.3390/nu12061583] [Medline: 32481594]
9. Ong JL, Lau TY, Massar SAA, Chong ZT, Ng BKL, Koek D, et al. COVID-19-related mobility reduction: heterogenous effects on sleep and physical activity rhythms. Sleep 2021 Feb 12;44(2):zsaa179 [FREE Full text] [doi: 10.1093/sleep/zsaa179] [Medline: 32918076]
10. Tison GH, Avram R, Kuhar P, Abreau S, Marcus GM, Pletcher MJ, et al. Worldwide effect of COVID-19 on physical activity: a descriptive study. Ann Intern Med 2020 Nov 03;173(9):767-770. [doi: 10.7326/m20-2665]
11. Sala G, Jopp D, Gobet F, Ogawa M, Ishioka Y, Masui Y, et al. The impact of leisure activities on older adults' cognitive function, physical function, and mental health. PLoS One 2019 Nov 8;14(11):e0225006 [FREE Full text] [doi: 10.1371/journal.pone.0225006] [Medline: 31703115]
12. Taylor HO, Taylor RJ, Nguyen AW, Chatters L. Social isolation, depression, and psychological distress among older adults. J Aging Health 2018 Feb 17;30(2):229-246 [FREE Full text] [doi: 10.1177/0898264316673511] [Medline: 28553785]
13. Schrempft S, Jackowska M, Hamer M, Steptoe A. Associations between social isolation, loneliness, and objective physical activity in older men and women. BMC Public Health 2019 Jan 16;19(1):74 [FREE Full text] [doi: 10.1186/s12889-019-6424-y] [Medline: 30651092]
14. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920. [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
15. Faulkner J, O'Brien WJ, McGrane B, Wadsworth D, Batten J, Askew CD, et al. Physical activity, mental health and well-being of adults during initial COVID-19 containment strategies: a multi-country cross-sectional analysis. J Sci Med Sport 2021 Apr;24(4):320-326 [FREE Full text] [doi: 10.1016/j.jsams.2020.11.016] [Medline: 33341382]
16. Meyer J, McDowell C, Lansing J, Brower C, Smith L, Tully M, et al. Erratum: Meyer, J., et al. Changes in physical activity and sedentary behavior in response to COVID-19 and their associations with mental health in 3052 US Adults. 2020 (18), 6469. Int J Environ Res Public Health 2020 Sep 23;17(19):6949 [FREE Full text] [doi: 10.3390/ijerph17196949] [Medline: 32977506]
17. Huang Y, Zhao N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. Psychiatry Res 2020 Jun;288:112954 [FREE Full text] [doi: 10.1016/j.psychres.2020.112954] [Medline: 32325383]
18. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]

19. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
20. Callow DD, Arnold-Nedimala NA, Jordan LS, Pena GS, Won J, Woodard JL, et al. The mental health benefits of physical activity in older adults survive the COVID-19 pandemic. Am J Geriatr Psychiatry 2020 Oct 30;28(10):1046-1057 [FREE Full text] [doi: 10.1016/j.jagp.2020.06.024] [Medline: 32713754]
21. Mattioli AV, Sciomer S, Cocchi C, Maffei S, Gallina S. Quarantine during COVID-19 outbreak: changes in diet and physical activity increase the risk of cardiovascular disease. Nutr Metab Cardiovasc Dis 2020 Aug 28;30(9):1409-1417 [FREE Full text] [doi: 10.1016/j.numecd.2020.05.020] [Medline: 32571612]
22. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. Proc Natl Acad Sci U S A 2013 Apr 09;110(15):5797-5801 [FREE Full text] [doi: 10.1073/pnas.1219686110] [Medline: 23530191]
23. 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]
24. Goethals L, Barth N, Guyot J, Hupin D, Celarier T, Bongue B. Impact of home quarantine on physical activity among older adults living at home during the COVID-19 pandemic: qualitative interview study. JMIR Aging 2020 May 07;3(1):e19007 [FREE Full text] [doi: 10.2196/19007] [Medline: 32356777]
25. Raglin JS. Exercise and mental health. Beneficial and detrimental effects. Sports Med 1990 Jun;9(6):323-329. [doi: 10.2165/00007256-199009060-00001] [Medline: 2192422]
26. Gordon BR, McDowell CP, Hallgren M, Meyer JD, Lyons M, Herring MP. Association of efficacy of resistance exercise training with depressive symptoms: meta-analysis and meta-regression analysis of randomized clinical trials. JAMA Psychiatry 2018 Jun 01;75(6):566-576 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.0572] [Medline: 29800984]
27. Ashdown-Franks G, Firth J, Carney R, Carvalho AF, Hallgren M, Koyanagi A, et al. Exercise as medicine for mental and substance use disorders: a meta-review of the benefits for neuropsychiatric and cognitive outcomes. Sports Med 2020 Jan 20;50(1):151-170. [doi: 10.1007/s40279-019-01187-6] [Medline: 31541410]
28. Carriedo A, Cecchini JA, Fernandez-Rio J, Méndez-Giménez A. COVID-19, psychological well-being and physical activity levels in older adults during the nationwide lockdown in Spain. Am J Geriatr Psychiatry 2020 Nov;28(11):1146-1155 [FREE Full text] [doi: 10.1016/j.jagp.2020.08.007] [Medline: 32919872]
29. Damiot A, Pinto A, Turner J, Gualano B. Immunological implications of physical inactivity among older adults during the COVID-19 pandemic. Gerontology 2020 Jun 25;66(5):431-438 [FREE Full text] [doi: 10.1159/000509216] [Medline: 32585674]
30. Krug S, Jordan S, Mensink G, Müters S, Finger J, Lampert T. Physical activity: results of the German Health Interview and Examination Survey for Adults (DEGS1). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013 May 25;56(5-6):765-771. [doi: 10.1007/s00103-012-1661-6] [Medline: 23703496]
31. Voelcker-Rehage C, Godde B, Staudinger UM. Activity, physical and psychological mobility in old age. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2006 Jun;49(6):558-566. [doi: 10.1007/s00103-006-1269-9] [Medline: 16673070]
32. Bailey L, Ward M, DiCosimo A, Baunta S, Cunningham C, Romero-Ortuno R, et al. Physical and mental health of older people while cocooning during the COVID-19 pandemic. QJM 2021 Nov 13;114(9):648-653 [FREE Full text] [doi: 10.1093/qjmed/hcab015] [Medline: 33471128]
33. Strutz N, Kiselev J, Lahmann N. Prädiktoren von Sturzereignissen in Pflegewohnheimen: eine Querschnittsstudie in Deutschland. HBScience 2020 Aug 04;11(3-4):44-51. [doi: 10.1007/s16024-020-00340-w]
34. Miller VJ, Fields NL, Anderson KA, Kusmaul N, Maxwell C. Nursing home social workers perceptions of preparedness and coping for COVID-19. J Gerontol B Psychol Sci Soc Sci 2021 Mar 14;76(4):e219-e224 [FREE Full text] [doi: 10.1093/geronb/gbaa143] [Medline: 32861215]
35. Li J, Wang J, Kong X, Gao T, Wu B, Liu J, et al. Person-centered communication between health care professionals and COVID-19-infected older adults in acute care settings: findings from Wuhan, China. J Gerontol B Psychol Sci Soc Sci 2021 Mar 14;76(4):e225-e229 [FREE Full text] [doi: 10.1093/geronb/gbaa190] [Medline: 33136158]
36. Flake R, Kochskämper S, Risius P, Seyda S. IW-Trends 3/2018 Fachkräfteengpass in der Altenpflege. URL: https://iwkoeln.de/fileadmin/user_upload/Studien/IW-Trends/PDF/2018/IW-Trends_2018-03-02_Pflegefallzahlen.pdf [accessed 2022-05-03]
37. Schmedes C. Emotionsarbeit in der Pflege: Beitrag zur Diskussion über die psychische Gesundheit Pflegender in der stationären Altenpflege. Berlin: Springer VS; 2021.
38. Geuter G. Bewegungsförderung und Gesundheit im Alter. Public Health Forum 2012;20(1):9.e1-9.e4. [doi: 10.1016/j.phf.2011.12.005]
39. John M, Kleppisch M. Digitale Gesundheitsanwendungen in der Prävention und Gesundheitsförderung ? Stand der Technik und Praxis. Springer Reference Pflege - Therapie - Gesundheit 2021:1019-1037. [doi: 10.1007/978-3-662-62426-5_108]
40. 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 25;63(6):698-707. [doi: 10.1007/s00103-020-03156-1] [Medline: 32451596]

41. Nowossadeck S, Wettstein M, Cengia A. Körperliche Aktivität in der Corona- Pandemie: Veränderung der Häufigkeit von Sport und Spazierengehen bei Menschen in der zweiten Lebenshälfte 2021 May. Berlin: Deutsches Zentrum für Altersfragen; 2021.
42. 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]
43. Wiemeyer J. Spielerische Förderung körperlicher Aktivität von Älteren. Präv Gesundheitsf 2018 Jul 5;13(4):285-291. [doi: 10.1007/s11553-018-0656-z]
44. Xu W, Liang H, Baghaei N, Wu Berberich B, Yue Y. Health benefits of digital videogames for the aging population: a systematic review. Games Health J 2020 Dec 24;9(6):389-404. [doi: 10.1089/g4h.2019.0130] [Medline: 32589482]
45. 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]
46. Vinolo Gil MJ, Gonzalez-Medina G, Lucena-Anton D, Perez-Cabezas V, Ruiz-Molinero MDC, Martín-Valero R. Augmented reality in physical therapy: systematic review and meta-analysis. JMIR Serious Games 2021 Dec 15;9(4):e30985. [doi: 10.2196/30985] [Medline: 34914611]
47. 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]
48. Wiemeyer J. Serious Games für die Gesundheit: Anwendung in der Prävention und Rehabilitation im Überblick (1. In: Springer Nature Switzerland. Switzerland: Springer Nature; 2016.
49. Vaziri DD, Aal K, Gschwind YJ, Delbaere K, Weibert A, Annegarn J, et al. Analysis of effects and usage indicators for a ICT-based fall prevention system in community dwelling older adults. Int J Hum-Comput Stud 2017 Oct;106:10-25. [doi: 10.1016/j.ijhcs.2017.05.004]
50. Stanmore EK, Mavroeidi A, de Jong LD, Skelton DA, Sutton CJ, Benedetto V, et al. The effectiveness and cost-effectiveness of strength and balance exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: a multi-centre, cluster randomised controlled trial. BMC Med 2019 Feb 28;17(1):49 [FREE Full text] [doi: 10.1186/s12916-019-1278-9] [Medline: 30813926]
51. Brox E, Konstantinidis, Evertsen G, Fernandez-Luque L, Rremartinez A, Oesch P. GameUp: exergames for mobility - a project to keep elderly active. 2016 Presented at: Medicon 2016; XIV Mediterranean Conference on Medical and Biological Engineering and Computing; April 2016; Paphos, Cypros. [doi: 10.1007/978-3-319-32703-7_236]
52. 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]
53. 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 20;73(4):794-811. [doi: 10.1111/jan.13160] [Medline: 27688056]
54. Bossen D, Broekema A, Visser B, Brons A, Timmerman A, van Etten-Jamaludin F, et al. Effectiveness of serious games to increase physical activity in children with a chronic disease: systematic review with meta-analysis. J Med Internet Res 2020 Apr 01;22(4):e14549 [FREE Full text] [doi: 10.2196/14549] [Medline: 32234697]
55. Siani A, Marley SA. Impact of the recreational use of virtual reality on physical and mental wellbeing during the Covid-19 lockdown. Health Technol (Berl) 2021 Feb 14;11(2):425-435 [FREE Full text] [doi: 10.1007/s12553-021-00528-8] [Medline: 33614391]
56. Parker K, Uddin R, Ridgers ND, Brown H, Veitch J, Salmon J, et al. The use of digital platforms for adults' and adolescents' physical activity during the COVID-19 pandemic (our life at home): survey study. J Med Internet Res 2021 Feb 01;23(2):e23389 [FREE Full text] [doi: 10.2196/23389] [Medline: 33481759]
57. Corregidor-Sánchez AI, Polonio-López B, Martin-Conty JL, Rodríguez-Hernández M, Mordillo-Mateos L, Schez-Sobrino S, et al. Exergames to prevent the secondary functional deterioration of older adults during hospitalization and isolation periods during the COVID-19 pandemic. Sustainability 2021 Jul 15;13(14):7932. [doi: 10.3390/su13147932]
58. Gao Z, Lee JE, McDonough DJ, Albers C. Virtual reality exercise as a coping strategy for health and wellness promotion in older adults during the COVID-19 pandemic. J Clin Med 2020 Jun 25;9(6):1986 [FREE Full text] [doi: 10.3390/jcm9061986] [Medline: 32630367]
59. Jeon M, Jeon H, Yi C, Kwon O, You S, Park J. Block and random practice: a Wii Fit dynamic balance training in older adults. Res Q Exerc Sport 2021 Sep 13;92(3):352-360. [doi: 10.1080/02701367.2020.1733456] [Medline: 32401683]
60. Wollersheim D, Merkes M, Shields N, Liamputtong P, Wallis L, Raynolds F. Physical and psycosocial effects of Wii video game use among older women. Int J Emerging Technol Soc 2010;8(2):85-98.
61. Son JS, Nimrod G, West ST, Janke MC, Liechty T, Naar JJ. Promoting older adults’ physical activity and social well-being during COVID-19. Leis Sci 2020 Jun 26;43(1-2):287-294. [doi: 10.1080/01490400.2020.1774015]
62. Tinetti ME, Franklin Williams T, Mayewski R. Fall risk index for elderly patients based on number of chronic disabilities. Am J Med 1986 Mar;80(3):429-434. [doi: 10.1016/0002-9343(86)90717-5]

63. Cohen J. Statistical Power Analysis for the Behavioral Sciences (2nd ed.). New York, NY: Routledge; 1988.
64. Bortz J, Schuster C. Statistik für Human- und Sozialwissenschaftler. Switzerland: Springer Nature; 2010.
65. Matzk S, Tsiasioti C, Behrendt S, Jürchott K, Schwinger A. Pflegebedürftigkeit in Deutschland. Pflege-Report 2020:239-277. [doi: 10.1007/978-3-662-61362-7_16]
66. Aung MN, Yuasa M, Koyanagi Y, Aung TNN, Moolphate S, Matsumoto H, et al. Sustainable health promotion for the seniors during COVID-19 outbreak: a lesson from Tokyo. J Infect Dev Ctries 2020 Apr 30;14(4):328-331 [FREE Full text] [doi: 10.3855/jidc.12684] [Medline: 32379708]
67. Caspersen C, Powell K, Christenson G. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131.
68. Kleschnitzki J, Beyer L, Beyer R, Großmann I. Study protocol, limitations and practicability for a representative controlled trial with seniors in care homes. Evaluating the effects of exergames on mental health adherence, cognitive and motor skills. Empirische Evaluationsmethoden 2020.
69. Breuer J, Tolks D. Grenzen von "Serious games for health". Präv Gesundheitsf 2018 Jul 12;13(4):327-332. [doi: 10.1007/s11553-018-0654-1]
70. Baranowski T. Games for health research—past, present, and future. Präv Gesundheitsf 2018 Jul 12;13(4):333-336. [doi: 10.1007/s11553-018-0657-y]

Abbreviations

CG: control group
IG: intervention group
WHO: World Health Organization

Edited by N Zary; submitted 02.02.22; peer-reviewed by S Muraki; comments to author 11.03.22; revised version received 29.03.22; accepted 07.04.22; published 10.05.22.

Please cite as:
Kleschnitzki JM, Grossmann I, Beyer R, Beyer L
Modification in the Motor Skills of Seniors in Care Homes Using Serious Games and the Impact of COVID-19: Field Study
JMIR Serious Games 2022;10(2):e36768
URL: https://games.jmir.org/2022/2/e36768
doi:10.2196/36768
PMID:35536610

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org