JMIR Serious Games

Journal Impact Factor (JIF) (2023): 3.8
Volume 9 (2021), Issue 2 ISSN 2291-9279 Editor-in-Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Viewpoint

Corrigenda and Addendas

Original Paper

Virtual Reality App for Treating Eating Behavior in Eating Disorders: Development and Usability Study

Billy Sundström Langlet[1], PhD; Dorothy Odegi[1], MSc; Modjtaba Zandian[1], PhD; Jenny Nolstam[2]; Per Södersten[1], PhD; Cecilia Bergh[1,2], PhD

[1]Division of Clinical Geriatrics, Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[2]Mandometer Clinic, Stockholm, Sweden

Corresponding Author:
Billy Sundström Langlet, PhD
Division of Clinical Geriatrics, Center for Alzheimer Research
Department of Neurobiology, Care Sciences and Society
Karolinska Institutet
Hälsovägen 7, 3rd floor
Stockholm, 141 52
Sweden
Phone: 46 852488034
Email: billy.langlet@ki.se

Related Article:

This is a corrected version. See correction statement: https://games.jmir.org/2021/2/e29686/

Abstract

Background: Anorexia nervosa is one of the more severe eating disorders, which is characterized by reduced food intake, leading to emaciation and psychological maladjustment. Treatment outcomes are often discouraging, with most interventions displaying a recovery rate below 50%, a dropout rate from 20% to 50%, and a high risk of relapse. Patients with anorexia nervosa often display anxiety and aversive behaviors toward food. Virtual reality has been successful in treating vertigo, anxiety disorder, and posttraumatic stress syndrome, and could potentially be used as an aid in treating eating disorders.

Objective: The aim of this study was to evaluate the feasibility and usability of an immersive virtual reality technology administered through an app for use by patients with eating disorders.

Methods: Twenty-six participants, including 19 eating disorder clinic personnel and 5 information technology personnel, were recruited through emails and personal invitations. Participants handled virtual food and utensils on an app using immersive virtual reality technology comprising a headset and two hand controllers. In the app, the participants learned about the available actions through a tutorial and they were introduced to a food challenge. The challenge consisted of a meal type (meatballs, potatoes, sauce, and lingonberries) that is typically difficult for patients with anorexia nervosa to eat in real life. Participants were instructed, via visual feedback from the app, to eat at a healthy rate, which is also a challenge for patients. Participants rated the feasibility and usability of the app by responding to the mHealth Evidence Reporting and Assessment checklist, the 10-item System Usability Scale, and the 20-point heuristic evaluation questionnaire. A cognitive walkthrough was performed using video recordings of participant interactions in the virtual environment.

Results: The mean age of participants was 37.9 (SD 9.7) years. Half of the participants had previous experience with virtual reality. Answers to the mHealth Evidence Reporting and Assessment checklist suggested that implementation of the app would face minor infrastructural, technological, interoperability, financial, and adoption problems. There was some disagreement on intervention delivery, specifically regarding frequency of use; however, most of the participants agreed that the app should be used at least once per week. The app received a mean score of 73.4 (range 55-90), earning an overall "good" rating. The mean score of single items of the heuristic evaluation questionnaire was 3.6 out of 5. The lowest score (2.6) was given to the "accuracy" item. During the cognitive walkthrough, 32% of the participants displayed difficulty in understanding what to do at the initial selection screen. However, after passing the selection screen, all participants understood how to progress through the tasks.

Conclusions: Participants found the app to be usable and eating disorder personnel were positive regarding its fit with current treatment methods. Along with the food item challenges in the current app, participants considered that the app requires improvement to offer environmental and social (eg, crowded room vs eating alone) challenges.

(JMIR Serious Games 2021;9(2):e24998) doi:10.2196/24998

KEYWORDS

feeding and eating disorders; anorexia nervosa; bulimia nervosa; binge eating disorder; immersive virtual reality; eating disorders; virtual reality

Introduction

Anorexia nervosa is an eating disorder characterized by restriction of energy intake, leading to a significantly low body weight, intense fear of gaining weight, disturbed body perception, and lack of insight into the seriousness of the disorder [1]. At least 90% of individuals with anorexia nervosa are women, with 40% of the identified patients ranging between 15 and 19 years of age [2,3]. According to the Diagnostic and Statistical Manual of Mental Disorders IV, the lifetime prevalence of anorexia nervosa varies from 0.3% to 1.0%, and the mean crude mortality rate is 5% per decade [4,5], varying from 0% to 15.6% [6,7]. Anorexia nervosa is associated with physical problems such as anemia, reduced brain volume, infertility, altered hormonal balance, loss of muscle mass, and osteoporosis. However, most of these physical and mental problems normalize with weight regain [8].

Long-term outcomes in anorexia nervosa treatments are often discouraging. A 2002 review concluded that only 46.9% of anorexia nervosa patients reached full recovery, 33.5% improved, and the disorder became chronic in 20.8% of cases [9]. In 2020, similar outcomes were found in anorexia nervosa patients receiving treatments consisting of eating disorder–focused structured individual therapies [10]. In addition, studies frequently report poor long-term treatment outcomes with high relapse rates [11], dropout rates ranging from 20.2% to 49.6% [12], and a propensity for the anorexia nervosa disorder to transition into other eating disorders [13]. Better treatment outcomes could be achieved when one part of the treatment consists of normalizing eating behavior through eat training [14]. Treatment based on this method typically begins after establishing a baseline eating behavior for each patient by having them eat food on a medical device (Mandometer) that records the cumulative food intake (in grams). In this treatment, patients eat their meals on the Mandometer following a reference curve for food intake over time that is displayed on their own smartphone. The food intake quantity is initially based on the baseline measure but is updated depending on the rate of recovery of the patient [15]. Another part of eat training consists of reducing food avoidance and apprehension around food, which are the key factors responsible for maintaining the starved state [16]. Virtual reality (VR) technology could offer an alternative approach for handling food and to practice eat training in a treatment setting [17].

VR treatment involves immersing an individual in a computer-generated 3D world that is customized according to treatment needs, where the individual can be safely exposed to stressors. Additional benefits of using VR are that it enables repetition as well as exposure control (internal validity) and has high generalizability to other contexts (external validity). Recent technological advances have greatly reduced the cost of VR technology, thereby increasing its scalability [18]. To date, intervention studies that employed VR have been successful in treating posttraumatic stress syndrome, anxiety syndrome, and smoking [19,20]. In the eating disorder context, most studies that have employed VR used the technology as an assessment tool for exploring body image perception, and the response of virtual foods and environments [21]. One study demonstrated the ecological validity of the approach, with VR food eliciting similar responses as real food in patients with eating disorders [22]. Most of the eating disorder interventions employing VR have aimed to correct a distorted body image [23]. However, VR exposure therapy has also been reported to reduce the anxiety response to food in patients with bulimia nervosa [24]. VR therapy may also be more acceptable to eating disorder patients than other treatment forms. In a study on phobias, VR exposure was more likely to be selected (76% of participants chose VR) and had lower refusal rates than in vivo exposure (3% vs 27%, respectively) [25]. These findings suggest that exposure therapy via VR has internal, external, and ecological validity in the target group, indicating that it can be effective in treating bulimia nervosa and perhaps also anorexia nervosa, and may be more acceptable for end users (patients with eating disorders) [25].

Evaluation of interventions is common; however, the quality of these studies is often poor, which leads to incorrect implementations and findings that are not reproducible [26]. The first step of establishing digital health interventions is usually to gather the functional requirements for development and testing [27]. The mHealth Evidence Reporting and Assessment (mERA) checklist is a useful method for reporting on digital interventions, which has the benefit of providing recommendations for reporting the feasibility of intervention strategies [27]. Another key factor for the successful adoption of a new technology is usability [28], which was defined as “the extent to which a product can be used to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” [29]. To ensure the acceptance and attitude of patients and clinicians, as well as to enhance well-being, reduce risk of harm, and increase accessibility for patients, health technologies should be appropriately designed to the end users’ needs before they are deployed as health interventions [30,31]. One review suggests using multiple methods in performing usability evaluations [29]. Another study found that the System Usability Scale (SUS) was the most common questionnaire used in such evaluations, and recommended the use of quantitative evaluation methods but concluded that further research is needed to identify which methods are best suited for different patient groups [28].

The aim of this study was to investigate the perceived usability of a newly developed VR app that simulates eat training and is intended for use in an eating disorder intervention study. As a first step in this development, usability was evaluated by staff involved in eating disorder treatment.

Methods

Participants

To be eligible for participation in the study, the individual had to be working in an eating disorder clinic and have daily interactions with eating disorder patients as a clinician or physician (eating disorder personnel) or as part of information technology (IT) service (IT personnel). Individuals were invited to participate in the study through emails and personal invitations. Eating disorder patients were not approached at this stage because of the potential risk that participation would result in poorer treatment outcomes and the assessment that many of the questions related to intervention delivery would be difficult for patients to answer.

Technology

The HTC VIVE VR system (HTC) was used as the immersive VR technology in this study, consisting of a headset (connected to a computer) through which the VR environment could be viewed, two hand controllers that enabled interaction with the VR environment, and two base stations that enabled motion tracking. To ensure proper performance, the room size should be at least 1.5×2.0 meters. The computer used a NVIDIA GeForce GTX 1060 graphic card, an Intel Core i5-4590 processor, 8 GB of RAM, and the Windows 7 operating system. The software was developed in the Unity engine and ran on the digital game distribution platform Steam (Valve Corporation), using the SteamVR app.

Instruments

Four methods were used for the usability evaluation: the mERA checklist, 10-Item SUS, 20-item heuristic evaluation questionnaire, and a cognitive walkthrough. Rather than being forced to respond to each question, participants could leave questions blank if unable to respond.

The mERA checklist is a 16-item checklist that aims to standardize reporting, and enables quick assessment of eHealth and mobile health apps. Adoption of this checklist is meant to highlight issues of generalizability and rigor in reporting and improving transparency [27]. The procedure and app allowed the researchers to objectively evaluate the infrastructure and technological platform, and to make a cost assessment (items 1, 2, and 9). A one-on-one interview was held with each participant regarding the items that could not be objectively evaluated (items 3, 4, 5, 7, 8, 10, 11, and 12). The specific question for each item was whether the participant could foresee any obstacles related to the item (eg, how easy it would be to integrate the app and VR system with already existing health care systems) if the app and VR system were to be used in an identical fashion to how eat training is currently performed in a clinical setting. Questions on interoperability and adaptability were answered by all participants (items 3 and 12), whereas questions on intervention delivery, content testing, accessibility, adoption, scalability, and user feedback were answered only by the eating disorder personnel (items 4, 5, 7, 8, 10, and 11). Since the app was at a formative stage, five questions on the checklist were not applicable (items 6, 13, 14, 15, and 16).

The SUS is a validated tool that is widely used to assess the perceived usability of a system. It consists of 10 statements such as "I found the system unnecessarily complex" and "I felt very confident using the system," which are answered on a 5-point Likert scale with verbal anchors at the extremes from 0 ("strongly disagree") to 4 ("strongly agree"). The sum result of the SUS is a score ranging from 0 to 100 [32].

The cognitive walkthrough is a usability evaluation method in which evaluators work through a series of tasks required by the system asking a set of questions from the perspective of the user [33]. The purpose of the cognitive walkthrough is to evaluate the ease of use of an interface design to new or infrequent users. In this study, the following four steps of human-computer interaction were evaluated: (1) goals to be completed in the system, (2) determination of currently available actions, (3) selection of actions to be taken, and (4) performance of the tasks and evaluation of the feedback given by the system. We determined whether a participant managed to perform a step for each task by reviewing video recordings of their interaction with the app (eg, Task 3. Read instructions, grab spoon, pick up meatballs, place meatballs on plate, and move to next task; see Multimedia Appendix 1). Video recordings of actions performed by the user in the app were made using the inherent Steam streaming software. The video recordings were complemented with audio recordings on a smartphone.

The heuristic evaluation is a method for evaluating the usability of computer software, focusing on identifying problems with the user interface. As the name implies, this questionnaire evaluates recognized usability principles (heuristics), constituting one of the more informal methods of human-computer interaction. The heuristic evaluation questionnaire used in this study was based on the Weinschenk and Barker [34] classification, which consist of 20 items, for example "User support: does the application provide additional assistance as needed or requested?" Each item was answered on a 5-point Likert scale with verbal anchors at the extremes from 1 ("strongly disagree") to 5 ("strongly agree").

Procedure

Each participant attended an information meeting, including a VR tutorial, VR eat training, and a usability evaluation (Figure 1). They were informed of the study and encouraged to ask questions, after which they signed a written consent form if they agreed to participate.

The participants were then taken to the VR lab to familiarize themselves with the VR equipment and environment (Figure 2). The VR lab tutorial started with calibration of the app so that a virtual table and chair were in the same position as a real table and chair in the VR lab. The participant was seated on the real chair in front of the real table and was then fitted with the

VR equipment. At this point, video and voice recording features were initiated. In the VR environment, the controllers were represented by hands. The tutorial was built in steps, with each step being supported by instructions on what tasks to perform to meet the goal to proceed to the next step, which were displayed on a virtual smart tablet present in the VR environment.

The success of each task was evaluated based on the performance of the participants (Figure 3). In the first step, participants used the controller to interact with the “next” button on the virtual smart tablet. Each subsequent step ended by pressing the “next” button. In the second step, a plate appeared on the table and the participant placed the plate at a specific position on the table. In the third step, a pan of meatballs and a spoon appeared; participants had to then transfer the meatballs from the pan to the plate using the spoon. In the fourth step, a pot of potatoes and a fork appeared; participants had to transfer the potatoes from the pot to the plate using the fork. In the fifth step, a sauce boat filled with gravy appeared; participants had to pour gravy from the sauce boat over the plate. In the sixth step, a bowl of lingonberry jam and a spoon appeared; participants had to transfer the lingonberry jam from the bowl to the plate using the spoon. In the seventh step, a jug of water and a glass appeared; participants had to pour water from the jug into the glass. In the final step of the tutorial, participants interacted freely with the served food, cutting and eating it.

Once familiarized with the app, the participant experienced VR eat training in the same manner as eat training with real food is practiced in the clinic. First, the participant was asked to place a healthy portion of food on the plate. Feedback on how close that portion was to a healthy portion was presented on the computer screen (expressed as percentages). Once the screen showed 100%, the participant started to eat. At this point, the training curve for food intake was displayed on the screen, and as the virtual meal progressed, the virtual food intake also emerged on the screen. Similar to treatment of real eating behavior, the participant tried to eat virtually following the training curve [15]. Once all of the virtual food had been consumed, the duration of the meal, amount of food eaten, and rate of eating were presented on the screen, and the participant was given the option to close the app.

After having experienced both the tutorial and VR eat training session, participants sat down with one researcher to answer the SUS, heuristic evaluation, and mERA checklist questionnaires. Initially, the participants were expected to answer the questions alone. However, it became immediately obvious that some questions required clarification and affirmation by the researcher to confirm proper interpretation. Therefore, a researcher was present in the room when the participant filled in the questionnaires to provide help when needed. After completing the questionnaires, participants received a cinema ticket as a reward and they were thanked for their participation.

Figure 1. Study protocol presented in chronological order from left to right. VR: virtual reality; mERA: mHealth Evidence Reporting and Assessment.

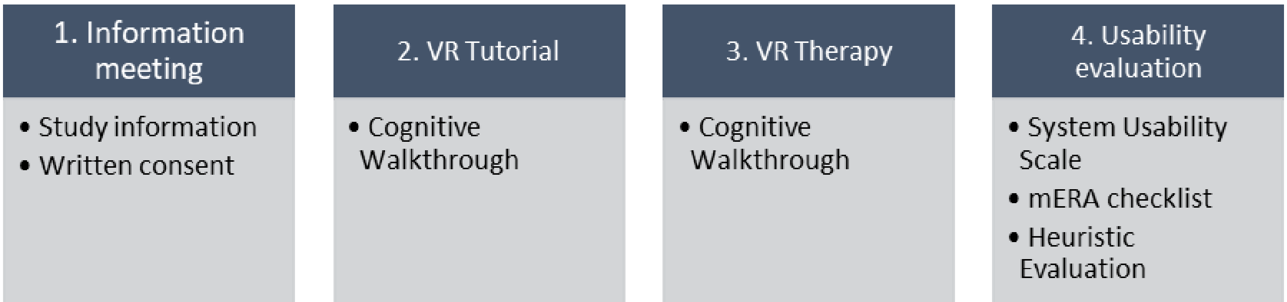


Figure 2. Participant interacting with virtual food on the plate (left) and participant eating virtual food (right).

Figure 3. Free interaction with food during the tutorial step of the app.

Data Interpretation

Results are expressed as percentage among responding participants. To provide context for the SUS score, the adjective rating scale proposed by Bangor et al [35] was used, where a score of 39.17 to 52.00 is considered "poor," 52.01 to 72.74 is considered "OK," 72.75 to 85.57 is considered "good," and 85.58 to 100.00 is considered "excellent."

Ethical Approval

All procedures were approved by the Swedish Ethical Review Authority (dnr 2019-04249) and followed the Helsinki Declaration. Each participant attended an information meeting and provided written consent for participation.

Results

Participants

A total of 24 participants (67% women) were recruited, including 19 eating disorder personnel and 5 IT personnel. Their mean age was 37.9 (SD 9.7) years. Eating disorder personnel had worked at an eating disorder clinic for 4.1 (SD 5.7) years. Twelve of the 24 participants (50%) had previously experienced VR, including 42% (8/19) of the clinicians and 80% (4/5) of the IT participants.

IT personnel were only able to answer 20% of the mERA questions intended for them, and these responses were therefore excluded from the analysis of mERA results. The reason given by IT personnel for not being able to answer the questions was that they are not directly involved in the clinical treatment aspects. Eating disorder personnel were able to answer 80% of the mERA questions intended for them and these responses were therefore included in the mERA questionnaire analysis.

mERA Checklist

Item 1: Infrastructure

In Swedish urban environments, the availability of power supplies and network connections is adequate to support the app; however, it is uncertain whether the infrastructure in rural environments would be sufficient. The size of the room required for the equipment makes it possible to implement in any facility. However, when the space is smaller than 1.5×2.0 meters, there is a risk of signals from the controller and headset not being transferred to the base stations, causing the app to pause.

Item 2: Technology Platform

The app performed well on a computer with low hardware specifications (see the Technology subsection of the Methods), but hardware requirements may be higher if other apps are expected to run simultaneously.

Item 3: Interoperability

All of the responding eating disorder personnel (100%, 10/10) thought that it would be easy to integrate the app with the clinic's medical record system, but would be more difficult to integrate with the national care information system. The main problem mentioned in relation to the national system is that visual presentations are difficult to implement, but simplified presentation of data would be possible.

Item 4: Intervention Delivery

All participants thought that during periods of app use, the frequency should be at least once a week. Participants thought that the current app should be used during treatment, with a higher frequency at the start of treatment. They considered that the addition of difficult environmental and social situations (challenges) would make it usable at a similar frequency throughout treatment.

Item 5: Content Testing

All eating disorder personnel thought that the app should complement real eat training using VR to expose patients to food, environment, and social challenges. The addition of challenges would require alterations to the app.

Item 7: User Feedback

All responding eating disorder personnel were pleased with the state of the app (100%, 14/14) and thought that future interventions would be easy due to the similarities of the app with current clinical treatment protocols.

Item 8: Access of Individual Participants

Nine out of 19 (47%) eating disorder personnel thought that the app was usable for the treatment of all eating disorders; the remainder were unsure but stated no specific barriers to use. Other potential barriers identified were age, culture (availability of only one meal), language (only an English version is currently available), and epilepsy. Overall, 16% (3/19) of the eating disorder personnel expressed a concern that the patients may replace real meals with virtual meals, thereby complicating treatment.

Item 9: Cost Assessment

The currently used setup (computer and VR set) costs approximately US $1600, which is affordable for both research and clinical purposes. Based on an intervention protocol of 1 hour per week for each patient, a clinic should be able to run an intervention using two VR devices, which would cost approximately US $3200. There is also the added cost of renting space and a part time clinician.

Item 10: Adoption Inputs

Eating disorder personnel thought that the app should be introduced early in treatment and be described as an aid to patients, similar to the medical device (Mandometer) currently used by the clinic, which measures food weight and provides feedback on eating behavior.

Item 11: Limitations for Delivery at Scale

Regarding limitations at scale, 32% (6/19) of the participants thought that providing technical support for installing, updating, and using the system could cause problems. Moreover, 21% (4/19) of the participants thought that implementation for clinics using other treatment forms could be problematic. Few mentioned costs (11%, 2/19) and time requirements (5%, 1/19) as potential problems for scaling up.

Item 12: Contextual Adaptation

Eating disorder personnel agreed that the most important aspect of the app should be eat training and exposure to food, but commented that including aspects of the environment (eg, school dining hall and birthday party) and social challenges (eg, eating with other people, eating when people are talking) could further improve the app. Other suggestions included reducing anxiety by providing relaxing environments. There was also a suggestion to add a reward system, where patients could score points if they managed to succeed in various challenges.

SUS Questionnaire

The SUS scores were similar between clinicians and IT participants, with a mean score of 73.4 (SD 9.2) and 73.5 (SD 11.3), respectively (for a total mean score of 73.4, range 55-90), earning an overall “good” rating. Although the sample sizes did not allow for a statistical comparison, the largest discrepancy between single items was found for the question “I think that I would need assistance to be able to use this system,” for which clinicians were less confident that patients would be able to use the app without support (1.3, mean difference –0.9). The highest score was given to the item “I would imagine that most people would learn to use this system very quickly,” with a mean score of 3.5 (maximum score 4).

Heuristic Evaluation

The mean score of single items of the heuristic evaluation questionnaire was 3.6 out of 5, and was very similar between eating disorder and IT personnel (3.6 and 3.7, respectively). Mean values of specific items ranged from 2.6 to 4.4. Both eating disorder and IT personnel provided low scores on “accuracy,” with a mean score of 2.6 (SD 1.1). Eating disorder personnel also provided a low score on “user support”, with a mean of 2.9 (SD 1.3), whereas the IT personnel provided low scores on “flexibility,” with a mean of 2.0 (SD 0.7).

Cognitive Walkthrough

The only task in which participants faced problems in understanding the goals to be completed in the app was at the initial selection screen (step 1), where 29% (7/24) of the participants needed instructions on how to select between the tutorial and eat training part of the app. All participants were able to determine the available actions (step 2) and select which actions to take (step 3). All participants required further instructions at least once when trying to perform the tasks (step 4). Based on the instructions required for the participant to complete the task, six problems were identified: (1) handling utensils that used “magnetic” properties (ie, where the food latched onto the utensil based on proximity) rather than gravity to interact with food (serving potatoes and lingonberries) was the most challenging issue; (2) the utensils had specific attributes tied to them (ie, the fork was used for picking up food and the knife was used for cutting), which created problems for participants who tried to pick things up with their hands in some cases (controllers not holding utensils) or used both utensils for picking up food; (3) it was not possible to divide the meatballs and the potatoes more than twice, which confused some participants; (4) only one food item (eg, potatoes or meatballs) could be placed on the fork at a time, and most participants tried to put both on the fork together initially; (5) when drinking, there was a risk of hitting the VR headset with the controller because the water glass was held closer to the controller compared with the fork holding food; and (6) when the VR environment (SteamVR) was improperly calibrated, it became difficult to interact with the items.

Discussion

Using the app tested in this study in a clinical setting would likely only face minor infrastructural, technological, interoperability, financial, or adoption problems. Eating disorder personnel seemed to be positive at the prospect of using the app in the clinic. However, there was some disagreement on the protocol of intervention delivery. SUS scores suggested that the system is passable with room for improvement. The heuristics of the user interface was acceptable, identifying user support, accuracy, and flexibility as potential weaknesses. Video and audio recordings of users' interaction with the app suggest that users knew the goal of each step of the app and understood when they had successfully completed each step. A suggestion for improvement of the app was to add environmental and social stressors, and to use them in a similar manner as the food stressors in the current app.

Despite its early stage of development, compared with other systems, the app performed above average on the SUS [35]. These findings are in line with answers on the mERA checklist, heuristic evaluation, and observations from the cognitive walkthrough. The results also indicate a willingness by clinicians to use the app in treatment, which is important to ensure proper intervention fidelity [36]. Regarding the heuristic evaluation, the lowest scores were given to "accuracy," "flexibility," and "user support." The low "accuracy" scores seem to have stemmed from difficulty in using the fork. This may have been caused by different types of cutlery having different functions associated with them. In practice, this meant that interactions with the fork worked on proximity, similar to a magnet, whereas interaction with the spoon was based on the regular gravitation properties of objects. Both methods of interaction are often used in VR apps, but mixing the two likely resulted in reduced accuracy. The low flexibility score indicates a need for user customization. The next version of the app should therefore include visual and audio information in both Swedish and English, as well as a more varied selection of cutlery (eg, spoon and fork) and dishes. The low score on user support suggests the requirement of technical support during the intervention. One problem addressed by eating disorder personnel is that there does not seem to be a safe way for individuals with photosensitive epilepsy to use VR equipment. However, this group only accounts for approximately 10% of epilepsy cases in the age range of 7-19 years [37]. Another worry of eating disorder personnel was that patients would replace real meals with virtual meals. Addressing this concern is beyond the scope of this study but should be considered when conducting VR interventions for eating disorder patients.

Responses to the mERA checklist suggested that infrastructural requirements were low using the current VR technology (HTC VIVE). However, some clinics could face issues, especially if rooms serve multiple purposes, which would require the system to be mobile. One solution would be to use more mobile alternatives such as Oculus Quest, which do not require base stations or a computer connection. The low hardware requirements also suggest that most clinics should be able to use the app with their current computers. Even if new computers are purchased for the intervention, the estimated cost of a VR study using current VR technology is low. Given the similarity of the data provided by the app to data already handled by the medical records system of the intended clinics, only minor modifications are required to incorporate the app in treatment. The reason that IT personnel rated the requirement for assistance lower than the eating disorder personnel may be because they were more likely to have used VR before. In the intended intervention, assistance requirements will not be a problem since patients will be assisted by eating disorder personnel. However, if future interventions intend for the app to be used unassisted or outside of the clinic, additional assistance provided by the app is likely required. Despite the concerns that users may need assistance when using the app, the highest usability score was given to the ease of learning how to use the system.

According to participant feedback, the app should be introduced at the beginning of treatment (intervention) and be used in parallel with real-life eat training. The app should be used throughout the treatment course, with a focus on food challenges at the beginning, and environment and social challenges introduced at the end. VR sessions should be administered at least once a week, for at least the duration of a single meal (around 12 minutes), but preferably for a few meal scenarios. In addition, each clinic should have one person on call for technical support when the VR sessions are performed. The proposed intervention protocol has a slightly lower frequency (once per week vs twice) but a longer duration (3 weeks vs treatment duration) than a study on binge eating performed by one of the more prominent groups focusing on VR-based treatment [38]. However, this protocol is similar to the current treatment [15], which eating disorder personnel thought should make implementation easy.

A strength of this study was the high number of eating disorder personnel included, which increases confidence in generalizability of the findings. However, it should be noted that the personnel all had experience in using a similar treatment protocol; thus, clinicians administering other treatment protocols may respond differently. A potential weakness was the use of eating disorder personnel rather than patients with eating disorders. However, many of the questions related to intervention delivery at this early stage of development would be difficult for patients to answer. This, along with the worry that patients may be affected negatively by the app, was why we recruited only eating disorder and IT personnel for this study. After ensuring that the VR technology and the app are safe for eating disorder patients to use, more mature versions of the app should be evaluated by the intended user group. Another potential weakness was that the heuristic evaluation was translated from English to Swedish but has not been validated in the translated language. Owing to the early stage of development, most mERA items could not be objectively evaluated, but instead an opinion was sought from participants.

This is not the intended use of the mERA checklist, which could lead to reduced reproducibility.

Future studies should aim to evaluate usability in eating disorder patients, investigate if VR is effective in changing behavior, and implement the app in an intervention. To ensure scalability to other clinics, especially those less experienced with handling electronic devices as a tool to help treat eating disorders, a clear protocol should be established, similar to that currently existing for handling a medical device (Mandometer) in eating disorder clinics [15]. Due to the immature state of VR as a treatment method, more general studies should also be performed to evaluate how various elements such as modes of information transfer (ie, tactile, visual, and auditory) influence usability and compliance.

The app was found to be usable and eating disorder personnel were positive regarding its incorporation in treatment. The app would fit well with current treatment, requiring only minor alterations. The overall consensus was that the app should be introduced at the beginning of treatment, be administered in parallel with real-life training, technical support should be available, and initial app use should be restricted to the clinic. Along with the food type dimension, there were requests by eating disorder personnel to allow changes to the environment and social context (eg, crowded room vs eating alone).

Acknowledgments

The material and methods used and the collection of data used in this study were financially supported by the Mandometer clinic. The funding organization had no role in designing the study, collecting and analyzing the data, or in preparing this manuscript.

Conflicts of Interest

PS and CB own shares in Mandometer AB, the company that financed the development of the VR app. JN was an employee of Mandometer AB. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Tasks performed in the app by participants.
[PDF File (Adobe PDF File), 909 KB - games_v9i2e24998_app1.pdf]

References

1. American Psychiatric Association. Diagnostic Statistical Manual of Mental Disorders, 5th Edition. Washington, DC: American Psychiatric Association Publishing; Jul 01, 2013.
2. Hoek HW, van Hoeken D. Review of the prevalence and incidence of eating disorders. Int J Eat Disord 2003 Dec;34(4):383-396. [doi: 10.1002/eat.10222] [Medline: 14566926]
3. Udo T, Grilo CM. Prevalence and correlates of DSM-5-defined eating disorders in a nationally representative sample of U.S. adults. Biol Psychiatry 2018 Sep 01;84(5):345-354 [FREE Full text] [doi: 10.1016/j.biopsych.2018.03.014] [Medline: 29859631]
4. Arcelus J, Mitchell AJ, Wales J, Nielsen S. Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. Arch Gen Psychiatry 2011 Jul;68(7):724-731. [doi: 10.1001/archgenpsychiatry.2011.74] [Medline: 21727255]
5. Smink FRE, van Hoeken D, Hoek HW. Epidemiology of eating disorders: incidence, prevalence and mortality rates. Curr Psychiatry Rep 2012 Aug;14(4):406-414 [FREE Full text] [doi: 10.1007/s11920-012-0282-y] [Medline: 22644309]
6. Zipfel S, Löwe B, Reas DL, Deter HC, Herzog W. Long-term prognosis in anorexia nervosa: lessons from a 21-year follow-up study. Lancet 2000 Mar 26;355(9205):721-722. [doi: 10.1016/S0140-6736(99)05363-5] [Medline: 10703806]
7. Dobrescu SR, Dinkler L, Gillberg C, Råstam M, Gillberg C, Wentz E. Anorexia nervosa: 30-year outcome. Br J Psychiatry 2020 Feb;216(2):97-104 [FREE Full text] [doi: 10.1192/bjp.2019.113] [Medline: 31113504]
8. Westmoreland P, Krantz MJ, Mehler PS. Medical complications of anorexia nervosa and bulimia. Am J Med 2016 Jan;129(1):30-37. [doi: 10.1016/j.amjmed.2015.06.031] [Medline: 26169883]
9. Steinhausen H. The outcome of anorexia nervosa in the 20th century. Am J Psychiatry 2002 Aug;159(8):1284-1293. [doi: 10.1176/appi.ajp.159.8.1284] [Medline: 12153817]
10. Treasure J, Duarte TA, Schmidt U. Eating disorders. Lancet 2020 Mar 14;395(10227):899-911. [doi: 10.1016/S0140-6736(20)30059-3] [Medline: 32171414]
11. Södersten P, Bergh C, Leon M, Brodin U, Zandian M. Cognitive behavior therapy for eating disorders versus normalization of eating behavior. Physiol Behav 2017 May 15;174:178-190 [FREE Full text] [doi: 10.1016/j.physbeh.2017.03.016] [Medline: 28322911]
12. Wallier J, Vibert S, Berthoz S, Huas C, Hubert T, Godart N. Dropout from inpatient treatment for anorexia nervosa: critical review of the literature. Int J Eat Disord 2009 Dec;42(7):636-647. [doi: 10.1002/eat.20609] [Medline: 19208386]
13. Fichter MM, Quadflieg N, Crosby RD, Koch S. Long-term outcome of anorexia nervosa: Results from a large clinical longitudinal study. Int J Eat Disord 2017 Sep;50(9):1018-1030. [doi: 10.1002/eat.22736] [Medline: 28644530]

14. Bergh C, Callmar M, Danemar S, Hölcke M, Isberg S, Leon M, et al. Effective treatment of eating disorders: results at multiple sites. Behav Neurosci 2013 Dec;127(6):878-889. [doi: 10.1037/a0034921] [Medline: 24341712]
15. Esfandiari M, Papapanagiotou V, Diou C, Zandian M, Nolstam J, Södersten P, et al. Control of eating behavior using a novel feedback system. J Vis Exp 2018 May 08(135):57432 [FREE Full text] [doi: 10.3791/57432] [Medline: 29806832]
16. Murray SB, Strigo IA. Anorexia nervosa, neuroimaging research, and the contextual salience of food cues: The food approach-avoidance conundrum. Int J Eat Disord 2018 Aug;51(8):822-825 [FREE Full text] [doi: 10.1002/eat.22883] [Medline: 29846004]
17. Clus D, Larsen ME, Lemey C, Berrouiguet S. The use of virtual reality in patients with eating disorders: systematic review. J Med Internet Res 2018 Apr 27;20(4):e157 [FREE Full text] [doi: 10.2196/jmir.7898] [Medline: 29703715]
18. Ferrer-García M, Gutiérrez-Maldonado J. The use of virtual reality in the study, assessment, and treatment of body image in eating disorders and nonclinical samples: a review of the literature. Body Image 2012 Jan;9(1):1-11. [doi: 10.1016/j.bodyim.2011.10.001] [Medline: 22119329]
19. 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 [FREE Full text] [doi: 10.1097/HRP.0000000000000138] [Medline: 28475502]
20. Gorini A, Riva G. Virtual reality in anxiety disorders: the past and the future. Expert Rev Neurother 2008 Mar;8(2):215-233. [doi: 10.1586/14737175.8.2.215] [Medline: 18271709]
21. de Carvalho MR, Dias TR, Duchesne M, Nardi AE, Appolinario JC. Virtual reality as a promising strategy in the assessment and treatment of bulimia nervosa and binge eating disorder: a systematic review. Behav Sci (Basel) 2017 Jul 09;7(3):43 [FREE Full text] [doi: 10.3390/bs7030043] [Medline: 28698483]
22. Gorini A, Griez E, Petrova A, Riva G. Assessment of the emotional responses produced by exposure to real food, virtual food and photographs of food in patients affected by eating disorders. Ann Gen Psychiatry 2010 Jul 05;9:30 [FREE Full text] [doi: 10.1186/1744-859X-9-30] [Medline: 20602749]
23. Ferrer-García M, Gutiérrez-Maldonado J. The use of virtual reality in the treatment of eating disorders. Stud Health Technol Inform 2012;181:17-21. [Medline: 22954820]
24. Toro J, Cervera M, Feliu MH, Garriga N, Jou M, Martinez E, et al. Cue exposure in the treatment of resistant bulimia nervosa. Int J Eat Disord 2003 Oct;34(2):227-234. [doi: 10.1002/eat.10186] [Medline: 12898559]
25. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007 Oct;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
26. 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]
27. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
28. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
29. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Mar;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
30. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
31. Huryk LA. Factors influencing nurses' attitudes towards healthcare information technology. J Nurs Manag 2010 Jul;18(5):606-612. [doi: 10.1111/j.1365-2834.2010.01084.x] [Medline: 20636510]
32. Jordan PW, Thomas B, Weerdmeester BA, McClelland IL. SUS - a quick and dirty usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
33. Andre TS, Hartson HR, Williges RC. Determining the effectiveness of the usability problem inspector: a theory-based model and tool for finding usability problems. Hum Factors 2003;45(3):455-482. [doi: 10.1518/hfes.45.3.455.27255] [Medline: 14702996]
34. Weinschenk D, Barker DT. Designing effective speech interfaces. Hoboken, NJ: John Wiley & Sons; 2000.
35. 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]
36. Moncher FJ, Prinz RJ. Treatment fidelity in outcome studies. Clin Psych Rev 1991 Jan;11(3):247-266. [doi: 10.1016/0272-7358(91)90103-2]
37. Quirk J, Fish DR, Smith SJ, Sander JW, Shorvon SD, Allen PJ. Incidence of photosensitive epilepsy: a prospective national study. Electroencephalogr Clin Neurophysiol 1995 Oct;95(4):260-267. [doi: 10.1016/0013-4694(95)00118-i] [Medline: 8529557]

38. Ferrer-García M, Gutiérrez-Maldonado J, Pla-Sanjuanelo J, Vilalta-Abella F, Riva G, Clerici M, et al. A randomised controlled comparison of second-level treatment approaches for treatment-resistant adults with bulimia nervosa and binge eating disorder: assessing the benefits of virtual reality cue exposure therapy. Eur Eat Disord Rev 2017 Nov;25(6):479-490. [doi: 10.1002/erv.2538] [Medline: 28804985]

Abbreviations

IT: information technology
mERA: mHealth Evidence and Reporting Assessment
SUS: System Usability Scale
VR: virtual reality

Edited by N Zary; submitted 18.10.20; peer-reviewed by R Murphy, CA Almenara; comments to author 08.01.21; revised version received 15.02.21; accepted 13.03.21; published 13.04.21.

Please cite as:
Langlet BS, Odegi D, Zandian M, Nolstam J, Södersten P, Bergh C
Virtual Reality App for Treating Eating Behavior in Eating Disorders: Development and Usability Study
JMIR Serious Games 2021;9(2):e24998
URL: https://games.jmir.org/2021/2/e24998
doi:10.2196/24998
PMID:33847593

Original Paper

Design and Evaluation of User-Centered Exergames for Patients With Multiple Sclerosis: Multilevel Usability and Feasibility Studies

Alexandra Schättin[1], PhD; Stephan Häfliger[1], MSc; Alain Meyer[1], MSc; Barbara Früh[1], MSc; Sonja Böckler[2], MA; Yannic Hungerbühler[2], BA; Eling D de Bruin[1,3], PhD; Sebastian Frese[4], PhD; Regula Steinlin Egli[5], BSc; Ulrich Götz[2], Dipl; René Bauer[2], Lic Phil; Anna Lisa Martin-Niedecken[2], PhD

[1]Department of Health Sciences and Technology, Institute of Human Movement Sciences and Sport, ETH Zurich, Zurich, Switzerland

[2]Department of Design, Subject Area in Game Design, Zurich University of the Arts, Zurich, Switzerland

[3]Division of Physiotherapy, Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden

[4]Technology and Innovation Unit and Department of Research, ZURZACH Care, Bad Zurzach, Switzerland

[5]Physiotherapy Langmatten, Binningen, Switzerland

Corresponding Author:
Alexandra Schättin, PhD
Department of Health Sciences and Technology, Institute of Human Movement Sciences and Sport
ETH Zurich
IBWS, HCP H, Leopold Ruzicka Weg 4
Zurich, 8093
Switzerland
Phone: 41 44 633 90 57
Email: alexandra.schaettin@hest.ethz.ch

Related Article:

This is a corrected version. See correction statement: https://games.jmir.org/2021/2/e30326/

Abstract

Background: Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system. Patients with MS experience a wide range of physical and cognitive dysfunctions that affect their quality of life. A promising training approach that concurrently trains physical and cognitive functions is video game–based physical exercising (ie, exergaming). Previous studies have indicated that exergames have positive effects on balance and cognitive functions in patients with MS. However, there is still a need for specific, user-centered exergames that function as a motivating and effective therapy tool for patients with MS and studies investigating their usability and feasibility.

Objective: The aim of this interdisciplinary research project is to develop usable and feasible user-centered exergames for the pressure-sensitive plate Dividat Senso by incorporating theoretical backgrounds from movement sciences, neuropsychology, and game research as well as participatory design processes.

Methods: Focus groups (patients and therapists) were set up to define the user-centered design process. This was followed by the field testing of newly developed exergame concepts. Two sequential usability and feasibility studies were conducted on patients with MS. The first study included a single exergaming session followed by measurements. Between the first and second studies, prototypes were iterated based on the findings. The second study ran for 4 weeks (1-2 trainings per week), and measurements were taken before and after the intervention. For each study, participants answered the System Usability Scale (SUS; 10 items; 5-point Likert Scale; score range 0-100) and interview questions. In the second study, participants answered game experience–related questionnaires (Flow Short Scale [FSS]: 13 items; 7-point Likert Scale; score range 1-7; Game Flow questionnaire: 17 items; 6-point Likert Scale; score range 1-6). Mixed methods were used to analyze the quantitative and qualitative data.

Results: In the first study (N=16), usability was acceptable, with a median SUS score of 71.3 (IQR 58.8-80.0). In the second study (N=25), the median SUS scores were 89.7 (IQR 78.8-95.0; before) and 82.5 (IQR 77.5-90.0; after), and thus, a significant decrease was observed after training (z=−2.077; P=.04; *r*=0.42). Moreover, high values were observed for the overall FSS (pre: median 5.9, IQR 4.6-6.4; post: median 5.8, IQR 5.4-6.2) and overall Game Flow Questionnaire (pre: median 5.0, IQR 4.7-5.3; post: median 5.1, IQR 4.9-5.3). A significant decrease was observed in the item *perceived importance* (FSS: z=−2.118; P=.03;

r=0.42). Interviews revealed that user-centered exergames were usable, well accepted, and enjoyable. Points of reference were identified for future research and development.

Conclusions: The project revealed that the newly developed, user-centered exergames were usable and feasible for patients with MS. Furthermore, exergame elements should be considered in the development phase of user-centered exergames (for patients with MS). Future studies are needed to provide indications about the efficacy of user-centered exergames for patients with MS.

(JMIR Serious Games 2021;9(2):e22826) doi:10.2196/22826

KEYWORDS

multiple sclerosis; exergame; motor; physical; cognition; usability; feasibility

Introduction

Background

Globally, approximately 2.3 million people have multiple sclerosis (MS) [1]. MS is an immune-mediated chronic inflammatory disease in which focal inflammation causes the degradation of myelin in the nerve fibers of the central nervous system (CNS), resulting in a wide range of symptoms and impairments [2-4]. Depending on the affected CNS regions and the degree of severity, patients with MS can have physical disabilities (eg, motor weakness, spasticity, sensory disturbances, ataxia, and visual loss), cognitive dysfunction (eg, information processing, attention, executive functions, and memory), and fatigue [5-7]. Symptoms and disabilities affect quality of life by increasing the risk of falls, mobility restrictions, and social isolation [5,6,8-14]. Moreover, patients with MS are often physically inactive or have a sedentary lifestyle as a consequence of the abovementioned symptoms and disabilities, initiating a vicious circle of deconditioning and worsening of symptoms [15,16]. MS is commonly diagnosed in young adults between 20 and 40 years of age and thus affects the early stages of their working lives [5]. All these factors lead to an increase in social and health care costs [17,18]. Therefore, there is a huge socioeconomic need to stabilize and counteract physical disabilities and cognitive dysfunctions by introducing effective therapies for patients with MS.

In general, physical exercise is a safe method that can yield beneficial effects such as depending on the training content, muscular strength, and aerobic capacity and, consequently, it improves mobility, fatigue, and quality of life in patients with MS [19-21]. A further training method that counteracts the aspect of cognitive decline is computer-based training. Specific computer-based training seems to positively influence different cognitive functions (eg, information processing, executive functions, and memory domains) in patients with MS [22-24]. However, both methods train the physical and cognitive components separately. A concurrent offering of both training components seems to be promising because this would promote the interplay of physical and cognitive functions and thus add *everyday life ecological validity* to the training approach [25].

An upcoming training method that concurrently combines the training of physical and cognitive functions is exergaming [26], "technology-driven physical activities, such as video game play, that require participants to be physically active or exercise in order to play the game" [27]. Typically, a player physically interacts with a video game represented on the screen via special controller technologies. Controllers track the player's movements and mediate them into a virtual game scenario that provides audio-visual feedback. In this way, commercially available exergames (eg, Nintendo Wii, Sony Move, or Microsoft Kinect) have successfully turned living rooms into playful training settings for approximately 10 years [28,29]. Apart from the entertainment market, video game–based training and therapy applications have also established themselves in the fitness and rehabilitation industry (eg, game-based, robot-assisted movement therapy [30,31]; virtually augmented climbing [32]; or exergame fitness training [29,33,34]). Besides the various beneficial effects of exergaming [35-37], the physical-cognitive interaction of exergames seems to trigger an alternating brain-body communication. Depending on the video game stimuli and the body-controller interaction, different cognitive and physical functions can be trained, which makes exergames a promising tool in MS therapy.

In recent years, researchers have started to evaluate exergames as a rehabilitation tool for patients with MS. Exergames proved to be an acceptable, feasible, safe, enjoyable, challenging, and self-motivating tool [38-40]. Kramer et al [41] concluded that the integration of exergames seemed to have a positive effect on training adherence and therefore could support the efficacy of long-term rehabilitation. Video game–based exercises, especially Nintendo Wii Fit, seem to improve static and dynamic balance as well as gait performance in patients with MS [41-44]. Intriguingly, these exercises led to improvements in the myelin sheaths of nerves in the brain areas involved in balance and movement [45]. Robinson et al [46] showed that the physical benefits of Nintendo Wii Fit training were comparable with traditional balance training in patients with MS. Furthermore, 2 recent systematic reviews concluded that exergaming enhanced cognitive functioning, in particular decision-making processes (executive functions) and visuospatial perception, in neurological patients who experience stroke, Parkinson disease, MS, or dementia [40,47]. However, many of the results so far stem from commercially available exergame systems (mainly Nintendo Wii and fewer Xbox Kinect and Sony PlayStation) that have not been developed for specific rehabilitation audiences. A review of exergame training in patients with MS suggested the development of exergames that target the training of a clinically identifiable need for this patient group [48]. For example, Nintendo Wii games did not appear to be entirely suitable for rehabilitation in MS because of a lack of flexibility and adaptability to the needs of patients with MS, which require special software development [49].

Human-computer interaction research, sports science, and human movement sciences offer numerous guidelines and frameworks aiming for more attractive and effective full-body motion games for different target populations [28,32,50-57]. Accordingly, these games should consider the needs and constraints of the target population [55,58,59]. One of these frameworks is the dual-flow concept that requires individual adaptable training features, thus ensuring that exergames are user-centered [55]. The dual-flow approach implies that exergame-based training provides an individual and optimal level of physical and cognitive challenge for every trainee throughout each training session by adapting the difficulty and complexity of the game to an individual's current physical, cognitive, and emotional states and needs in real time. Furthermore, the technology-based system of exergames allows the systematic and individual integration of training principles such as intensity, volume, progression, tailoring, and feedback [60-62]. Specific software algorithms continuously analyze and rate performance, thus allowing real-time adaptations. Recent findings of the international game research debate indicate that a player can be optimally motivated and stimulated with an adaptive game mechanic [29,34,55,63-65]. In combination with an audio-visually appealing exergame scenario (visuals, sound, story, etc), players' motivation can be increased [64]. *Having fun while training* with interactive games might have a huge impact on engagement and compliance [66]. Thus, a holistic exergame design approach can achieve an attractive and effective training experience by considering the levels of body, controller, and game scenario [33,67].

Objectives

In summary, there is a huge potential for developing effective and attractive user-centered exergames that combine training principles with elements of game design and focus on disease-specific deficits to increase motivation and performance and thus to ensure the possibility of successful training. The overall aim of the interdisciplinary research and development work presented here is to develop and evaluate user-centered exergames for the game controller Dividat Senso by incorporating a theoretical background from movement sciences, neuropsychology, and game research, as well as participatory design processes with patients with MS and their therapists. This work aims to contribute specifically to the following: (1) research-based, iterative, co-designed user-centered exergames for patients with MS and (2) the usability and feasibility testing of newly developed exergames by field testing and study trials.

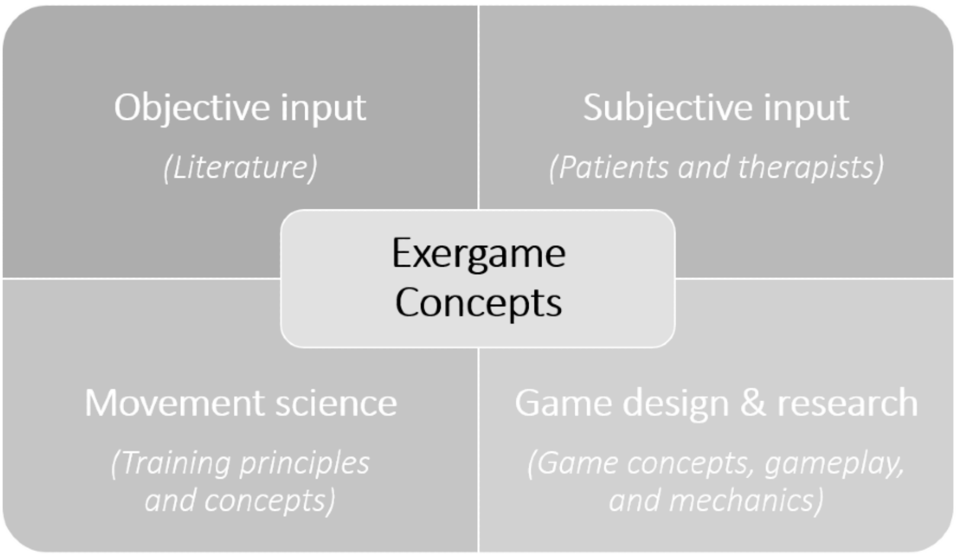
Methods

MS Exergame Concept

Design Process

As a first step, the iterative and research-based development process of the exergame concepts considered the knowledge gained from different user perspectives (patients with MS and therapists) and disciplines (human movement science and neuropsychology as well as game design and research) to holistically generate a potentially attractive and effective user-centered exergame training for cognitive-motor therapy in patients with MS (Figure 1). The multilevel design approach covered important aspects of the exergame concept: the hardware (the Dividat Senso plate), training concept (input movements, training principles, and cognitive tasks), and software (virtual game scenario).

Figure 1. Iterative and research-based development process.



This interdisciplinary research and development project developed new exergame concepts for the game controller Dividat Senso (Dividat; Figure 2). The Dividat Senso is a pressure-sensitive plate that serves as a game-input device. It uses specific lower body movements (eg, footsteps or weight shifts) to control various game scenarios presented on a screen. Several high-resolution sensors in the plate measure the force dynamically through body movements. The Dividat Senso plate further allows the generation of multidimensional sensory stimuli (eg, auditory, visual, and tactile). To support the trainee and for safety reasons, the plate is surrounded by a handrail.

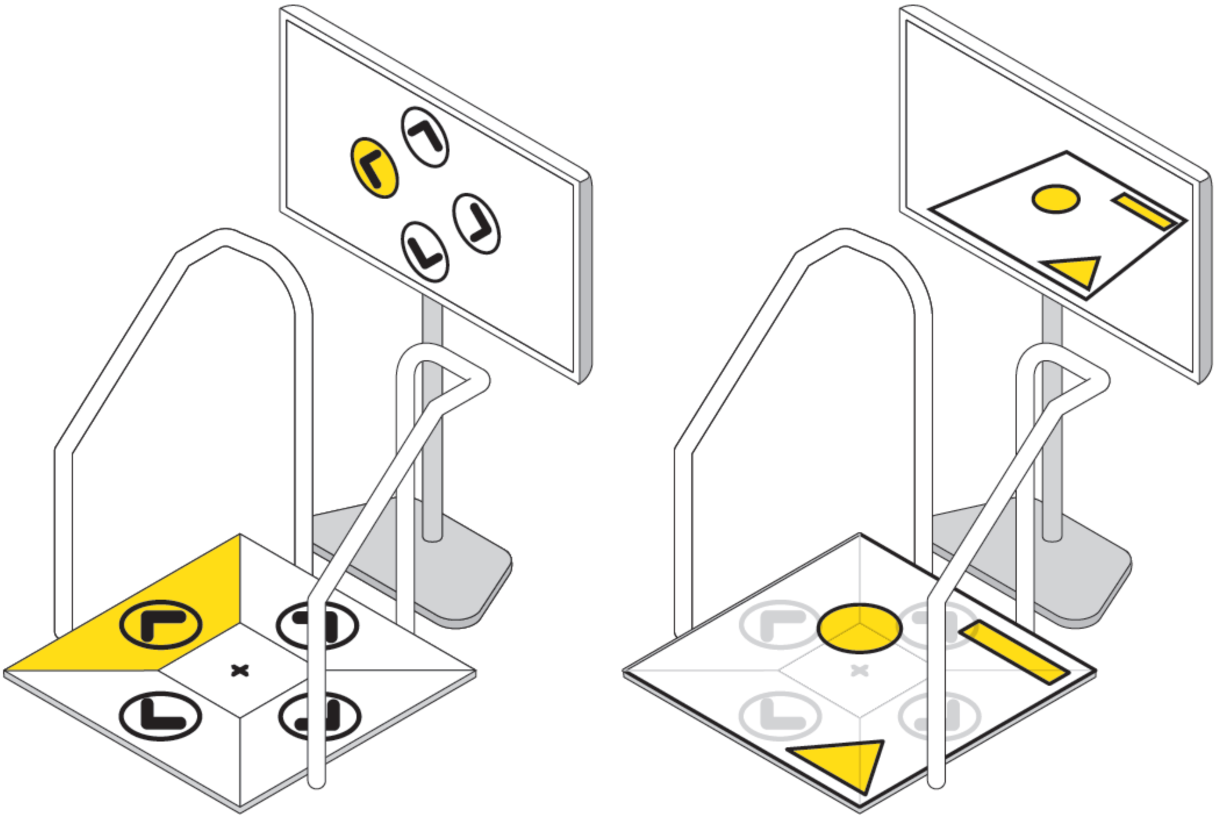
Figure 2. Original setup of the Dividat Senso.

Rethinking the Dividat Senso Plate (Hardware) and Game Designs (Software)

The design process started by analyzing the existing system and determining its technical opportunities, focusing on the Dividat Senso plate and the game collection, as they were not designed with or for the specific requirements of patients with MS. In this context, the project team visited certain therapy settings (rehabilitation center and physiotherapy) with neurologically impaired patients (MS and Parkinson disease) using the existing system in a therapy session. Furthermore, project members tested the plate and existing games themselves.

The most important finding was that patients often showed similar interaction patterns while playing on the Dividat Senso; patients first focused on the screen to receive the visual game stimuli and then tended to look down at the plate to step on the plate area to trigger the respective game input. This process seems to be important for patients with MS, as the motor learning process can be triggered via cognitive and motor information processing and realization [68]. However, the game control did not leave much room for maneuver, required very precise stepping, and did not make use of the whole plate. Such usage might interrupt movement dynamics and game flow [29,34,64,65] and leave certain gameplay options unused. Therefore, the plate layout was reconsidered, aiming for more intuitive, natural, and everyday-like patterns [34,69,70]. The focus was on using the entire pressure-sensitive plate, allowing the player to keep focusing on the game scenario and thus to stay uninterrupted in the game flow (Figure 3).

Figure 3. Rethinking the Dividat Senso plate. Concepts for more intuitive and natural input movements and flow are shown.



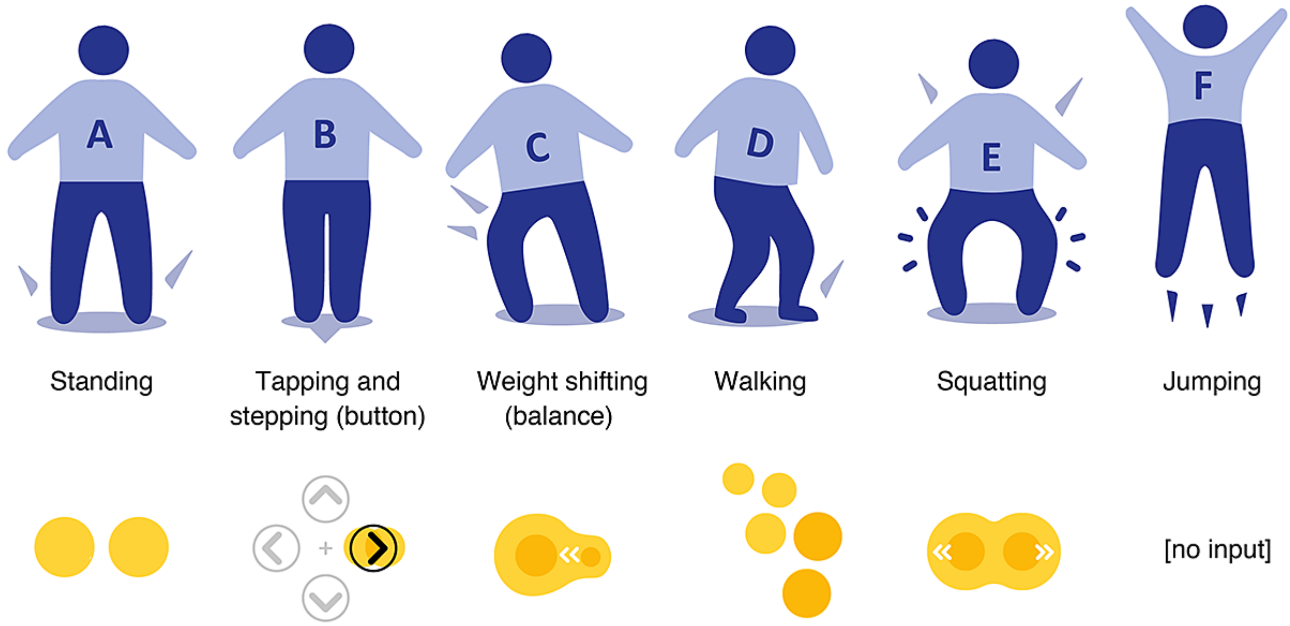
Moreover, some of the existing games did not necessarily follow a *meaningful design* [34] in terms of player perspectives [71] and the audio-visual representations of the cognitive stimuli and the respective motor challenges. For example, a virtual skier skies downhill while avoiding crashing into obstacles. The skier is represented on the screen in a third-person perspective with a top-down view and descends from above the screen but is controlled by sideways movements on the Dividat Senso plate where the left hand and right hand are flipped.

Rethinking the Training Concepts (Cognitive and Motor Tasks)

On the basis of the above reflections (usage and interaction patterns), the existing training concepts were also reconsidered, focusing on MS-specific motor and cognitive disabilities (eg, balance and coordination) and disease-specific deficits (eg, degeneration of myelin). Overall, the training concepts were developed and integrated by considering the following specific training principles: (1) type and specificity, (2) intensity, (3) progression, (4) variability, and (5) feedback [60-62]. Literature on exergaming in a therapeutic context was also considered [72].

In this process, some motor functions were considered that seem to be beneficial for patients with MS. Patients with MS often experience, to a variable extent, muscle weakness, diminished dexterity, spastic paresis, sensory dysfunction, gait disturbances, and fall risk, as well as fatigue and depression [5,6,73,74]. Therefore, the training concepts aimed to integrate motor control components, focusing on static and dynamic balance and coordination skills. Figure 4 shows the preexisting and reconsidered input movement. In terms of cognitive stimulation, the training concept aimed to integrate cognitive functions that may be affected in patients with MS, such as information processing, attention, decision making, error correction, executive functions, and memory [7,73-75].

Figure 4. Input movements, including existing patterns (A, B, and C) and rethought patterns (D, E, and F). Input movements are presented as body models and as patterns that are registered by the pressure-sensitive plate.



A further training concept for exergames that must be mentioned is the dual-task approach. Study findings indicate that patients with MS have impaired dual- or multi-task performances that could result from their deficits in divided attention, resource capacity overload, or differential neural activation [76-80]. In this case, exergames allow the concurrent processing and synchronization of cognitive and motor stimuli and therefore seem to support constant body-brain communication. These processes might be advantageous as they are close to day-to-day activities, such as walking in an enriched real-world environment.

Furthermore, the reconsidered training concepts considered both games that endorse motor learning [68] and games that require moderate continuous exercise performance [81] in order to replicate preliminary findings of physical training on myelin sheath regeneration as well as to specifically target important disability-related structural deficits seen in patients with MS.

Focus Groups: Cocreating New Exergame Concepts

Following the rethinking process, new exergame scenarios were designed. To ensure that the concepts were user-centered, the target group (patients with MS and their therapists) was involved from the outset. A semistructured interview guideline was developed based on questions about all elements of the exergame environment (eg, body, controller, and virtual game scenarios). The aim of the focus group interviews was to explore the target group's experiences with exergames and technology in the context of therapy, as well as to define needs, preferences, and expectations for an optimal exergame setup and its integration into an MS therapy setting. The focus group surveys took approximately 90 minutes and were carried out with 4 physiotherapists experienced in MS therapy, 9 patients with MS, and 2 specialists in neuropsychology. In addition to a list of specific questions, participants' thoughts and specific wishes for the look and feel of future exergames were assessed using 3 different sketches of potential game scenarios (Figure 5).

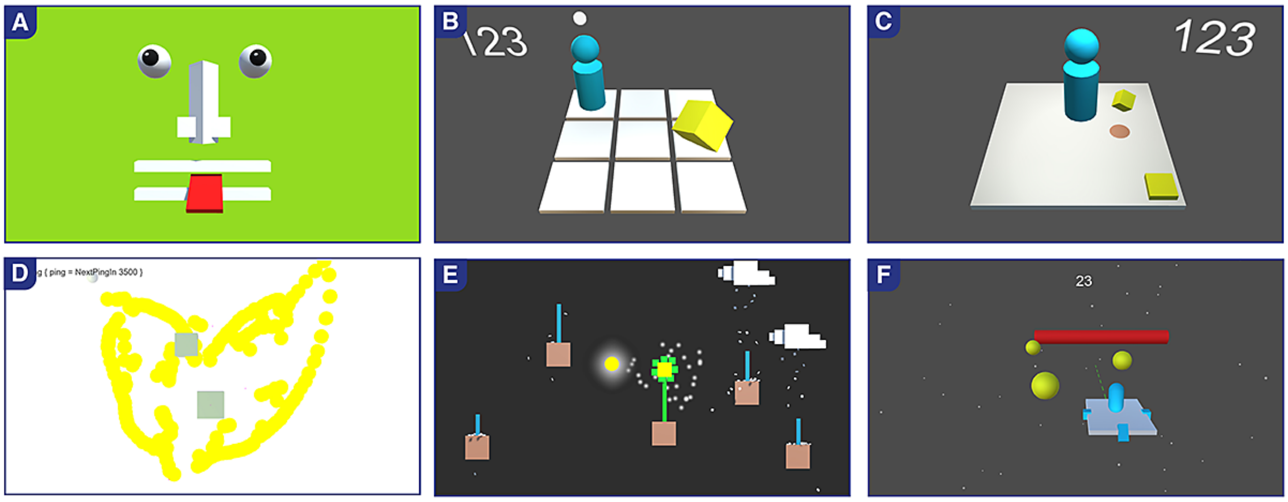
Figure 5. Three sketches of potential game scenarios. Different gameplay options, game mechanics, and perspectives served as inspiration during focus groups. The Puddle Jump sketch (A), the Gentle Giant sketch (B), and the Owl Flight sketch (C).

On the basis of the results of the focus groups, personas for the 2 target audiences were developed. The primary aim was to provide patients with MS (predominantly adult females of all ages, ranging from high to low fitness) an attractive and effective training. The secondary aim was to provide physiotherapists (who are open to the use of technology in movement therapy) with a flexible supplementary tool to their traditional therapy methods. Among other outcomes, the focus groups revealed that the design should not be restricted to a specific age or gender group nor to a single game style and input movement concept, because the MS disease pattern is very heterogeneous. Therefore, different exergame scenarios were designed, including different game mechanics, narratives, perspectives, and input movements with the Dividat Senso. Each scenario provided slightly different cognitive and motor challenges and aimed at patients with MS aged around 30-85 years who fulfilled further requirements (see the study criteria in *Recruitment and Participants*).

Field Research: Initial Concept Testing

In total, 6 box prototypes (Figure 6) were modeled using the game engine Unity 3D and showcased at numerous neurorehabilitation trade shows. After visitors of the trade shows, especially therapists and patients, tested the box prototypes, mock-ups of different themes (street, kitchen, alpine, underwater, forest, garden, oriental, and sci-fi) were shown to them. People could rate their favorite game and choose the theme that would suit them best. Out of the 6 box prototypes, the 3 concepts that were most promising and best rated were retained. The survey showed that both patients and therapists of different gender and age groups rated natural, garden-like game settings the highest.

Figure 6. Unity 3D box prototypes. Based on the input from the focus groups, different game scenarios and mechanics were designed. A and D: Two playful, toy-like 2D prototypes allowing the feet to move freely on the Dividat Senso plate to draw and play with a face. E: 2D scenario allowing free steps or weight shifting. B and C: Two 3D images of the Dividat Senso plate acting as a virtual playground, allowing free steps and jumps. F: 3D Racer scenario with a weight shifting input.



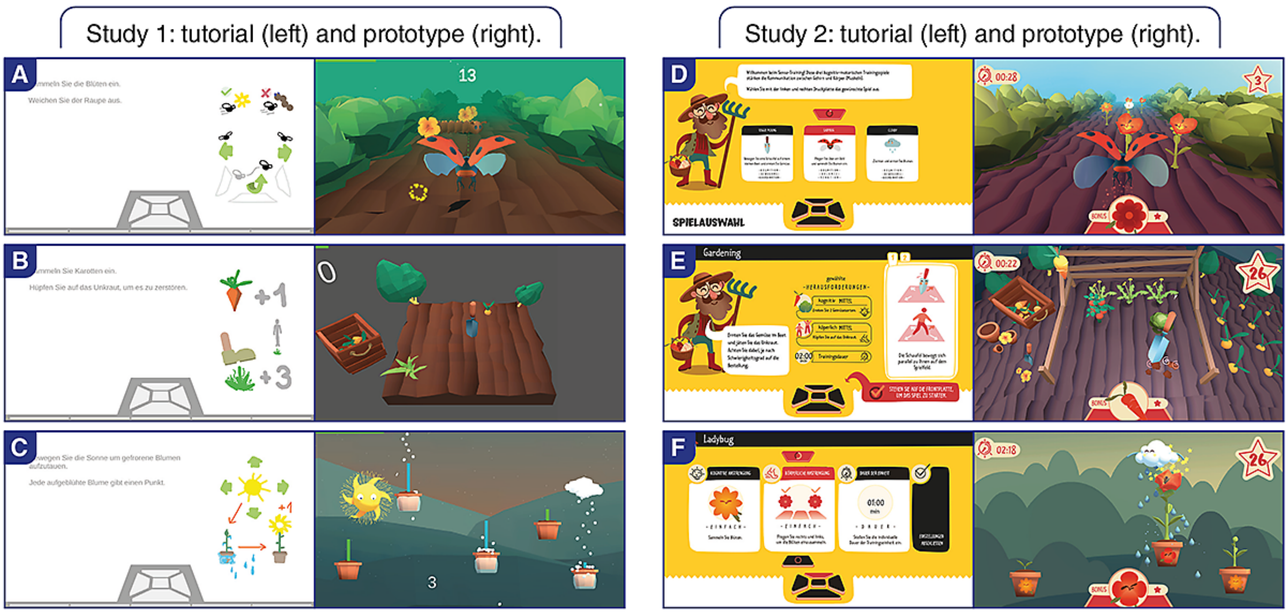
Game Concepts: Design, Redesign, and Finalization

Following the preliminary field research, 3 exergame concepts were designed, including different virtual game scenarios and game mechanics, each demanding other input movements on the Dividat Senso plate. The specific descriptions of the video games, visualization of the input movements, and visual progression overview can be found in Table 1, Figure 4, and Figure 7, respectively. In all 3 exergames, the following training principles of motor learning [68] were integrated specifically to train MS-specific disabilities (eg, balance and coordination) and disease-specific deficits (eg, degeneration of myelin): (1) type and specificity (MS-specific motor and cognitive components; see also *Rethiking the Training Concepts* and Table 1); (2) intensity and progression (level adjustment and in-exergame adaptation [movement speed avatar, Ladybug] allowing for moderate continuous exercise experiences) [81]; (3) variability (3 exergames to capture different training foci; Table 1); and (4) feedback (scoring and sound effects).

Table 1. Game concepts for the game controller Dividat Senso.

Exergames	Ladybug	Scooper	Cloudy
Description	Navigation of a ladybug to collect randomly allocated flowers and avoid collisions with obstacles	Harvesting garden vegetables	Setting the position of the sun (Study 1) or a rain cloud (Study 2) to grow flowers
Motor components	Static balance and coordination	Dynamic balance, coordination, accuracy, and strength	Static balance, coordination, accuracy, and strength
Cognitive components	Information processing, anticipation, selective attention, and visual-spatial orientation	Information processing, planning, selective attention, and visual-spatial orientation	Information processing and selective attention
Motor-level settings (Study 2)	• Level 1: Side stepping, tapping or weight shifting • Level 2: Side stepping, tapping or weight shifting and stepping to the front to avoid obstacles (stones) • Level 3: Side stepping, tapping or weight shifting and stepping to the front to avoid obstacles (caterpillars)	• Level 1: Walking and standing on objects for collection • Level 2: Walking and squatting on objects for collection • Level 3: Walking and jumping on objects for collection	• Level 1: Side stepping or tapping • Level 2: Side stepping or tapping and squatting to make the cloud rain • Level 3: Side stepping or tapping and jumping to make the cloud rain
Cognitive-level settings (Study 2)	• Level 1: Pick all flowers • Level 2: Pick bonus flower (2 colors) • Level 3: Pick bonus flower (3 colors)	• Level 1: Pick all vegetables • Level 2: Pick bonus vegetables (2 colors) • Level 3: Pick bonus vegetables (3 colors)	• Level 1: Water all flowers • Level 2: Water bonus flower (2 colors) • Level 3: Water bonus flower (3 colors)

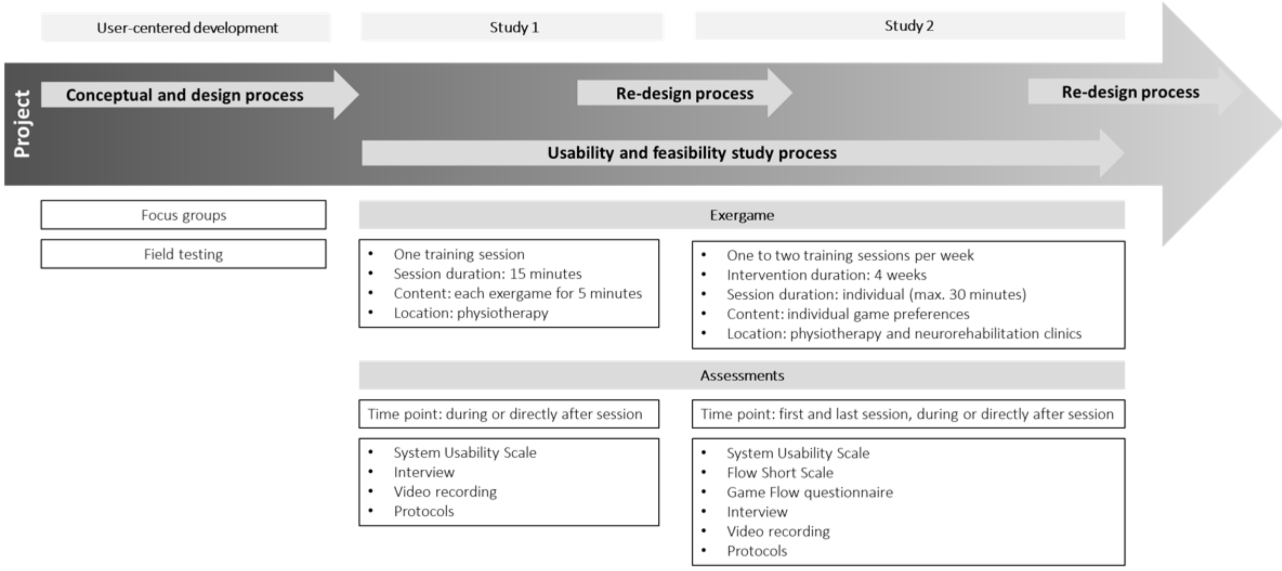
Figure 7. Study setup and in-game screenshots of the tutorial and game tested in the first study (A, B, and C) and in the second study (D, E, and F).



Study Design

Two usability studies were conducted to evaluate the usability and feasibility of the newly developed user-centered exergames in patients with MS. From January to February 2019, the measurements for the first study were taken, and from April to May 2019, the training sessions and measurements for the second study were conducted. Figure 8 shows the project process, including the 2 user studies.

Figure 8. Project schedule.



In the first study, patients with MS tested each exergame concept (Figure 7) for 5 minutes in a random order. Video recordings and observation protocols for exergame performance and interaction were assessed by trained study investigators at a physiotherapy center (Physiotherapy Langmatten, Binningen, Switzerland). After the exergame sessions, patients rated the System Usability Scale (SUS) and answered predefined interview questions.

In the second study, patients with MS played the redesigned exergame concepts (Figure 7) over a period of 4 weeks. Each patient was trained 1 to 2 times per week at a physiotherapy center (Physiotherapy Langmatten) or at one of the neurorehabilitation centers (ZURZACH Care, Rehaklinik Bad Zurzach, Bad Zurzach, Switzerland, and Reha Rheinfelden, Rheinfelden, Switzerland). In the first training session, the participants tested all 3 exergame concepts at level 1 for motor and cognitive adaptations. In the following sessions, patients could decide which exergames they wanted to play and for how long. This procedure was chosen to obtain an impression of the patient's preferences. Regarding training progression, levels for motor and cognitive functions were individually adapted from session to session, aiming for moderate training intensities (values between 3 and 4 on the modified Borg scale, which ranges from 1 to 10) over 4 weeks. Furthermore, training time was individually increased from week to week for each patient while ensuring a minimum training time of 20-25 minutes per session. However, as the daily state of patients with MS was unpredictable, the level and training time fluctuated in some cases. In the last training session, each patient replayed each of the exergame concepts by starting from where they had left off at the last training session to familiarize themselves with the concepts before the postmeasurements. Measurements were taken during the first training session and at the last training session. During the exergame performance, video recordings and observation protocols for exergame performance and interaction were assessed. After the exergame performance, patients rated the SUS and answered the Flow Short Scale (FSS), Game Flow questionnaire, and predefined interview questions. The training sessions were supervised by trained researchers and physiotherapists, and the measurements were taken by trained researchers.

The ethics committee of ETH Zurich, Switzerland, approved both study protocols (EK 2018-N-85 and EK 2018-N-124). Before any measurements were taken, all eligible patients provided written informed consent according to the Declaration of Helsinki. Withdrawal for no stated reason was permitted at any time during the study.

Recruitment and Participants

In the first study, potential participants were recruited by physiotherapists from a physiotherapy center (Physiotherapy Langmatten). In the second study, participants were recruited by physiotherapists and study investigators from specialized centers for neurological physiotherapy (Physiotherapy Langmatten) and rehabilitation (ZURZACH Care, Rehaklinik Bad Zurzach and Reha Rheinfelden). In both studies, all interested patients were fully informed about the study procedure and the inclusion criteria by physiotherapists and study investigators before screening. Patients who met the initial eligibility criteria and signed the informed consent form participated in a personal interview to screen for mental and physical health. Screened data included demographic data and medical information regarding MS (eg, MS type, leg spasticity, and fatigue). Furthermore, the following 2 questionnaires were assessed to define prevalent MS-related restrictions: the MS Impact Scale [82] and Activities-specific Balance Confidence scale [83].

For the first and second study, the same eligibility criteria were set. Patients fulfilling all the following inclusion criteria were eligible: (1) female or male; (2) aged 25-80 years; (3) clinical diagnosis of MS, including all forms (relapsing or remitting, primary-progredient, secondary-progredient, and progressive-relapsing); (4) stationary and ambulant; (5) able to provide written informed consent and understand instructions; (6) able to stand at least for 10 minutes with the aid of a handrail; and (7) visual acuity including correction sufficient to work on a television screen. Any of the following criteria led to exclusion: (1) conditions that precluded stepping exercise (severe spasticity that prevents a person from taking a full step or severe musculoskeletal injury), (2) excessive fatigue that prevented training participation, and (3) exercise intolerance that prevented training participation.

Assessments

Table 2 illustrates the assessments used for the first and second studies.

Table 2. Study assessments.

Category	Explanation
Feasibility	
Training adherence and attrition rate	• Compliance with training sessions • Participants lost at follow-up (dropouts)
Usability	
System Usability Scale	• Reliable and valid tool providing a global view of subjective usability [84-86] • A score of at least 70 for an "acceptable" solution, below 50 is "unacceptable," and 50-70 is "marginally acceptable" [86] • 10 items (5-point Likert Scale), score range 0-100
Flow Short Scale	• Used to retrospectively get a typical flow-score for specific kinds of actions or situations [87] • 13 items (7-point Likert Scale), score range 1-7 • Dimensions: flow (items 1-10), fluency (item 2, 4, 5, 7, 8, and 9), absorption (items 1, 3, 6, and 10), and perceived importance (items 11-13)
Game Flow questionnaire	• Derived from the Sweetser and Wyeth [53] "Game Flow" model, which determines the key elements of player enjoyment • 17 items (6-point Likert Scale), score range 1-6 • 7 main items (items 1-7) building the dimension Game Flow and 10 additional explorative exergame-specific items (items 8-17)
Feasibility and usability	
Guideline-based interview	• Qualitative evaluation of the user's game play experiences • Categories: (1) overall experience, (2) game scenario, (3) Dividat Senso plate (game controller), (4) body and mind, (5) motivation, (6) training, (7) comparison to conventional movement therapy, and (8) others
Video recording and monitoring protocol	• Exergame performance • Same categories as for the interview
Training parameters	
Physical and cognitive exertion	• Modified Borg Scale from 1 to 10 [88]
Number of trainings	• Range from 4 to 8 trainings
Training time	• How long participants trained per session
Play preferences	• How often each exergame was played

Data Analysis

For quantitative data, statistical analysis was conducted using SPSS (IBM SPSS 26). The level of significance was set at $P<.05$. The data were compared using the Wilcoxon signed-rank test, as the assumptions for parametric statistics were not met (nonnormally distributed data). The effect size (r) was calculated using the following equation [89]:

$$r = z/\sqrt{(N)}$$

An effect size of 0.10–0.29 indicates a small effect, an effect size of 0.30–0.49 indicates a medium effect, and $r \geq 0.50$ indicates a large effect [89]. The interviews were assessed by 5 of the authors (1 game researcher and 4 movement scientists) following an iterative thematic coding approach based on qualitative content analysis [90]. For all interviews, the coders individually transcribed and coded the data according to the categories of the interview guidelines. In 2 iterations, the coders discussed the emerging results until an agreement was reached. Finally, two of the authors (1 game researcher and 1 movement scientist) further summarized the findings. A preliminary explorative analysis was conducted on the observation protocols and videos, but for the purpose of this paper, they were only used to check certain findings from the interview analysis.

Results

Participants

The participant characteristics are shown in Table 3. At the beginning of the second study, 29 patients with MS were included, while 4 patients with MS dropped out (attrition rate: 4/29, 14%) during the study period. The reasons for dropout were disease-related weakness, physical condition, early clinical release, and scheduling conflicts. In total, participants completed 70 training sessions (mean 4.8 training per participant, SD 1.1) with the exergames. Of the 25 patients with MS, 4 patients with MS missed a training session once and 1 patient with MS missed a training session twice, leading to an attendance rate of 95% (120/126). The reasons for missed training were overload and fatigue after training, illness, absence, date conflict, and holiday. Considering the game preferences in the second study, the participants mostly played *Ladybug* (1.51 sessions per training), followed by *Scooper* (1.19 sessions per training), and the least *Cloudy* (0.91 sessions per training). Overall, no adverse events were recorded in the first and second studies.

Table 3. Baseline and training data characteristics.

Characteristics	Study 1 (N=16)	Study 2 (N=25)
Gender, n (%)		
Female	10 (62)	15 (60)
Male	6 (38)	10 (40)
Age (years), mean (SD)	62.1 (13.0)	57.3 (11.2)
Types of MS[a] , n (%)		
RR[b]	7 (44)	11 (44)
SP[c]	2 (12)	10 (40)
PP[d]	7 (44)	3 (12)
Not applicable	0 (0)	1 (4)
Therapy stay, n (%)		
Ambulant	16 (100)	19 (76)
Stationary	0 (0)	6 (24)
Diagnosis since (years), mean (SD)	22.73 (13.1)	16.6 (11.7)
MSIS[e], mean (SD)	33.2 (16.7)	34.3 (15.0)
MSIS physical, mean (SD)	34.2 (20.1)	36.2 (15.5)
MSIS psychological, mean (SD)	29.7 (23.0)	32.1 (19.2)
ABC[f], mean (SD)	74.7 (11.7)	69.4 (18.2)
Exergame experience, n (%)	2 (13)	10 (40)
Number of trainings per participant, mean (SD)	1 (0)	4.8 (1.1)
Training time per session (min), mean (SD)	15 (0)	19.1 (3.9)
Borg motor, mean (SD)	3.3 (1.2)	3.8 (1.8)
Borg cognitive, mean (SD)	4.0 (1.8)	3.5 (1.9)

[a]MS: multiple sclerosis.
[b]RR: relapsing-remitting.
[c]SP: secondary-progressive.
[d]PP: primary-progressive.
[e]MSIS: multiple sclerosis impact scale.
[f]ABC: Activities-specific Balance Confidence scale.

Quantitative Data

In the first study, the median SUS score was 71.3 (IQR 58.8-80.0). The SUS and questionnaire pre-post comparisons of the second study are presented in Table 4.

Table 4. Questionnaire data (N=25).

Questionnaires[a]	Pre, median (IQR)	Post, median (IQR)	z	P value	*r*
System Usability Scale	89.7 (78.8-95.0)	82.5 (77.5-90.0)	−2.077	.04[b]	0.42
Flow Short Scale[c]	5.9 (4.6-6.4)	5.8 (5.4-6.2)	−0.400	.69	0.08
Fluency	5.7 (4.4-6.6)	5.6 (4.8-6.6)	−0.325	.75	0.07
Absorption	5.8 (5.1-6.5)	6.0 (5.1-6.6)	−0.485	.63	0.10
Perceived importance[d]	2.0 (1.5-3.8)	1.3 (1.0-3.5)	−2.118	.03[b]	0.42
Game Flow[c]	5.0 (4.7-5.3)	5.1 (4.9-5.3)	−0.473	.64	0.09
Concentration	5.0 (5.0-6.0)	6.0 (5.0-6.0)	−0.775	.44	0.16
Challenge	4.0 (2.5-4.5)	4.0 (3.0-4.8)	−0.210	.83	0.04
Skills or abilities	5.0 (4.0-5.0)	5.0 (4.0-5.0)	−0.277	.78	0.06
Control	5.0 (4.5-5.0)	5.0 (4.5-6.0)	−0.732	.46	0.15
Aim	6.0 (6.0-6.0)	6.0 (6.0-6.0)	−0.816	.41	0.16
Feedback	6.0 (5.0-6.0)	6.0 (6.0-6.0)	−1.030	.30	0.21
Immersion	5.0 (5.0-6.0)	5.0 (5.0-6.0)	−0.811	.42	0.16
Pleasure and liking	6.0 (5.0-6.0)	6.0 (5.0-6.0)	−0.264	.79	0.05
Dual flow over—challenge[d]	1.0 (1.0-2.5)	1.0 (1.0-2.0)	−0.577	.56	0.12
Dual flow under—challenge[d]	1.0 (1.0-3.0)	2.0 (1.0-2.8)	−0.418	.68	0.08
System control	5.0 (4.3-5.0)	5.0 (5.0-6.0)	−1.604	.11	0.32
Movement	5.0 (5.0-6.0)	5.0 (5.0-6.0)	−0.351	.73	0.07
Motivation	6.0 (5.0-6.0)	6.0 (5.0-6.0)	−0.816	.41	0.16
Physical exertion[e]	4.0 (2.0-5.0)	4.0 (2.0-5.0)	−0.158	.88	0.03
Cognitive exertion[e]	3.0 (2.0-4.5)	3.0 (2.0-4.0)	−0.042	.97	0.01
Optimal challenge	5.0 (4.0-5.0)	4.0 (4.0-5.0)	−0.842	.40	0.17
Spatial presence	5.0 (3.5-6.0)	5.0 (4.0-6.0)	−0.361	.72	0.07

[a]Data were analyzed using Wilcoxon signed-rank test.

[b]P<.05.

[c]The higher the scores, the better the results. This counts for all items that are not specifically marked.

[d]The lower the scores, the better the results.

[e]The more in the middle field, the better the results.

Qualitative Data

Findings from the guideline-based interviews of both studies are reported for *overall experience* (Figure 9), *body and mind* (Figure 10), *games, gameplay experience, and hardware* (Figure 11), *motivation* (Figure 12), and the *comparison of exergames with conventional therapy* (Figure 13).

In summary, all participants reported an enjoyable, motivating, varied, and fun experience with the exergames, which was a completely new thing for most of them (Figure 9, Figure 11, and 12). They also reported that, in addition to having a lot of fun while being challenged, they felt a clear improvement in the handling (coordination and physical interaction) of the new technology over time (Figures 10 and 11), which made them feel more confident in using it (Figures 9 and 11). On the level of body and mind, participants clearly focused on the virtual gaming world, which distracted them from physical exertion and made it seem very pleasant, albeit challenging, but by no means overstraining (Figure 10). By immersing in the game world, patients were able to forget their everyday worries (often associated with the disease) for the moment (Figure 10). Regarding the potential use of exergames as a therapeutic device, most participants saw the added value of the novel training solution in terms of distraction from everyday life, fun, and the combined body and brain training approach, even though traditional therapy measures were also described very positively and were difficult to compare (Figure 13). A complementary integration of the exergames into therapy could be imagined very well by all patients. Further development of the exergames over the 2 studies was also perceived positively.

Figure 9. Interview data focusing on overall experience. (Some and minority = at least 30% of the participants; many = at least 50% of the participants; most and majority = at least 80% of the participants).

1st study

All participants were eager and **motivated to explore** the exergames.

Feelings:
Some participants
- were **surprised** by their own performance
- **forgot** everything
- **were totally immersed**
- **had fun**

Overall, most participants showed interest in participating in the follow-up study, including a training intervention over 4 weeks.

2nd study before

Participants commented that the game session was
- **playful**
- **interesting**
- **fun**
- **entertaining**
- **not dull**
- **physically activating**
- **agile**
- **challenging**

Feelings:
Participants **enjoyed experiencing something new** and had mainly positive feelings. They felt
- **concentrated**
- **amused**
- **pleased by colors**
- **curious**
- **full of anticipation**
- **comfortable**
- **reminded of childhood**

Some participants reported feelings such as **clumsiness and uncertainty** at the beginning. After a short familiarization phase, these feelings **turned into immersion, excitement and ambition** to play and to correctly play the exergame.

Remembrance:
- Participants often mentioned **things related to their feelings** during the game sessions.
- Some addressed the **physical activity and agility.**
- Participants often **thought of one specific exergame as the most rememberable.**

2nd study after

Participants had **fun**, and most perceived the exergame to be
- **interesting**
- **exciting**
- **brilliant**
- **enjoyable**
- **new**

Many participants were glad
- to have had **the chance to be part** of the study.
- to **contribute** to this project.

Feelings:
Participants' **feelings** throughout the study were **very positive**, (in line with pre-interview). Almost all participants
- **enjoyed** the games.
- were transported back to their **childhood.**
- felt **more and more at ease, the more familiar** they became.
- enjoyed the **deflection from daily-life problems**.

Remembrance:
- Some addressed the **physical–cognitive interaction** performed on the plate.
- Others reported that due to this activity, they could **forget time** and **daily-life issues**.
- One participant mentioned the **varying response capacity and efficiency** depending on the time of day.

Changes before/after:
Overall, the study **did not create noteworthy changes** in physical–cognitive conditions because of
- **short training time**
- **limited number of training sessions**

Participants **commented the most** on **mental adjustments,** such as easier handling of the games or higher levels of confidence.

One participant experienced a **dramatic physical improvement** and could use new training methods.

Figure 10. Interview data focusing on body and mind. (Some and minority = at least 30% of the participants; many = at least 50% of the participants; most and majority = at least 80% of the participants).

1st study

Perceived physical demand and load: All participants perceived the exertion as **easy to moderate**.

Perceived cognitive demand and load: Many participants had **high concentration levels**.

Focus: Many participants **focused** on the game, but were still **aware of their own bodies** and movements, **switching their focus** between game and body.

Steering movements: Movements felt **intuitive** but more or less **unnatural** (eg, due to a participant's impairments).

Challenge: Challenge was **neither too high nor too low**.

Progression:

- **Difficulty** was higher in the beginning.
- After **familiarization,** it was easier to play until the game became more complex and faster.

Multiple sclerosis (MS)-specific limitations:

- **No limitations** due to MS were reported.

Some participants cognitively **processed** what was happening in the game, **but could not physically react** as fast as required or as intended (eg, to score).

2nd study before

Perceived physical demand and load:

- Participants **primarily** perceived the exertion as **challenging**.
- Participants **enjoyed** controlling the games by **body movements**.

Perceived cognitive demand and load:

- Participants **secondarily** perceived the cognitive load as **challenging**.

Focus:

- Most participants **mainly focused** on the game and **barely noticed their surroundings**.
- Some participants **set their focus differently on the 3 games** (depending on their impairments and fitness).

Steering movements:

- Movements felt **intuitive, comfortable** and fairly **natural**.

Challenge:

- **Physical aspects** seemed to be the main challenge (although physical challenge was **not too high**). Participants were confronted with **handling balance, stability and coordination**.
- Few participants (additionally) mentioned **cognitive aspects** that challenged them most, such as **mental fatigue, concentration on the screen and comprehension** of how to play the game.

Progression:

- **Difficulty** was higher in the beginning.
- After **familiarization,** it was easier to play.
- Most participants **perceived in-game feedback** but **not all understood it properly** (eg, the meaning of scores).

Multiple sclerosis (MS)-specific limitations:

- Some participants could **process** what was happening in the game, **but could not transfer it into physical movements** as fast as required or as intended.
- Others **forgot the restrictions** caused by their impairment and were **surprised how well** they performed specific movements during the game session (eg, walking backwards).

2nd study after

Perceived physical demand and load:

- Some participants mentioned a focus on **physical aspects** (eg, balance and precise performance).

Perceived cognitive demand and load:

- At the end of the study, participants initially set their **focus on the mind** (eg, they concentrated on the game process).

Focus:

- Most participants **mainly focused** on the games and **barely noticed their surroundings.**
- Some participants **lost their sense of time**.

Challenge:

- A third of all participants agreed that it was the **combination of physical and cognitive aspects that challenged them the most**.

Progression:

- **Subjectively perceived improvements** were mainly seen in the game **"Ladybug" (better motor control, physical precision, improved tactics, and more self-confidence).**

Multiple sclerosis (MS)-specific limitations:

- Participants were less concerned and **enjoyed a distraction from daily life and impairments**.

Figure 11. Interview data focusing on games, gameplay experience, and hardware. (Some and minority = at least 30% of the participants; many = at least 50% of the participants; most and majority = at least 80% of the participants).

1st study

Favorite game:

- All 3 game scenarios were considered a favorite by different participants.
- The **favorite game** was the game participants were the **most successful** in playing.
- In some cases, they chose the **game** with which they had the most difficulties that had **provided the biggest challenge**.

Gameplay experience:

- Games were described as **comprehensible** and generally **easy to understand.**
- Appearance of the game scenarios was described as **enjoyable and appealing**.
- Need for **instructions and an introductory tutorial** was claimed.

Hardware:

- Participants **mastered the Dividat Senso game control very well.**
- They **quickly understood** how to use the Dividat Senso plate.
- One participant was **slightly afraid of falling off** the Dividat Senso plate.

2nd study before

Favorite game:

All 3 game scenarios appealed to the participants.

- The game that was **liked the most** and was likeliest to **be played again** was **"Ladybug" (great dynamics, rapidity and reaction)**.
- The second favorite game and second likeliest to be played again was **"Scooper" (slowness and comfortable to play).**
- The least favorite game was **"Cloudy" (meaningfulness, sense of purpose in the game, and easy to play).**

Most participants most enjoyed the game that provided them with the **most difficulties** and the **biggest challenges.** They still **liked the option of going for an easier game** on days when they did not feel fit enough.

Gameplay experience:

- Many participants **lost their sense of time** as they were immersed in the moment.
- If participants experienced time as passing rather slowly or fast, they related it to the **newness** or their **fitness over time**.
- Games were described as **comprehensible** and generally **easy to understand.**
- In-game instructions were helpful and **easy to understand**.
- Appearance of the game scenarios was described as **appealing** (participants liked the colors as well as the clear, bright and delightful pictures).

Hardware:

- The comprehension of the game control was evaluated as good – participants generally **understood quickly how to handle the in-game navigation**.
- The menu navigation was rated **controversially** (likes and dislikes for the confirmation step to the front plate).
- There was a **satisfying sense of the system's sensitivity** (participants reported that the system reacted well and immediately).
- Some participants were a bit **afraid to step or walk backwards.**

2nd study after

Favorite game:

- The game that was **liked the most** and was likeliest to **be played again** was **"Ladybug" (dynamics and fast movements)**.
- The second favorite game and second likeliest to be played again was **"Scooper" (movement variations, cognitive challenge and realism of scenario).**

Most participants enjoyed and wanted to replay the game that was the **most challenging and fun**.

Gameplay experience:

- Some participants **surpassed themselves** (eg, they jumped not knowing they were able to).
- Participants **managed** and **appreciated** the game **difficulty adjustment**.
- Appearance of the game scenarios was described as **attractive, lively, colorful and comprehensible**.

Hardware:

- After the study period, the game control was without exception rated as **good** and **comprehensive.**
- Most participants gained **more confidence and psychological security** with increasing experience.
- The menu navigation was rated **controversial** (likes and dislikes for the confirmation step to the front plate).
- Participants who required the **handrail** for more stability generally considered the handrail **helpful and appropriate.**

Figure 12. Interview data focusing on motivation. (Some and minority = at least 30% of the participants; many = at least 50% of the participants; most and majority = at least 80% of the participants).

1st study
Motivation: • Most participants were **motivated and had fun playing.** • Participants could imagine **training with the exergames once or twice a week for 10 to 15 minutes.** **Long-term motivation:** • Majority of the participants wanted **increasing game difficulty** to obtain a constant challenge (eg, complex gameplay, and increased speed and broader game collection. • Some participants stated that **games would never get boring** because their **daily condition varied (flare-ups)** and the challenge would therefore be maintained.

2nd study before	2nd study after
Motivation: • Most participants were **motivated and interested, curious to play** the games and wanted to **perform well.**	**Motivation:** Participants were driven • **by the ambition to perform well physically** and by the opportunity to **be physically active** again. • by the **curiosity** about what the other game's levels were like. • by the research project contribution. **Only for some participants, collecting points** was the motivator (some felt encouraged to break their records, and others did not notice their scores).
Long-term motivation: • Participants could imagine the exergames **being motivating and fun to play even after long-term use.** • For some participants, the main motivation was to **get better** and **increase their physical performance over time**.	**Long-term motivation:** Some participants wanted **new and slightly more challenging** game levels to keep a certain level of challenge over time.

Figure 13. Interview data focusing on the comparison of exergames with conventional therapy. (Some and minority = at least 30% of the participants; many = at least 50% of the participants; most and majority = at least 80% of the participants).

1st study

Exergames:
For most participants, the exergames were more motivating than the conventional therapy, because they were
- **fun** to play.
- great **brain training.**
- a **distraction from daily problems** and concerns.
- a **new** training environment.

Therapy:
- For some participants, conventional therapy and exergames were completely **different** and **hard to compare.**
- Some participants preferred a **combination of both** trainings for maximum benefit.

A few participants **preferred the conventional therapy** because
- they could train **more specifically.**
- they could **feel and know the materials** they are training with.
- they **liked the interaction** with the therapist or animals.

Integration of exergames into therapy: Keep it simple and effortless because patients already have quite a busy daily program or need help to get to therapy or to leave the house.

2nd study before

Exergames:
The exergame sessions were seen as
- **fun** and **challenging.**
- an active **combination of physical and cognitive** challenge.
- a **multiple-task training.**
- an **unconstrained opportunity to move** and be active.

Compared to medical **fitness training**, participants felt **less physically challenged** during the gameplay.

Therapy:
- Physiotherapy is generally perceived to be **less active, beneficial and soothing**.
- The **focus is mainly on physical functions** and the aim is to **relearn processes** and efficiently use the remaining **body** functions.
- Participants **emphasized the cooperation between therapist and patient** and the **definition of a long-term goal.**
- Some mentioned that **mistakes and failures** in physical progress **were better addressed** in therapy.

2nd study after

Exergames:
The exergame sessions
- were seen as a potential **supplementary part of therapy** by the majority of participants.
- were a **well-received** and **well-balanced combination of body and brain training.**
- **distracted from daily problems** and concerns.
- were **enjoyable.**
- **caused loss of sense of time** as participants were immersed in the moment
- **increased self-confidence and physical fitness in daily-life activities.**

Compared with conventional therapy, the exergames were **more fun, entertaining and motivating for movement.**

Therapy:
- After the study period, participants compared the exergame sessions with conventional therapy in a **similar** way **to before the intervention**.
- Participants mentioned that conventional therapy was **more individual.**

Integration of exergames into therapy:
- Participants would welcome a **combination of both training approaches.**

Discussion

Overview

This project aimed to contribute specifically to (1) develop research-based, iterative, and co-designed user-centered exergames for patients with MS and (2) determine the usability and feasibility of the newly developed exergames. This was only possible by incorporating the theoretical background from human movement sciences, neuropsychology, and game research, as well as practical skills from game design. Furthermore, this iterative and participatory design process was carried out in close collaboration with patients with MS and their therapists.

In the following sections, the quantitative and qualitative results of the user studies are discussed and set in the context of related work and knowledge in game research and movement science, as well as research in the field of MS. Quantitative and qualitative data revealed certain exergame elements that are specific to patients with MS and can become key features for the further development of user-centered exergames for this heterogeneous target group. An outlook on future approaches in user-centered MS-specific exergame development and research will be provided.

Shift of Focus

After the second study, patients often reported a *shift in their focus* from the physical to the cognitive level when playing exergames. Some patients even reported a shift of focus from their impairments to their actual skills and abilities, which they found to increase over the period studied. A study in older adults showed that exergame training increased the participants' confidence and research connected this confidence with increased *self-efficacy* [91,92]. One participant could even use advanced training methods in his regular therapy at the end of the second study. The exergames allowed the patients a sense of control over their tasks, as described by Sweetser and Wyeth [53]. The more familiar participants became with the exergames and the more they trained their own gameplay strategies and body movements, the more secure, confident, immersed, absorbed, and "in the flow" they became with the exergame. The flow feeling was described not only in the interviews but also in the FSS and Game Flow questionnaire, illustrated by a high rating in several questionnaire items as well as by a significant decrease in the questionnaire item *perceived importance*. The decreased *perceived importance* item seems to indicate that the gaming challenge of the exergames was more enjoyable, as patients have attached less importance to the gaming outcomes [93]. This might have been caused by the shift of focus, the increased sense of control, the familiarization process, and higher flow feeling. Furthermore, some patients reported that gaming time distracted them from their daily-life problems and their MS-related impairments. This is in line with the findings of related studies [53,91,94].

Heterogeneity of Patients With MS

The heterogeneity of patients with MS, including the individual course of the disease (eg, wide range of symptoms and unpredictable flare-ups), as well as demographic details (eg, wide range of age), was also reflected in the interviews. Patients reported that game content, challenge, and progression should always be adaptable to their *individual physical, cognitive, and mental requirements and their daily form* [51,72,95]. Therefore, an exergame for patients with MS should allow an individually adaptive training focus, taking into account physical, cognitive, and mental aspects, to correspond with the heterogeneity and fluctuations of the disease pattern. The exergames covered 3 different types of game control and content; each exergame included 3 levels for motor and cognitive functions. Another relevant aspect is *security*, especially in therapeutic environments [72]. In this project, the patients could use the handrail to support exergame performance due to the insecurity of their physical stability and capacity. This security support was greatly appreciated, as presented in the results of our study. Overall, no adverse events were recorded during the entire duration of the project. The wide age range in patients with MS brings very different previous experiences in using technology [96]. Thus, even for older adult patients without previous technology use, the exergames need to be *self-explanatory and easy to use* (including help from therapists). In terms of the system's usability in the heterogeneous study group, the iterative, participative, and interdisciplinary design process of this project was very successful as the SUS increased from study 1 to study 2. In study 2, the SUS dropped from pre- to postmeasurement. The novelty of the exergame might have distracted patients' focus away from the usability barriers, explaining the high SUS score at the premeasurement. Furthermore, with each additional session, participants had more time to test the system and explore usability barriers. Nevertheless, the SUS in study 2 remained at a level that can be described as a usable exergame system for patients with MS [86,97].

Training Motivation and Challenges

Most patients were motivated to train by exergames and enjoyed the requirement of physical activity for playing them. This is in line with a previous study that interviewed patients with MS about Nintendo Wii Fit [39]. However, due to the user-centered development steps and therapy focus, it may be that the training motivation was even higher than in studies that used conventional exergames [48]. One of the main *motivational drivers* was to improve the player's body functions, to be immersed in another world, and to be distracted from daily life for the duration of the exergame session [53]. Interestingly, interviews revealed that most patients preferred more challenging games (but still not overchallenging). This was also reflected in the number of sessions in which patients chose to play the most challenging exergames. This challenging situation, in combination with the skill balance of the exergames, may have facilitated the abovementioned flow state during the training sessions [98]. Exergames should provide individually challenging but still feasible gaming experiences to increase training motivation and therefore possible training-related improvements [53]. To maintain their motivation, patients also wished for more challenging and different games or levels over time in future trials.

Training Intensity and Progress

An exergame should be able to adapt to the individual patient at a physical and cognitive level to meet the heterogeneous and individual requirements of patients with MS and to allow for an *optimal training zone* [51,95]. For this reason, the design integrated individual levels for physical and cognitive functions into the exergames, allowing for an individually challenging game for patients with MS. To extend the playfulness and effectiveness of the exergames in the future, the assessment of certain motor and cognitive parameters (objective) or rating scales (subjective) could help to define an individual training area [99-102]. The integration of in-exergame, real-time adaptation could help to maintain a predefined optimal training zone in a training session and over a longer period (progression) [51,102,103]. In this project, participants had to rate each training session for physical and cognitive perceived exertion, allowing the training load to be adapted for the upcoming sessions. The results of the perceived exertion ratings showed that the newly developed exergames allowed for a moderate training load on the cognitive and physical levels in a single training session and over the training period. For aerobic and strength exercises, moderate training is recommended in patients with MS [104]. However, it is possible that high training intensities, such as those used in high-intensity interval training, might be even more beneficial [105]. Nevertheless, a *moderate training intensity* seems to be an appropriate approach for exergames to trigger possible motor learning processes without negatively influencing movement execution in patients with MS [68,81].

Exergame as an MS Therapy Tool

Interviews showed a strong acceptance of the exergames by patients (even in the first study). The majority would *welcome the integration* of exergames into their conventional therapy because of their appealing nature and beneficial motor-cognitive training approach [40,106]. The combined training regimen allows for the concurrent processing and synchronization of cognitive and motor stimuli and therefore can trigger brain-body communication. Patients with MS can have impaired dual- or multi-task performance due to possible deficits in divided attention, resource capacity overload, or differential neural activation [76-80]. Furthermore, exergames allow the integration of the patient's conventional therapy progress in physical and cognitive functions and provide a daily-life environment in terms of the combined cognitive-motor training. However, some patients missed the *social component* and *interaction with the therapists*. Therefore, it might be interesting to specifically integrate the therapist(s) into the exergame experience by in-exergame interaction, allowing training adaptation and support. This finding is in line with recent exergame studies in patients with MS and older adults that emphasize the importance of social interaction in exergames to increase training motivation [91,103]. Moreover, social interaction is a part of the Game Flow model proposed by Sweetser and Wyeth [53]. Overall, user-centered exergames seem to be a very promising therapy tool for patients with MS, considering the abovementioned aspects of training and design principles.

As a next step, further research and development work will deepen the knowledge of design principles in MS exergames and reveal additional insights. To meet the heterogeneous spectrum of MS and to provide an individually attractive and effective training and therapy tool, the newly developed exergames will be further iterated and extended based on the findings of the usability and feasibility studies. Furthermore, new types of use will be implemented, such as playing a multitask version of the exergames that involve upper-body input movements or sitting in a wheelchair. Moreover, further balancing game mechanics will be implemented, as well as extending the types of input, movement ranges, and tracking zone.

Limitations

There are some limitations that can be reported for this study. In the first study, participants were trained only once with the exergames, whereas they trained multiple times in the second study. Therefore, participants might have had the chance to reflect more on and better familiarize themselves with the games in the second study, while they had only one attempt in the first study. Additionally, their feedback might have been influenced by the novelty effect. Furthermore, study testing was conducted at various clinics and institutions and it did not focus on measures of effectiveness. However, it should be emphasized that these studies should be conducted in the context of developing a complex intervention for health care settings. Within this context, intervention development contains different mandatory steps that should be taken in a sequential order [107]. In that sense, this study reflects a preintervention stage in which important principles and necessary actions for this stage were considered [107]. These findings justify continuing with studies that focus on the outputs and effects in clinical trials [107].

Conclusions

The aim of the presented research and development work was to take the first step in the new field of user-centered exergames for patients with MS, to evaluate the usability and feasibility of the newly developed exergame concepts, to learn from the findings, and to derive design guidelines for future research and development projects in this field.

The quantitative and qualitative results of this project showed that the developed exergames were usable, feasible, well accepted, and enjoyable for patients with MS. Furthermore, the results indicated preliminary positive effects regarding the attractiveness of the newly developed, user-centered exergames. Participants enjoyed the motivating, varied, and fun experience with the exergames, which were both fun and physically as well as cognitively challenging and allowing them to forget their everyday worries (often associated with the disease) for the moment. Moreover, specific exergame elements were identified: control mechanisms through audio-visual design, adaptation of the individual difficulty level, game concept diversity addressing the patients' heterogeneity, involvement of training principles, and considerations of the interaction of physical and cognitive impairments, especially brain-body communication.

Considering the points of discussion and design guidelines, user-centered exergames can be a promising training approach to improve physical and cognitive functions, especially brain-body communication in patients with MS. Thus, user-centered exergames might have positive effects on quality of life by reducing the risk of falling, mobility restrictions, and social isolation. Furthermore, the strengthening of body functions such as balance, coordination, and cognition seems to be a promising way to break the vicious circle of deconditioning. The evaluation of the effects of a user-centered exergame will show how far a user-centered exergame might complement or even surpass the results of conventional (exergame) approaches in patients with MS.

Acknowledgments

The authors want to thank the Swiss Innovation Agency Innosuisse for funding the project (grant number: 41968.1 IP-LS). Furthermore, they thank their postgraduate student Tiziana Schwarz for instructing trainings and helping with data acquisition. Finally, the authors thank all the patients with MS for their participation in this project and the physiotherapists who supported the studies.

Authors' Contributions

AMN and AS conceptualized, designed, and drafted the manuscript. EDB and SF contributed substantially to the conception and design of the manuscript. AMN, AS, SB, SH, and YH created the study design, compiled the training protocols, and selected the assessment methods for the first study. SH conducted the study (supervised by AMN, AS, and RS). For the second study, AMN, AM, AS, BF, SB, and YH created the study design, compiled the training protocols, and selected the assessment methods. AM and BF conducted the study (supervised by AMN, AS, RS, and SF). SB and YH designed the exergame environments for both studies (supported by AMN, RB, and UG). AMN and AS led data analysis and interpretation; EDB and SF contributed to the latter. All authors critically reviewed and approved the final manuscript.

Conflicts of Interest

EDB was a cofounder of Dividat, the spin-off company that developed the exergame plate used in this study, and is associated with the company as an external advisor. No revenue was paid (or promised to be paid) directly to EDB or his institution over the 36 months before the submission of the work.

References

1. Wallin MT, Culpepper WJ, Nichols E, Bhutta ZA, Gebrehiwot TT, Hay SI, et al. Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology 2019 Mar;18(3):269-285. [doi: 10.1016/s1474-4422(18)30443-5]
2. Hemmer B, Nessler S, Zhou D, Kieseier B, Hartung H. Immunopathogenesis and immunotherapy of multiple sclerosis. Nat Rev Neurol 2006 Apr;2(4):201-211. [doi: 10.1038/ncpneuro0154]
3. Hauser SL, Chan JR, Oksenberg JR. Multiple sclerosis: Prospects and promise. Ann Neurol 2013 Oct 09;74(3):317-327. [doi: 10.1002/ana.24009]
4. Filippi M, Bar-Or A, Piehl F, Preziosa P, Solari A, Vukusic S, et al. Multiple Sclerosis. Nat Rev Dis Primers 2018 Nov 08;4(43):1-27. [doi: 10.1038/s41572-018-0046-z]
5. Hauser S, Oksenberg L, Baranzini S. Multiple Sclerosis. In: Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease (5th Edition). London: Academic Press Elsevier; 2015:1001-1014.
6. Compston A, McDonald I, Noseworthy J, Lassmann H, Miller D, Smith K, et al. McAlpine's Multiple Sclerosis (4th Edition). London: Churchill Livingstone Elsevier; 2005.
7. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. The Lancet Neurology 2008 Dec;7(12):1139-1151. [doi: 10.1016/s1474-4422(08)70259-x]
8. Benito-León J, Manuel Morales J, Rivera-Navarro J, Mitchell AJ. A review about the impact of multiple sclerosis on health-related quality of life. Disability and Rehabilitation 2003;25(23):1291-1303. [doi: 10.1080/09638280310001608591]
9. Motl RW, McAuley E. Symptom Cluster and Quality of Life: Preliminary Evidence in Multiple Sclerosis. Journal of Neuroscience Nursing 2010;42(4):212-216. [doi: 10.1097/jnn.0b013e3181e26c5f]
10. Stolze H, Klebe S, Zechlin C, Baecker C, Friege L, Deuschl G. Falls in frequent neurological diseases. Journal of Neurology 2004;251(1):79-84. [doi: 10.1007/s00415-004-0276-8]
11. Beghi E, Gervasoni E, Pupillo E, Bianchi E, Montesano A, Aprile I, et al. Prediction of Falls in Subjects Suffering From Parkinson Disease, Multiple Sclerosis, and Stroke. Arch Phys Med Rehabil 2018;99(4):641-651. [doi: 10.1016/j.apmr.2017.10.009] [Medline: 29102438]
12. Sosnoff J, Gappmaier E, Frame A, Motl R. Influence of spasticity on mobility and balance in persons with multiple sclerosis. J Neurol Phys Ther 2011;35(3):129-132. [doi: 10.1097/NPT.0b013e31822a8c40] [Medline: 21934374]
13. Flachenecker P, Henze T, Zettl UK. Spasticity in patients with multiple sclerosis--clinical characteristics, treatment and quality of life. Acta Neurol Scand 2014;129(3):154-162. [doi: 10.1111/ane.12202] [Medline: 24256407]
14. Gullo HL, Fleming J, Bennett S, Shum DH. Cognitive and physical fatigue are associated with distinct problems in daily functioning, role fulfilment, and quality of life in multiple sclerosis. Mult Scler Relat Disord 2019;31:118-123. [doi: 10.1016/j.msard.2019.03.024] [Medline: 30981190]
15. Motl R. Physical activity and irreversible disability in multiple sclerosis. Exerc Sport Sci Rev 2010;38(4):186-191. [doi: 10.1097/JES.0b013e3181f44fab] [Medline: 20871235]
16. Motl RW, McAuley E, Snook EM. Physical activity and multiple sclerosis: a meta-analysis. Mult Scler 2005;11(4):459-463. [doi: 10.1191/1352458505ms1188oa] [Medline: 16042230]
17. Kobelt G. Health economic issues in MS. Int MS J 2006;13(1):17-26, 16. [Medline: 16420781]

18. Kobelt G, Berg J, Lindgren P, Fredrikson S, Jönsson B. Costs and quality of life of patients with multiple sclerosis in Europe. J Neurol Neurosurg Psychiatry 2006;77(8):918-926 [FREE Full text] [doi: 10.1136/jnnp.2006.090365] [Medline: 16690691]
19. Latimer-Cheung AE, Pilutti LA, Hicks AL, Martin Ginis KA, Fenuta AM, MacKibbon KA, et al. Effects of Exercise Training on Fitness, Mobility, Fatigue, and Health-Related Quality of Life Among Adults With Multiple Sclerosis: A Systematic Review to Inform Guideline Development. Archives of Physical Medicine and Rehabilitation 2013;94(9):1800-1828.e3. [doi: 10.1016/j.apmr.2013.04.020]
20. Motl RW, Pilutti LA. The benefits of exercise training in multiple sclerosis. Nat Rev Neurol 2012;8(9):487-497. [doi: 10.1038/nrneurol.2012.136]
21. Pilutti LA, Platta ME, Motl RW, Latimer-Cheung AE. The safety of exercise training in multiple sclerosis: A systematic review. Journal of the Neurological Sciences 2014;343(1-2):3-7. [doi: 10.1016/j.jns.2014.05.016]
22. Dardiotis E, Nousia A, Siokas V, Tsouris Z, Andravizou A, Mentis AA, et al. Efficacy of computer-based cognitive training in neuropsychological performance of patients with multiple sclerosis: A systematic review and meta-analysis. Multiple Sclerosis and Related Disorders 2018;20:58-66. [doi: 10.1016/j.msard.2017.12.017]
23. Hancock LM, Bruce JM, Bruce AS, Lynch SG. Processing speed and working memory training in multiple sclerosis: A double-blind randomized controlled pilot study. Journal of Clinical and Experimental Neuropsychology 2015;37(2):113-127. [doi: 10.1080/13803395.2014.989818]
24. Bonavita S, Sacco R, Della Corte M, Esposito S, Sparaco M, d'Ambrosio A, et al. Computer-aided cognitive rehabilitation improves cognitive performances and induces brain functional connectivity changes in relapsing remitting multiple sclerosis patients: an exploratory study. J Neurol 2015;262(1):91-100. [doi: 10.1007/s00415-014-7528-z] [Medline: 25308631]
25. 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]
26. Oh Y, Yang S. Proceedings of Meaningful Play 2010:1-17.
27. Witherspoon L. ACSM Information on Exergaming. American College of Sports Medicine 2013:1.
28. Mueller F, Khot RA, Gerling K, Mandryk R. Exertion Games. FNT in Human–Computer Interaction 2016;10(1):1-86. [doi: 10.1561/1100000041]
29. Martin-Niedecken A, Mekler E. The ExerCube: Participatory Design of an Immersive Fitness Game Environment. 2018 Presented at: Joint International Conference on Serious Games; 2018; Darmstadt, Germany. [doi: 10.1007/978-3-030-02762-9_28]
30. Martin A, Götz U, Bauer R. Development of task-specific RehabGame settings for robot-assisted pédiatrie movement therapies. 2014 Presented at: IEEE Games Media Entertainment; 2014; Toronto, ON, Canada. [doi: 10.1109/GEM.2014.7048090]
31. Martin A, Götz U, Bauer R. "Gabarello v.1.0" and "Gabarello v.2.0": Development of motivating rehabilitation games for robot-assisted locomotion therapy in childhood. 2014 Presented at: Games for Health; 2014; Boston p. 101-104. [doi: 10.1007/978-3-658-07141-7_13]
32. Kajastila R, Hämäläinen P. Motion games in real sports environments. interactions 2015 Feb 25;22(2):44-47. [doi: 10.1145/2731182]
33. Martin-Niedecken A, Marquez Segura E, Rogers K, Niedecken S, Turmo Vidal L. Towards Socially Immersive Fitness Games: An Exploratory Evaluation Through Embodied Sketching. 2019 Presented at: CHI PLAY; 2019; Barcelona p. 525-534. [doi: 10.1145/3341215.3356293]
34. Martin-Niedecken A, Rogers K, Turmo Vidal L, Mekler E, Marquez Segura E. ExerCube vs. Personal Trainer: Evaluating a Holistic, Immersive, and Adaptive Fitness Game Setup. 2019 Presented at: CHI Conference on Human Factors in Computing Systems; 2019; Glasgow p. 1-15. [doi: 10.1145/3290605.3300318]
35. Stojan R, Voelcker-Rehage C. A Systematic Review on the Cognitive Benefits and Neurophysiological Correlates of Exergaming in Healthy Older Adults. JCM 2019;8(5):734. [doi: 10.3390/jcm8050734]
36. Sween J, Wallington S, Sheppard V, Taylor T, Llanos A, Adams-Campbell L. The role of exergaming in improving physical activity: a review. J Phys Act Health 2014;11(4):864-870 [FREE Full text] [doi: 10.1123/jpah.2011-0425] [Medline: 25078529]
37. Byrne AM, Kim M. The Exergame as a Tool for Mental Health Treatment. Journal of Creativity in Mental Health 2019;14(4):465-477. [doi: 10.1080/15401383.2019.1627263]
38. Thomas S, Fazakarley L, Thomas P, Collyer S, Brenton S, Perring S, et al. Mii-vitaliSe: a pilot randomised controlled trial of a home gaming system (Nintendo Wii) to increase activity levels, vitality and well-being in people with multiple sclerosis. BMJ Open 2017;7(9):1-16 [FREE Full text] [doi: 10.1136/bmjopen-2017-016966] [Medline: 28954791]
39. Forsberg A, Nilsagård Y, Boström K. Perceptions of using videogames in rehabilitation: a dual perspective of people with multiple sclerosis and physiotherapists. Disabil Rehabil 2015;37(4):338-344 [FREE Full text] [doi: 10.3109/09638288.2014.918196] [Medline: 24833535]

40. Mura G, Carta M, Sancassiani F, Machado S, Prosperini L. Active exergames to improve cognitive functioning in neurological disabilities: a systematic review and meta-analysis. European Journal of Physical and Rehabilitation Medicine 2017;54(3):450-462. [doi: 10.23736/s1973-9087.17.04680-9]
41. Kramer A, Dettmers C, Gruber M. Exergaming With Additional Postural Demands Improves Balance and Gait in Patients With Multiple Sclerosis as Much as Conventional Balance Training and Leads to High Adherence to Home-Based Balance Training. Archives of Physical Medicine and Rehabilitation 2014;95(10):1803-1809. [doi: 10.1016/j.apmr.2014.04.020]
42. Nilsagård Y, Forsberg A, von Koch L. Balance exercise for persons with multiple sclerosis using Wii games: a randomised, controlled multi-centre study. Mult Scler 2013;19(2):209-216 [FREE Full text] [doi: 10.1177/1352458512450088] [Medline: 22674972]
43. Prosperini L, Fortuna D, Giannì C, Leonardi L, Marchetti MR, Pozzilli C. Home-Based Balance Training Using the Wii Balance Board. Neurorehabil Neural Repair 2013;27(6):516-525. [doi: 10.1177/1545968313478484]
44. Brichetto G, Spallarossa P, de Carvalho MLL, Battaglia MA. The effect of Nintendo® Wii® on balance in people with multiple sclerosis: a pilot randomized control study. Mult Scler 2013;19(9):1219-1221. [doi: 10.1177/1352458512472747]
45. Prosperini L, Fanelli F, Petsas N, Sbardella E, Tona F, Raz E, et al. Multiple Sclerosis: Changes in Microarchitecture of White Matter Tracts after Training with a Video Game Balance Board. Radiology 2014;273(2):529-538. [doi: 10.1148/radiol.14140168]
46. Robinson J, Dixon J, Macsween A, van Schaik P, Martin D. The effects of exergaming on balance, gait, technology acceptance and flow experience in people with multiple sclerosis: a randomized controlled trial. BMC Sports Sci Med Rehabil 2015;7(1):8 [FREE Full text] [doi: 10.1186/s13102-015-0001-1] [Medline: 25969739]
47. Stanmore E, Stubbs B, Vancampfort D, de Bruin ED, Firth J. The effect of active video games on cognitive functioning in clinical and non-clinical populations: A meta-analysis of randomized controlled trials. Neuroscience & Biobehavioral Reviews 2017;78:34-43. [doi: 10.1016/j.neubiorev.2017.04.011]
48. Taylor M, Griffin M. The use of gaming technology for rehabilitation in people with multiple sclerosis. Mult Scler 2014;21(4):355-371. [doi: 10.1177/1352458514563593]
49. Pau M, Coghe G, Corona F, Leban B, Marrosu MG, Cocco E. Effectiveness and Limitations of Unsupervised Home-Based Balance Rehabilitation with Nintendo Wii in People with Multiple Sclerosis. Biomed Res Int 2015:1-8 [FREE Full text] [doi: 10.1155/2015/916478] [Medline: 26583146]
50. Wüest S, van de Langenberg R, de Bruin ED. Design considerations for a theory-driven exergame-based rehabilitation program to improve walking of persons with stroke. Eur Rev Aging Phys Act 2014;11(2):119-129 [FREE Full text] [doi: 10.1007/s11556-013-0136-6] [Medline: 25309631]
51. Hardy S, Dutz T, Wiemeyer J, Göbel S, Steinmetz R. Framework for personalized and adaptive game-based training programs in health sport. Multimed Tools Appl 2014;74(14):5289-5311. [doi: 10.1007/s11042-014-2009-z]
52. Benzing V, Schmidt M. Exergaming for Children and Adolescents: Strengths, Weaknesses, Opportunities and Threats. J Clin Med 2018;7(11):422 [FREE Full text] [doi: 10.3390/jcm7110422] [Medline: 30413016]
53. Sweetser P, Wyeth P. GameFlow: a model for evaluating player enjoyment in games. Computers in Entertainment 2005;3(3):3-3. [doi: 10.1145/1077246.1077253]
54. Hoffmann K. , J. Wiemeyer, and S. Hardy, Prediction and control of the individual Heart Rate response in Exergames. 2016 Presented at: Proceedings of the 10th International Symposium on Computer Science in Sports (Iscss); 2016; Loughborough, UK p. 171-178. [doi: 10.1007/978-3-319-24560-7_22]
55. Sinclair J, Hingston P, Masek M. Considerations for the design of exergames. 2007 Presented at: GRAPHITE '07: Proceedings of the 5th international conference on Computer graphics and interactive techniques; 2007; Australia and Southeast Asia p. 289-295. [doi: 10.1145/1321261.1321313]
56. Segura E, Turmo Vidal L, Rostami A, Waern A. Embodied Sketching. 2016 Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; 2016; San Jose, CA, USA p. 6014-6027. [doi: 10.1145/2858036.2858486]
57. Isbister K, Mueller. Guidelines for the Design of Movement-Based Games and Their Relevance to HCI. Human–Computer Interaction 2014;30(3-4):366-399. [doi: 10.1080/07370024.2014.996647]
58. IJsselsteijn W, de Kort Y, Poels K, Jurgelionis A, Bellotti F. Characterising and Measuring User Experiences in Digital Games. 2007 Presented at: International conference on advances in computer entertainment technology; 2007; Salzburg Austria.
59. Gerling K, Schild J, Masuch M. Exergame Design for Elderly Users: The Case Study of SilverBalance. 2010 Presented at: ACE '10: Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology; 2010; Taipei Taiwan p. 66-69. [doi: 10.1145/1971630.1971650]
60. Gibson A, Wagner D, Heyward V. Advanced Fitness Assessment and Exercise Prescription. Champaign, IL: Human kinetics; 2019.
61. Hoffman J. Physiological aspects of sport training and performance. Champaign, IL: Human kinetics; 2014.
62. Healy AF, Kole JA, Bourne LE. Training principles to advance expertise. Front Psychol 2014;5:131 [FREE Full text] [doi: 10.3389/fpsyg.2014.00131] [Medline: 24600425]

63. Martin-Niedecken A, Götz U. Design and Evaluation of a Dynamically Adaptive Fitness Game Environment for Children and Young Adolescents. 2016 Presented at: CHI PLAY Companion '16; 2016; Austin Texas USA p. 205-212. [doi: 10.1145/2968120.2987720]
64. Martin-Niedecken A, Götz U. Go with the Dual Flow: Evaluating the Psychophysiological Adaptive Fitness Game Environment “Plunder Planet”. 2017 Presented at: Joint International Conference on Serious Games; 2017; Valencia, Spain p. 32-43. [doi: 10.1007/978-3-319-70111-0_4]
65. Martin-Niedecken A. Designing for bodily interplay: engaging with the adaptive social exertion game "plunder planet". 2018 Presented at: Interaction Design and Children; 2018; Trondheim, Norway p. 19-30. [doi: 10.1145/3202185.3202740]
66. Read JL, Shortell SM. Interactive games to promote behavior change in prevention and treatment. JAMA 2011;305(16):1704-1705. [doi: 10.1001/jama.2011.408] [Medline: 21447802]
67. Martin-Niedecken A, Rogers K, Turmo Vidal L, Mekler E, Marquez Segura E. ExerCube vs. Personal Trainer: Evaluating a Holistic, Immersive, and Adaptive Fitness Game Setup. 2019 Presented at: CHI; 2019; Glasgow Scotland Uk p. 1-15. [doi: 10.1145/3290605.3300318]
68. Bacmeister C, Barr H, McClain C, Thornton M, Nettles D, Welle C, et al. Motor learning promotes remyelination via new and surviving oligodendrocytes. Nat Neurosci 2020;23(7):819-831 [FREE Full text] [doi: 10.1038/s41593-020-0637-3] [Medline: 32424285]
69. Pasch M, Bianchi-Berthouze N, van Dijk B, Nijholt A. Movement-based sports video games: Investigating motivation and gaming experience. Entertainment Computing 2009;1(2):49-61. [doi: 10.1016/j.entcom.2009.09.004]
70. Shafer DM, Carbonara CP, Popova L. Controller Required? The Impact of Natural Mapping on Interactivity, Realism, Presence, and Enjoyment in Motion-Based Video Games. Presence: Teleoperators and Virtual Environments 2014;23(3):267-286. [doi: 10.1162/pres_a_00193]
71. Denisova A, Caims P. First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion? 2015 Presented at: CHI; 2015; Seoul Republic of Korea p. 145-148. [doi: 10.1145/2702123.2702256]
72. Wiemeyer J, Deutsch J, Malone LA, Rowland JL, Swartz MC, Xiong J, et al. Recommendations for the Optimal Design of Exergame Interventions for Persons with Disabilities: Challenges, Best Practices, and Future Research. Games for Health Journal 2015;4(1):58-62. [doi: 10.1089/g4h.2014.0078]
73. Hauser SL, Oksenberg JR. The Neurobiology of Multiple Sclerosis: Genes, Inflammation, and Neurodegeneration. Neuron 2006;52(1):61-76. [doi: 10.1016/j.neuron.2006.09.011]
74. White LJ, Dressendorfer RH. Exercise and Multiple Sclerosis. Sports Medicine 2004;34(15):1077-1100. [doi: 10.2165/00007256-200434150-00005]
75. Sumowski JF, Benedict R, Enzinger C, Filippi M, Geurts JJ, Hamalainen P, et al. Cognition in multiple sclerosis: State of the field and priorities for the future. Neurology 2018;90(6):278-288. [doi: 10.1212/wnl.0000000000004977]
76. Hamilton F, Rochester L, Paul L, Rafferty D, O'Leary C, Evans J. Walking and talking: an investigation of cognitive—motor dual tasking in multiple sclerosis. Mult Scler 2009;15(10):1215-1227. [doi: 10.1177/1352458509106712]
77. Leone C, Patti F, Feys P. Measuring the cost of cognitive-motor dual tasking during walking in multiple sclerosis. Mult Scler 2015;21(2):123-131. [doi: 10.1177/1352458514547408] [Medline: 25178543]
78. Wajda DA, Motl RW, Sosnoff JJ. Dual task cost of walking is related to fall risk in persons with multiple sclerosis. Journal of the Neurological Sciences 2013;335(1-2):160-163. [doi: 10.1016/j.jns.2013.09.021]
79. Learmonth YC, Ensari I, Motl RW. Cognitive Motor Interference in Multiple Sclerosis: Insights From a Systematic Quantitative Review. Archives of Physical Medicine and Rehabilitation 2017;98(6):1229-1240. [doi: 10.1016/j.apmr.2016.07.018]
80. Wajda DA, Sosnoff JJ. Cognitive-motor interference in multiple sclerosis: a systematic review of evidence, correlates, and consequences. Biomed Res Int 2015:1-8 [FREE Full text] [doi: 10.1155/2015/720856] [Medline: 25839039]
81. Feter N, Freitas M, Gonzales N, Umpierre D, Cardoso R, Rombaldi A. Effects of physical exercise on myelin sheath regeneration: A systematic review and meta-analysis. Science & Sports 2018;33(1):8-21. [doi: 10.1016/j.scispo.2017.06.009]
82. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001;124(5):962-973. [doi: 10.1093/brain/124.5.962] [Medline: 11335698]
83. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) Scale. The Journals of Gerontology 1995;50A(1):M28-M34. [doi: 10.1093/gerona/50a.1.m28] [Medline: 7814786]
84. Brooke J. SUS-A quick and dirty usability scale. Usability evaluation in industry 1996;189(194):4-7. [doi: 10.1201/9781498710411-35]
85. Tullis T, Albert W. Measuring the user experience: collecting, analyzing, and presenting usability metrics. Amsterdam: Elsevier Inc; 2013.
86. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
87. Rheinberg F, Vollmeyer R, Engeser S. Die Erfassung des Flow-Erlebens. In: Diagnostik von Motivation und Selstkonzept. Göttingen: Hogrefe; 2003:261-279.
88. Borg G. Borg's perceived exertion and pain scales. In: Borg's perceived exertion and pain scales. Champaign, IL: Human kinetics; 1998.

89. Cohen J, Cohen P, West S, Alken L. Applied multiple regression/correlation analysis for the behavioral sciences. Mahwah, New Jersey: Lawrence Erlbaum Association; 2003.
90. Mayring P. Qualitative Content Analysis: Theoretical Background and Procedures. In: Approaches to Qualitative Research in Mathematics Education. Advances in Mathematics Education. Dordrecht: Springer; 2015:365-380.
91. Meekes W, Stanmore EK. Motivational Determinants of Exergame Participation for Older People in Assisted Living Facilities: Mixed-Methods Study. J Med Internet Res 2017;19(7):e238. [doi: 10.2196/jmir.6841]
92. Schutzer KA, Graves BS. Barriers and motivations to exercise in older adults. Prev Med 2004;39(5):1056-1061. [doi: 10.1016/j.ypmed.2004.04.003] [Medline: 15475041]
93. Engeser S, editor. Advances in Flow Research. New York: Springer Science+Business Media; 2012.
94. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, et al. Measuring and defining the experience of immersion in games. International Journal of Human-Computer Studies 2008;66(9):641-661. [doi: 10.1016/j.ijhcs.2008.04.004]
95. Streicher A, Smeddinck J. Personalized and Adaptive Serious Games. In: Entertainment Computing and Serious Games. Cham: Springer; 2016:332-377.
96. O'brien MA, Rogers WA, Fisk AD. Understanding age and technology experience differences in use of prior knowledge for everyday technology interactions. ACM Trans. Access. Comput 2012;4(2):1-27. [doi: 10.1145/2141943.2141947]
97. Bangor A. , P. Kortum, and J. Miller, Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies 2009;4(3):114-123.
98. Jackson S, Csikszentmihalyi M. Flow in sports: The keys to optimal experiences and performances. Champaign, IL: Human kinetics; 1999.
99. Meckel Y, Zach S, Eliakim A, Sindiani M. The interval-training paradox: Physiological responses vs. subjective rate of perceived exertion. Physiology & Behavior 2018;196:144-149. [doi: 10.1016/j.physbeh.2018.08.013]
100. Vasilyev V, Borisov V, Syskov A. Biofeedback Methodology: a Narrative Review. 2019 Presented at: International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON); 2019; Novosibirsk, Russia. [doi: 10.1109/sibircon48586.2019.8958019]
101. Smith M, Fransen R, Coutts A. Inducing and assessing cognitive fatigue. 2014 Presented at: European College of Sport Science Annual Congress; 2014; Amsterdam.
102. Martin-Niedecken A, Schwarz T, Schättin A. Comparing Training Intensity and Experience of an Adaptive Exergame-Based Functional HIIT in the ExerCube Based on Individually Determined Versus Calculated Heart Rate in Healthy Young Adults. Frontiers in Psychology 2021:1-14 (forthcoming).
103. Octavia JR, Coninx K. Adaptive Personalized Training Games for Individual and Collaborative Rehabilitation of People with Multiple Sclerosis. BioMed Research International 2014;2014:1-22. [doi: 10.1155/2014/345728]
104. Kim Y, Lai B, Mehta T, Thirumalai M, Padalabalanarayanan S, Rimmer JH, et al. Exercise Training Guidelines for Multiple Sclerosis, Stroke, and Parkinson Disease. Am J Phys Med Rehabil 2019;98(7):613-621. [doi: 10.1097/phm.0000000000001174]
105. Campbell E, Coulter EH, Paul L. High intensity interval training for people with multiple sclerosis: A systematic review. Multiple Sclerosis and Related Disorders 2018;24:55-63. [doi: 10.1016/j.msard.2018.06.005]
106. Maggio MG, Russo M, Cuzzola MF, Destro M, La Rosa G, Molonia F, et al. Virtual reality in multiple sclerosis rehabilitation: A review on cognitive and motor outcomes. Journal of Clinical Neuroscience 2019;65:106-111. [doi: 10.1016/j.jocn.2019.03.017]
107. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019;9(8):e029954. [doi: 10.1136/bmjopen-2019-029954]

Abbreviations

CNS: central nervous system
FSS: Flow Short Scale
MS: multiple sclerosis
SUS: System Usability Scale

Edited by N Zary; submitted 25.09.20; peer-reviewed by C Werner, L Becker; comments to author 08.12.20; revised version received 03.02.21; accepted 12.02.21; published 07.05.21.

Please cite as:
Schättin A, Häfliger S, Meyer A, Früh B, Böckler S, Hungerbühler Y, de Bruin ED, Frese S, Steinlin Egli R, Götz U, Bauer R, Martin-Niedecken AL
Design and Evaluation of User-Centered Exergames for Patients With Multiple Sclerosis: Multilevel Usability and Feasibility Studies
JMIR Serious Games 2021;9(2):e22826
URL: https://games.jmir.org/2021/2/e22826
doi:10.2196/22826
PMID:33960956

Original Paper

A Prediction Model for Detecting Developmental Disabilities in Preschool-Age Children Through Digital Biomarker-Driven Deep Learning in Serious Games: Development Study

Ho Heon Kim[1*], RN, MS; Jae Il An[1*], MS; Yu Rang Park[1], PhD

Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

*these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics
Yonsei University College of Medicine
50-1 Yonsei-ro
Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2228 2493
Email: yurangpark@yuhs.ac

Abstract

Background: Early detection of developmental disabilities in children is essential because early intervention can improve the prognosis of children. Meanwhile, a growing body of evidence has indicated a relationship between developmental disability and motor skill, and thus, motor skill is considered in the early diagnosis of developmental disability. However, there are challenges to assessing motor skill in the diagnosis of developmental disorder, such as a lack of specialists and time constraints, and thus it is commonly conducted through informal questions or surveys to parents.

Objective: This study sought to evaluate the possibility of using drag-and-drop data as a digital biomarker and to develop a classification model based on drag-and-drop data with which to classify children with developmental disabilities.

Methods: We collected drag-and-drop data from children with typical development and developmental disabilities from May 1, 2018, to May 1, 2020, via a mobile application (DoBrain). We used touch coordinates and extracted kinetic variables from these coordinates. A deep learning algorithm was developed to predict potential development disabilities in children. For interpretability of the model results, we identified which coordinates contributed to the classification results by applying gradient-weighted class activation mapping.

Results: Of the 370 children in the study, 223 had typical development, and 147 had developmental disabilities. In all games, the number of changes in the acceleration sign based on the direction of progress both in the x- and y-axes showed significant differences between the 2 groups ($P<.001$; effect size >0.5). The deep learning convolutional neural network model showed that drag-and-drop data can help diagnose developmental disabilities, with an area under the receiving operating characteristics curve of 0.817. A gradient class activation map, which can interpret the results of a deep learning model, was visualized with the game results for specific children.

Conclusions: Through the results of the deep learning model, we confirmed that drag-and-drop data can be a new digital biomarker for the diagnosis of developmental disabilities.

(JMIR Serious Games 2021;9(2):e23130) doi:10.2196/23130

KEYWORDS

developmental delay; diagnosis prediction; deep learning; serious games; digital health; digital phenotyping; digital biomarkers

Introduction

Developmental disabilities are a set of common heterogeneous disorders developing in 10%-15% of preschool-age children and characterized by difficulties in one or more domains, including learning, behavior, and self-care [1-3]. The prevalence trend of all developmental disabilities increased from 1997 to 2017 in the United States, and the trend in low- and middle-income countries has also increased in the number of children surviving high-risk neonatal conditions from improved obstetric and neonatal care [4,5]. Although the etiology and cause of developmental disabilities are complicated and not well understood, early intervention is conventionally considered as an effective clinical treatment [6,7]. Early detection of developmental disabilities is key because early intervention can improve a child's prognosis due to rapid brain growth and neuroplasticity [8-10]. However, early detection or screening has multiple challenges, including time constraints, financial burden, scarcity of human resources, lack of consensus on the tools for the general childhood population, and diagnostic stability [11,12]. Given the phenotypical nature of developmental disabilities, the assessment processes show high variability [13,14]. Neuropsychological tests are often difficult and tedious for preschool-age children to complete, leading to inaccurate assessment [14]. Moreover, although it is important to perform continuous clinical examinations and comprehensive tracking for more accurate assessment [15-17], poor follow-up adherence rates have been reported. This low follow-up rate can induce a loss of chance for early intervention [18].

Meanwhile, a growing body of evidence has indicated a relationship between developmental disability and motor control, because the cerebellum is closely related to higher cognitive function [19]. Motor skill is considered to be a factor in the early diagnosis of developmental disability [20,21]. Despite this evidence, the measurement of motor skill requires expensive laboratory resources or clinical expertise and is not easily applicable in repeated measurements. As an alternative to measuring motor skill and without the constraints of time and place, a serious game that is able to capture upper extremity movements while touching a display could help in detecting children with developmental disabilities.

Therefore, this study aimed to identify the possibility of drag-and-drop data as a digital biomarker and to develop a classification model based on drag-and-drop data with which to classify children with developmental disabilities.

Methods

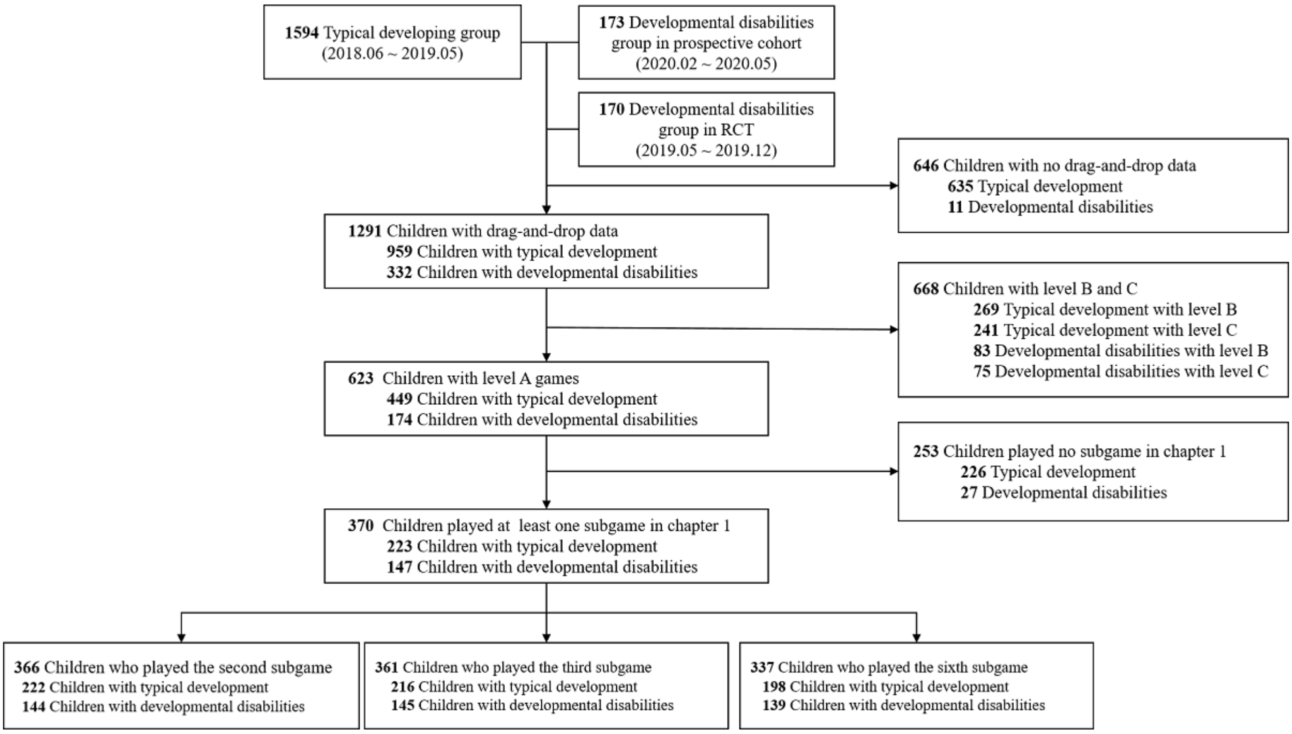
Serious Games

This study included children who had experiences with a serious game known as DoBrain (DoBrain Inc). DoBrain is a mobile-based game that provides programs for the cognitive development of children. The games of this application consist of chapters of 7 to 8 subgames targeting spatial awareness, perceptual speed, repair, creativity, reasoning, composition, memory, and visual discrimination of the cognitive area. Each subgame can be classified into a tapping game where users have to solve a problem by touching objects to answer a question or a drag-and-drop game where users have to drag and drop cartoon objects with their fingers. In addition, the difficulty of the game is divided into 3 levels (A, B, and C) depending on the cognitive level of the user. Chapter 1 comprises 7 subgames, including 4 tapping games or non–drag-and-drop games (first, fourth, fifth, and seventh) and 3 drag-and-drop games (second, third, and sixth). The second game is an imitation game in which users must infer the correct answer from similar images of an object and is designed to improve logical reasoning. The third subgame requires the user to infer the correct answer from remnant images and is designed to improve memory function. The sixth subgame is designed to improve spatial awareness by requiring the user to locate the object on the target region (Multimedia Appendix 1).

Study Design

In this retrospective study, we obtained deidentified participant data from 3 studies: (1) a retrospective study conducted between June 1, 2018, and June 1, 2020, with children having profile information in the application; (2) randomized clinical trials conducted from March 1, 2019, to December 30, 2019, for evaluation of cognitive improvement in children with developmental disabilities; and (3) a prospective study conducted from February 1, 2020, on a development classification model. The profiles of children with typical development included in the first study were entered by their parents, and the children included in the second and third studies were diagnosed by pediatric psychologists. In each study, 1594 children with typical development and 343 children with developmental disabilities (173 and 170 children, respectively) were included. Among the 1937 children with valid profiles, we also excluded children without drag-and-drop data due to server instability or unexpected shutdown of the game (n=646). Moreover, in 1291 children with drag-and-drop data, we included only those children who played games at difficulty level A to make experienced games homogenous because games at levels B and C have more objects to drag and drop than do games at level A (n=623). Finally, we included children who played at least one subgame among the second, third, and sixth subgames because the other subgames in chapter 1 are played with tapping answer objects (n=370). In these drag-and-drop subgames, we analyzed the drag-and-drop log data to classify those children with typical development and those with disabilities (Figure 1).

Figure 1. Eligible user selection flow.



Ethics Statement

The retrospective study was approved by the institutional review board of Yonsei University College of Medicine in South Korea (no. Y-2020-0076). The Division of Biomedical System Informatics (Department of Medicine of Yonsei University) and DoBrain work together as a nonprofit, joint research group for the early detection of disabilities in children and improved cognitive function. To conduct this research, we obtained deidentified data from DoBrain, and we have no conflicts of interest related to our dealings with DoBrain Inc.

Analytical Procedure

In our study, we compared the baseline characteristics of children with and without developmental disabilities using the *t* test or Mann-Whitney test for continuous variables (eg, age and device size) and used chi-square test for categorical variables (eg, sex). Furthermore, we derived features related to drag-and-drop data to capture children's kinetics (Multimedia Appendix 2). We additionally analyzed these derived data to identify differences in features from finger strokes between the 2 groups. Before comparative tests, we explored the normality of distribution by visual methods and statistical tests (Kolmogorov-Smirnov test) [22]. We determined normality with consideration to the shape of the histogram and the results of statistical tests [23]. In addition, we conducted an *F* test for homoscedasticity (equal variance in 2 populations). We then conducted a *t* test for normally distributed data and a Mann-Whitney test for nonnormally distributed data.

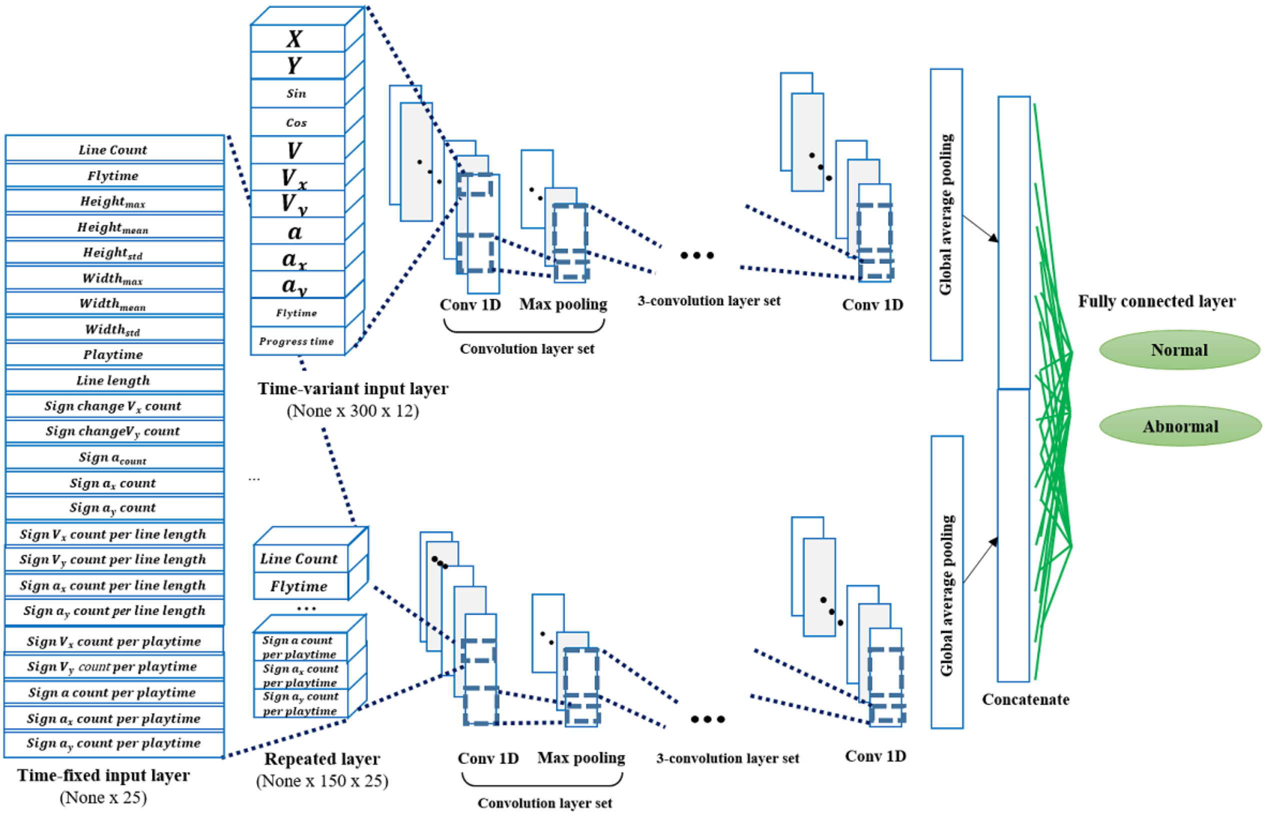
P values <.05 were considered statistically significant for 2-sided hypothetical tests. In addition, we calculated effect sizes to determine the possibility of type I statistical error. Cohen *d* for continuous variables for normally distributed data and for categorical variables were considered small depending on the type of effect size ($\eta^2 \approx 0.01$; $-0.20 < \text{Cohen } d < 0.20$) [24]. For nonparametric comparative methods, such as the Mann-Whitney test, the common language effect size (CLES) was calculated to identify the probability that a score sampled at random from one distribution would be greater than a score sampled from another distribution. CLES reflects the chance that a value for a randomly selected child with typical development would be higher than that from one with developmental disabilities [25].

For detecting children with developmental disabilities, we developed a deep learning classification model based on a 1D convolutional neural network for drag data. Using drag data, we tried to leverage multiple inputs (time variant variables: touch coordinates and their derived variables; time-fixed variables: statistics acquired at the end of the game, such as total touch area or demographic data) by joint fusion. We subsequently modeled the classification algorithm using deep learning and not conventional machine learning. In addition, we applied a strategy to decompose coordinates (fine motor movement) along the x- and y-axes. Through this decomposition of coordinates, we were able to leverage information along each axis by creating derived variables, such as velocity and acceleration, along the axes and volatility of sign change. Because this approach did not use positional information (contextual information) in 2D, we developed a model with a 1D convolutional neural network.

Drag-and-drop data, including all touch traces, were leveraged in our model. The traces of individual touch attempts were captured and stored in the forms of logs when the user touched an object on the display. We tried to use drag-and-drop data regardless of intention to touch. Because the direction is guided through sounds at the beginning of the game, we built the model to capture unnecessary touches. Therefore, we used all touch records to classify children with or without developmental disabilities.

At the end of the model, multiple inputs were concatenated by the individual user ID, and the output node calculated the binary prediction of developmental disabilities using a fully connected layer. For time-variant variables, derived variables calculated from each drag-and-drop feature were input. In contrast, for the time-fixed variable, demographic characteristics and game results that were generated after the end of the subgame were input (Figure 2).

Figure 2. Deep learning model architecture. conv: convolution.



For hyperparameter optimization, we searched hyperparameters with a grid search. A detailed description of this search is included in Multimedia Appendix 3. Deep learning models based on each subgame were also evaluated using 10-fold cross-validation. The area under the receiver operating characteristics curve (AUROC), area under the precision-recall curve, *F* score, precision, recall (sensitivity), and specificity for each subgame were calculated as the aggregation of the 10-fold cross-validation results.

Finally, we not only focused on building a classification model for developmental delay but also developed an interface with which to examine the network's decision to assess the fine motor movement feature detected by the network. We applied the gradient-weighted class activation mapping (Grad-CAM) method to determine which coordinates of fine motor movement touch were useful for predicting developmental disabilities in children [26]. For this, we overlaid coordinates with an attention map from the Grad-CAM, which showed the coordinates of positive correlates with the output of the network. All statistical analyses and model development were conducted using Python 3.6.8 and TensorFlow 1.14.0.

Ethics Statements

Our relationship with DoBrain Inc. involves no conflicts of interest. We and DoBrain Inc. have experience in conducting national research projects together ("Cognitive learning service for children with developmental disabilities using on AI" [unpublished data], 2019-2020) with funding from the National Information Society Agency, South Korea. Within these national projects, we obtained the deidentified data from DoBrain Inc.

Results

Baseline Characteristics

A total of 368 eligible children were included in the study. Of these, 366, 361, and 337 played the second, third, and sixth subgames, respectively (overall, 223 children had typical development, and 147 had developmental disabilities). There was a statistical difference in the chronological age of the 2 groups (P<.001; CLES=0.839) and in the mean playtime (P<.001; CLES=0.687). In each played subgame, a difference in the ratio of children who played games was only observed for the sixth subgame (P=.04; CLES=0.002; Table 1).

Table 1. Demographic characteristics in eligible users.

Variables	Children with typical development (n=223)	Children with developmental disabilities (n=147)	Total (n=370)	*P* value	Effect size
Age (months), median (IQR)	40.0 (12.0)	72.0 (32.5)	45.0 (26.5)	<.001	0.839[a]
Diagnosis, n (%)					
Intellectual disability	0 (0.0)	44 (0.33)	44 (0.11)	N/A[b]	N/A
Autism spectrum disorder	0 (0.0)	41 (0.27)	41 (0.11)	N/A	N/A
Developmental disorder	0 (0.0)	33 (0.22)	33 (0.09)	N/A	N/A
Brain lesions	0 (0.0)	25 (0.17)	25 (0.09)	N/A	N/A
Monogenic disorder	0 (0.0)	4 (0.02)	4 (0.01)	N/A	N/A
Children playing subgame, n (%)					
Second subgame	222 (99.55)	144 (97.96)	366 (98.92)	.03	<0.001
Third subgame	216 (97.31)	145 (98.64)	361 (97.84)	.24	<0.001
Sixth subgame	198 (88.79)	139 (95.56)	337 (91.98)	.04	0.002
Device size (inches), median (IQR)	6.1 (4.2)	6.1 (4.2)	6.1 (4.2)	<.001	0.469[a]
Game playtime, (s/per game), median (IQR)	11.8 (8.5)	8.1 (6.5)	10.3 (8.2)	<.001	0.687[a]
Games played, median (IQR)	9.0 (9.0)	1.0 (14.2)	7.0 (12.5)	<.001	0.682[a]

[a]Common language effect size of continuous variables; η^2 for effect size of categorical variables.

[b]N/A: not applicable.

Characteristics of Drag-and-Drop Game Play

Although playtime did not consistently show statistically significant differences in each subgame, the playtime in children with typical development was significantly longer than that of children with disabilities for the sixth subgame (*P*<.001; CLES=0.616; Table 2). In the touch region, the variables related to measurements of the touch area did not show a statistically significant difference in median values. However, in the sixth subgame, the median max of height and width that children used for playing games showed a significant difference (*P*<.001; CLES>0.65). In addition, although the median change of velocity sign change along the x-axis did not show a difference (*P*=.40, *P*=0.17, and *P*=0.08, respectively) in all subgames, the IQR of children with typical development was smaller than the that of children with developmental disabilities (Multimedia Appendix 4). In 4 accelerator variables, including sign change of acceleration along the x-axis and y-axis, with both count and per line count, the number of sign changes for children with typical development was larger than that for children with developmental disabilities.

Table 2. Comparison of movement features between children with typical development and children with developmental disabilities in each game.

Characteristic	Children with typical development	Children with developmental disabilities	*P* value	CLES[a]
Second game				
Distribution, n	222	144	N/A[b]	N/A
Play information, median (IQR)				
Playtime (second/game)	2.03 (6.39)	1.98 (5.11)	.33	0.513
Line number (n/game)	1.0 (2.0)	2.0 (2.0)	.24	0.376
Line length (n/game)	3.04 (3.17)	3.1 (4.24)	.28	0.518
Release to touch time (sec)	0.0 (3.31)	0.52 (4.46)	.15	0.409
Touch region, median (IQR)				
Height mean, ratio (%)	0.38 (0.1)	0.38 (0.08)	.26	0.540
Height max, ratio (%)	0.58 (0.08)	0.57 (0.06)	.10	0.520
Width max, ratio (%))	0.74 (0.08)	0.72 (0.08)	.13	0.534
Height mean, ratio (%)	0.38 (0.1)	0.38 (0.08)	.006	0.540
Third game				
Distribution, n	216	145	N/A	N/A
Play information, median (IQR)				
Playtime (second/game)	2.75 (6.63)	2.44 (5.99)	.22	0.524
Line number (n/game)	1.0 (2.0)	1.0 (2.0)	.46	0.344
Line length (n/game)	3.12 (4.23)	3.62 (6.28)	.04	0.553
Release to touch time (sec)	0.0 (2.84)	0.0 (3.11)	.45	0.372
Touch region, median (IQR)				
Height mean, ratio (%)	0.29 (0.09)	0.27 (0.06)	.02	0.520
Height max, ratio (%)	0.35 (0.23)	0.35 (0.17)	.26	0.565
Width max, ratio (%))	0.72 (0.1)	0.71 (0.08)	.18	0.528
Height mean, ratio (%)	0.29 (0.09)	0.27 (0.06)	.04	0.520
Sixth game				
Distribution, n	198	139	N/A	N/A
Play information, median (IQR)				
Playtime (second/game)	7.85 (15.25)	4.33 (10.03)	<.001	0.616
Line number (n/game)	3.0 (3.0)	2.0 (3.0)	.001	0.514
Line length (n/game)	4.64 (8.5)	4.5 (7.66)	.41	0.507
Release to touch time (sec)	2.9 (8.33)	1.4 (6.27)	.001	0.549
Touch region, median (IQR)				
Height mean, ratio (%)	0.43 (0.06)	0.34 (0.16)	<.001	0.695
Height max, ratio (%)	0.53 (0.09)	0.49 (0.2)	<.001	0.766
Width max, ratio (%))	0.78 (0.1)	0.73 (0.09)	<.001	0.658
Height mean, ratio (%)	0.43 (0.06)	0.34 (0.16)	<.001	0.695

[a]CLES: common language effect size.

[b]N/A: not applicable.

Model Performance

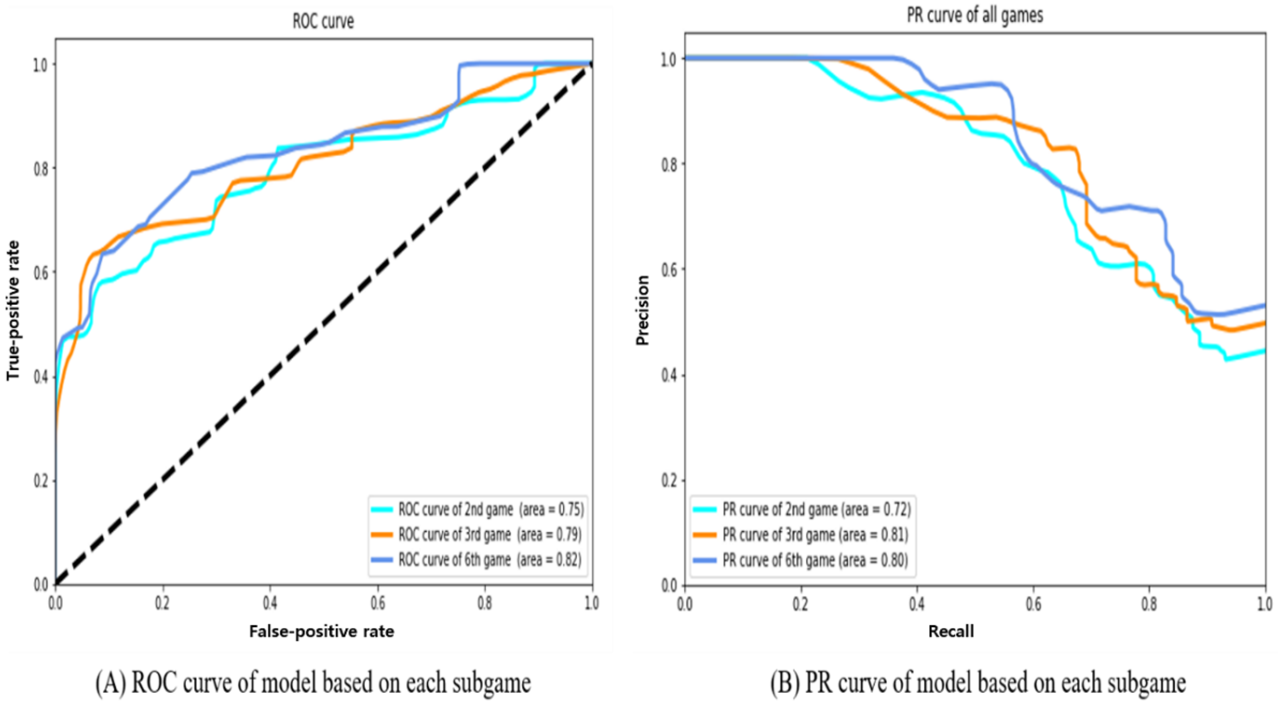
Overall, the average AUROCs for the second, third, and sixth games were calculated as 0.746 (σ=0.116), 0.793 (σ=0.117), and 0.817 (σ=0.070), respectively, in a 10-fold cross-validation; meanwhile, average *F* scores of the deep learning models for each targeted subgame were calculated as 0.627, 0.675, and 0.708, respectively (Table 3). The model for the sixth subgame showed relatively high performance compared to other models using the second or third subgames in terms of average model performance metrics (AUROC, accuracy, *F* score, precision, recall). Recall (also called sensitivity), which refers to the probability of a positive test given that the patient has a disease, was highest in the model for the sixth game (0.757; σ=0.123) (Figure 3). Specificity, which refers to the probability of a negative test given that the patient is normal, was relatively better in the model for the third subgame (0.783; σ=0.069) than in the other models using the second or sixth subgame (second subgame: 0.755, σ=0.089; sixth subgame: 0.740, σ=0.099).

Table 3. Average model performance result by 10-fold cross-validation in each drag-and-drop subgame.

Deep learning model for individual subgame	Performance, mean (SD)					
	AUROC[a]	Accuracy	*F* score	Precision	Recall	Specificity
Second subgame	0.746 (0.116)	0.719 (0.094)	0.627 (0.165)	0.616 (0.136)	0.683 (0.151)	0.755 (0.089)
Third subgame	0.793 (0.117)	0.747 (0.082)	0.675 (0.126)	0.686 (0.142)	0.688 (0.167)	0.783 (0.069)
Sixth subgame	0.817 (0.070)	0.769 (0.078)	0.708 (0.153)	0.675 (0.183)	0.757 (0.123)	0.740 (0.099)

[a]AUROC: area under the receiver operating characteristics curve.

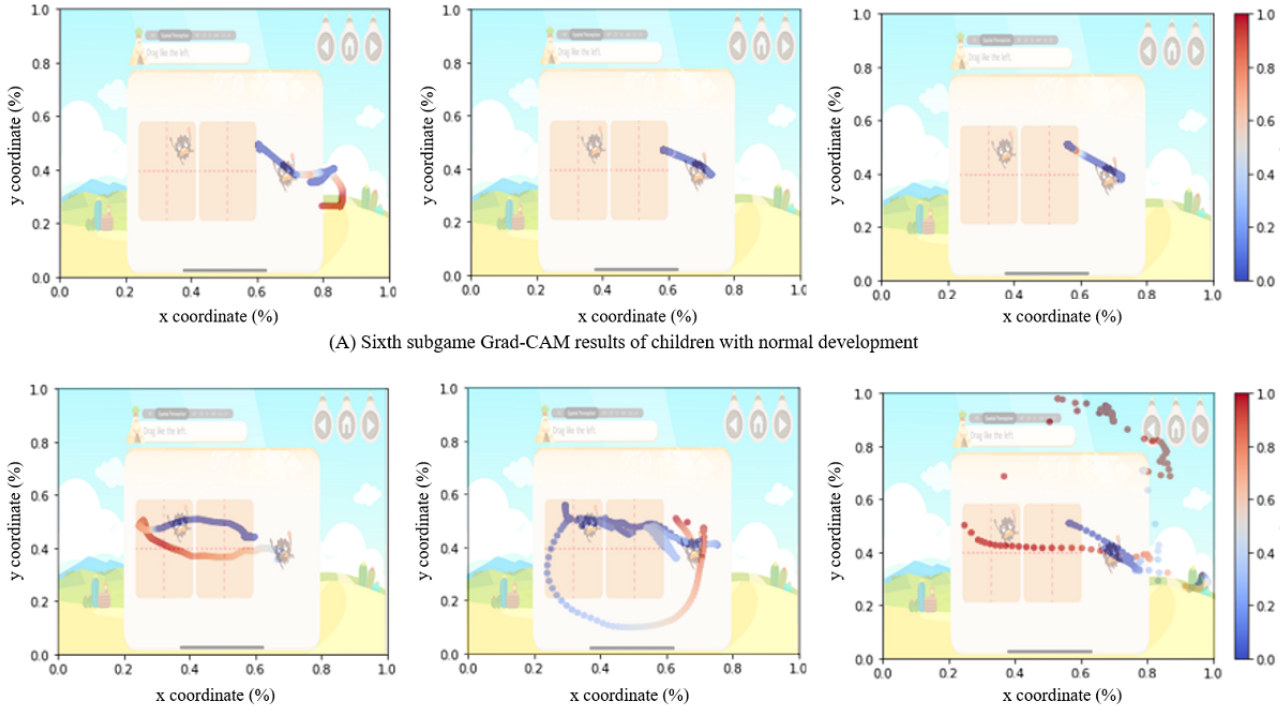
Figure 3. ROC curves and PR curves of the deep learning model. PR: precision-recall. ROC: receiver operating characteristic.



Prediction of Developmental Disabilities From Finger Strokes

The visualization of variable weights for developmental disabilities on game result images relied on a gradient-based projection of the classification scores to the input pixels. Children with typical development correctly implemented the optimal path proximity to answers, as recorded by the game; in Figure 4A, coordinates similar to the optimal path are visualized in blue, as presented in Grad-CAM. In contrast, we could confirm that children with developmental disabilities played games and drew gestures in various locations before drawing the optimal path (Figure 4B). Unlike the coordinates of children with typical development, coordinates located in the nonoptimal path are shown in red, as presented by Grad-CAM.

Figure 4. Sixth subgame with Grad-CAM results of children with normal and abnormal development disability. Grad-CAM: gradient-weighted class activation mapping.



(A) Sixth subgame Grad-CAM results of children with normal development

(B) Sixth subgame Grad-CAM results of children with developmental disabiltiy

Discussion

Principal Findings

This study showed that fine motor movements captured from touching a mobile display can be a novel digital biomarker for the classification of developmental delay. Our classification model leveraged moment-by-moment drag-and-drop data to capture fine motor movements for serious game play. This suggests that serious games can be used as diagnostic assistant tools or screen tools for detecting developmental disabilities. Moreover, our model can be adaptive to clinicians because it can visualize how much each coordinate of fine motor movement contributes to classification.

Drag-and-Drop Data as a Digital Biomarker for Detecting Developmental Delay

Early motor delay is often a sign of neurological dysfunction [27]. Fine motor skills are related to the use of upper extremities to engage and manipulate the environment [15,27]. This function is necessary for children to play or accomplish work. At the age of 3 years, children can copy circles and imitate a cross in the course of reaching developmental milestones [28]. Similarly, the games in our study require the user to select an object of the same color or solve the game by imitating a given environment (such as a picture on the left in Multimedia Appendix 1). Although the screening delivery platform has been changed into a digital device, it still seems capable of reflecting developmental milestones.

In the clinic or community, parent- or teacher-reported screening tools are used for large populations, including the Behavior Rating Inventory of Executive Function II (BRIEF-II) and the Denver Developmental Screening Tests II (DDST-II),. These screening tools collect information on the kinesthetic ability of children from parents or teachers. However, as these tools leverage information obtained from parents or teachers to calculate scores reflecting kinesthetic abilities, the results of the screening tools may have inter- or intraobserver variability. [29] Despite this, the use of drag-and-drop data in our model has a strength in that our model presents more objective results.

Previous studies have reported that children with autism show a rate of change in acceleration in movements that is significantly greater than that in children with typical development [30,31]. Although previous studies have reported on acceleration of the device (inertia) captured from a gyroscope in the device or infrared motion tracking system, our study showed higher rates of sign changes in acceleration (in all subgames) and velocity (in the sixth subgame) in finger strokes for dragging and dropping objects in children with typical development compared to children with developmental disabilities. Given that children with developmental disabilities had shorter playtimes and less variability, children with developmental disabilities might not have intuitively understood the problem and indiscriminately ran their finger across the screen. It is also possible that the difficulty of the game may be insufficient to screen problems even in children with developmental disabilities because the purpose of the game is to drag objects in a single straight line.

Early Detection and Diagnostic Stability

There are multiple challenges to screening children with developmental disabilities in routine clinical practice [2]. In step with the high clinical demand, there are long wait times for children who require examination by specialists. If screening tools are negative for developmental delay and the parents continue to be concerned about their child's behavior, a more intensive follow-up plan is needed, such as shorter-interval, repetitive screening tests. Other barriers include a lack of consensus on the best screening tools and insufficient physician confidence [2,32,33]. A previous study on diagnostic stabilities in children with and without autism spectrum disorder in the United States showed that 21% of children were initially not diagnosed with autism spectrum disorder [12]. From this point of view, developmental disability classification using mobile devices can be used as an element to potentially overcome these challenges.

Further, for early identification of developmental disability, cross-culturally appropriate and affordable tools are important, although tools satisfying these conditions are limited [34]. Applying tools developed in Western-based norms to other cultural contexts can induce overdetection in children, as there is a disparity in global pediatric mental health, especially in low- and middle-income countries [35]. Considering that many children do not regularly visit medical or mental health professionals in low- and middle-income countries [36], a screening tool that is quick and inexpensive would be desirable. In addition, tools with child-friendly characteristics for children to complete by themselves and under repetitive use with lower resources would be suitable. In this respect, a serious game with drag-and-drop data can be a candidate for tools that satisfy these requirements.

Limitations

This study has several limitations. First, the diagnostic profile of children with typical development used in our study was based on patient reports. Because children with typical development were not confirmed by physicians, this could have led to bias in the performance or results of the model due to the reliability of the label. However, previous studies included clinical trials using only limited samples, whereas our study analyzed hundreds of children with developmental disabilities diagnosed by a pediatrician. Therefore, compared to previous studies, our study represents an improvement in the robustness of the prediction results. Second, this study excluded children with developmental disabilities who were unable to control the mobile device by themselves. Because data acquisition assumed children could handle the mobile device and understand the instructions to the games, the characteristics of children with severe or moderate developmental disabilities were not considered in this study. Further research needs to be conducted after analyzing children with consideration of the degree of developmental disorders.

Conclusions

As continuous and comprehensive tracking for more accurate assessment is important in screening developmental disabilities, a screening tool that can be easily, repetitively, and objectively used is needed. To the best of our knowledge, this retrospective study is the first to show that a deep learning-based screening model leveraging digital biomarkers could be feasible for detecting developmental disabilities in children. Therefore, finger strokes on a mobile touch display can be a novel digital biomarker of use in screening for developmental disabilities.

Authors' Contributions

HHK, JIA, and YRY had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

HHK and JIA conceived and designed the study and performed statistical analyses. All authors contributed to the acquisition, analysis, or interpretation of data. HHK drafted the manuscript, and all authors critically revised the manuscript for important intellectual content. YRP provided administrative, technical, and material support, and supervision.

This study was supported by a grant from the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute and the Ministry of Health & Welfare, Republic of Korea (grant no. HI19C1015).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Captured images of DoBrain chapter 1 subgames.
[DOCX File , 755 KB - games_v9i2e23130_app1.docx]

Multimedia Appendix 2
Features used as deep learning inputs and their descriptions.
[DOCX File , 111 KB - games_v9i2e23130_app2.docx]

Multimedia Appendix 3
Velocity and acceleration of movement.
[DOCX File , 21 KB - games_v9i2e23130_app3.docx]

Multimedia Appendix 4
Hyperparameter optimization.
[DOCX File , 13 KB - games_v9i2e23130_app4.docx]

References

1. Zablotsky B, Black L, Blumberg S. Estimated prevalence of children with diagnosed developmental disabilities in the United States, 2014–2016. NCHS Data Brief. 2017. URL: https://www.cdc.gov/nchs/products/databriefs/db291.htm?fbclid=IwAR0MwhWDgawPU [accessed 2021-05-17]
2. Vitrikas K, Savard D, Bucaj M. Developmental delay: when and how to screen. Am Fam Physician 2017 Jul 01;96(1):36-43 [FREE Full text] [Medline: 28671370]
3. Choo Y, Agarwal P, How C, Yeleswarapu S. Developmental delay: identification and management at primary care level. Singapore Med J 2019 Mar;60(3):119-123 [FREE Full text] [doi: 10.11622/smedj.2019025] [Medline: 30997518]
4. Zablotsky B, Black LI, Maenner MJ, Schieve LA, Danielson ML, Bitsko RH, et al. Prevalence and trends of developmental disabilities among children in the United States: 2009-2017. Pediatrics 2019 Oct;144(4):e20190811 [FREE Full text] [doi: 10.1542/peds.2019-0811] [Medline: 31558576]
5. Kohli-Lynch M, Tann CJ, Ellis ME. Early intervention for children at high risk of developmental disability in low- and middle-income countries: a narrative review. Int J Environ Res Public Health 2019 Nov 13;16(22):44-49 [FREE Full text] [doi: 10.3390/ijerph16224449] [Medline: 31766126]
6. Bradshaw J, Steiner AM, Gengoux G, Koegel LK. Feasibility and effectiveness of very early intervention for infants at-risk for autism spectrum disorder: a systematic review. J Autism Dev Disord 2015 Mar;45(3):778-794. [doi: 10.1007/s10803-014-2235-2] [Medline: 25218848]
7. Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-Vanderweele J. A systematic review of early intensive intervention for autism spectrum disorders. Pediatrics 2011 May;127(5):e1303-e1311. [doi: 10.1542/peds.2011-0426] [Medline: 21464190]
8. Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH, Martin CL. A cross-disorder method to identify novel candidate genes for developmental brain disorders. JAMA Psychiatry 2016 Mar;73(3):275-283 [FREE Full text] [doi: 10.1001/jamapsychiatry.2015.2692] [Medline: 26817790]
9. Magnusson D, Palta M, McManus B, Benedict RE, Durkin MS. Capturing unmet therapy need among young children with developmental delay using national survey data. Acad Pediatr 2016 Mar;16(2):145-153. [doi: 10.1016/j.acap.2015.05.003] [Medline: 26183004]
10. Roberts MY, Kaiser AP. Early intervention for toddlers with language delays: a randomized controlled trial. Pediatrics 2015 Apr;135(4):686-693 [FREE Full text] [doi: 10.1542/peds.2014-2134] [Medline: 25733749]
11. Mackrides PS, Ryherd SJ. Screening for developmental delay. Am Fam Physician 2011 Sep 01;84(5):544-549 [FREE Full text] [Medline: 21888305]
12. Pierce K, Gazestani VH, Bacon E, Barnes CC, Cha D, Nalabolu S, et al. Evaluation of the diagnostic stability of the early autism spectrum disorder phenotype in the general population starting at 12 months. JAMA Pediatr 2019 Jun 01;173(6):578-587 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.0624] [Medline: 31034004]
13. Mathews TL, King ML, Kupzyk KA, Lake CM. Findings and implications of developmental screening for high-risk children referred to a tertiary developmental disability center. J Pediatr Health Care 2014;28(6):507-515. [doi: 10.1016/j.pedhc.2014.03.002] [Medline: 24793986]
14. Flynn RM, Colón-Acosta N, Zhou J, Bower J. A game-based repeated assessment for cognitive monitoring: initial usability and adherence study in a summer camp setting. J Autism Dev Disord 2019 May;49(5):2003-2014. [doi: 10.1007/s10803-019-03881-w] [Medline: 30656527]
15. Bellman M, Byrne O, Sege R. Developmental assessment of children. BMJ 2013 Jan 15;346:e8687. [doi: 10.1136/bmj.e8687] [Medline: 23321410]
16. Aslam A. The State of the World's Children 2013: Children With Disabilities. New York: United Nations Children's Fund (UNICEF); 2013:9280646567.
17. Rydz D, Srour M, Oskoui M, Marget N, Shiller M, Birnbaum R, et al. Screening for developmental delay in the setting of a community pediatric clinic: a prospective assessment of parent-report questionnaires. Pediatrics 2006 Oct;118(4):e1178-e1186. [doi: 10.1542/peds.2006-0466] [Medline: 17015506]
18. de Souza N, Sardessai V, Joshi K, Joshi V, Hughes M. The determinants of compliance with an early intervention programme for high-risk babies in India. Child Care Health Dev 2006 Jan;32(1):63-72. [doi: 10.1111/j.1365-2214.2006.00576.x] [Medline: 16398792]
19. Lobo MA, Paul DA, Mackley A, Maher J, Galloway JC. Instability of delay classification and determination of early intervention eligibility in the first two years of life. Res Dev Disabil 2014 Jan;35(1):117-126 [FREE Full text] [doi: 10.1016/j.ridd.2013.10.017] [Medline: 24176257]
20. Matson JL, Mahan S, Fodstad JC, Hess JA, Neal D. Motor skill abilities in toddlers with autistic disorder, pervasive developmental disorder-not otherwise specified, and atypical development. Research in Autism Spectrum Disorders 2010 Jul;4(3):444-449. [doi: 10.1016/j.rasd.2009.10.018]

21. Whyatt CP, Craig CM. Motor skills in children aged 7-10 years, diagnosed with autism spectrum disorder. J Autism Dev Disord 2012 Sep;42(9):1799-1809. [doi: 10.1007/s10803-011-1421-8] [Medline: 22180003]
22. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. Int J Endocrinol Metab 2012;10(2):486-489 [FREE Full text] [doi: 10.5812/ijem.3505] [Medline: 23843808]
23. Elliott A, Woodward W. Statistical Analysis Quick Reference Guidebook: With SPSS Examples. London: SAGE Publications, Inc; 2007.
24. Osteen P, Bright C. Effect sizes and intervention research. University of Maryland Baltimore. URL: https://archive.hshsl.umaryland.edu/bitstream/handle/10713/3582/effect?sequence=1 [accessed 2021-05-18]
25. McGraw KO, Wong SP. A common language effect size statistic. Psychological Bulletin 1992;111(2):361-365. [doi: 10.1037/0033-2909.111.2.361]
26. Selvaraju R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. 2017 Presented at: Proceedings of the IEEE international conference on computer vision; Oct 22-29; Venice. [doi: 10.1109/iccv.2017.74]
27. Tervo RC. Identifying patterns of developmental delays can help diagnose neurodevelopmental disorders. Clin Pediatr (Phila) 2006 Jul;45(6):509-517. [doi: 10.1177/0009922806290566] [Medline: 16893855]
28. Gerber RJ, Wilks T, Erdie-Lalena C. Developmental milestones: motor development. Pediatr Rev 2010 Jul;31(7):267-76; quiz 277. [doi: 10.1542/pir.31-7-267] [Medline: 20595440]
29. Weed LL. Medical records that guide and teach. N Engl J Med 1968 Mar 21;278(12):652-7 concl. [doi: 10.1056/NEJM196803212781204] [Medline: 5637250]
30. Anzulewicz A, Sobota K, Delafield-Butt JT. Toward the Autism Motor Signature: Gesture patterns during smart tablet gameplay identify children with autism. Sci Rep 2016 Aug 24;6:31107 [FREE Full text] [doi: 10.1038/srep31107] [Medline: 27553971]
31. Cook JL, Blakemore S, Press C. Atypical basic movement kinematics in autism spectrum conditions. Brain 2013 Sep;136(Pt 9):2816-2824 [FREE Full text] [doi: 10.1093/brain/awt208] [Medline: 23983031]
32. Drotar D, Stancin T, Dworkin PH, Sices L, Wood S. Selecting developmental surveillance and screening tools. Pediatr Rev 2008 Oct;29(10):e52-e58. [doi: 10.1542/pir.29-10-e52] [Medline: 18829768]
33. Warren R, Kenny M, Bennett T, Fitzpatrick-Lewis D, Ali MU, Sherifali D, et al. Screening for developmental delay among children aged 1-4 years: a systematic review. CMAJ Open 2016 Jan 25;4(1):E20-E27 [FREE Full text] [doi: 10.9778/cmajo.20140121] [Medline: 27226967]
34. Goldfeld S, Yousafzai A. Monitoring tools for child development: an opportunity for action. Lancet Glob Health 2018 Mar;6(3):e232-e233 [FREE Full text] [doi: 10.1016/S2214-109X(18)30040-8] [Medline: 29433654]
35. Marlow M, Servili C, Tomlinson M. A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries. Autism Res 2019 Feb;12(2):176-199. [doi: 10.1002/aur.2033] [Medline: 30707000]
36. Biasini FJ, De Jong D, Ryan S, Thorsten V, Bann C, Bellad R, et al. Development of a 12 month screener based on items from the Bayley II Scales of Infant Development for use in Low Middle Income countries. Early Hum Dev 2015 Apr;91(4):253-258 [FREE Full text] [doi: 10.1016/j.earlhumdev.2015.02.001] [Medline: 25734979]

Abbreviations

AUROC: area under the receiver operating characteristics curve
BRIEF-II: Behavior Rating Inventory of Executive Function II
CLES: common language effect size
DDST-II: Denver Developmental Screening Tests II
Grad-CAM: gradient-weighted class activation mapping

Edited by R Kukafka; submitted 02.08.20; peer-reviewed by P Caserman, T Ueno; comments to author 30.08.20; revised version received 15.10.20; accepted 09.04.21; published 04.06.21.

Please cite as:
Kim HH, An JI, Park YR
A Prediction Model for Detecting Developmental Disabilities in Preschool-Age Children Through Digital Biomarker-Driven Deep Learning in Serious Games: Development Study
JMIR Serious Games 2021;9(2):e23130
URL: https://games.jmir.org/2021/2/e23130
doi:10.2196/23130
PMID:34085944

Original Paper

Brain Exercising Games With Consumer-Grade Single-Channel Electroencephalogram Neurofeedback: Pre-Post Intervention Study

Pasin Israsena[1*], PhD; Suwicha Jirayucharoensak[1*], PhD; Solaphat Hemrungrojn[2,3*], MD; Setha Pan-Ngum[4*], PhD

[1]National Electronics and Computer Technology Center, National Science and Technology Development Agency, Pathumthani, Thailand

[2]Department of Psychiatry, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

[3]Cognitive Fitness Research Center, Chulalongkorn University, Bangkok, Thailand

[4]Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand

[*]all authors contributed equally

Corresponding Author:
Pasin Israsena, PhD
National Electronics and Computer Technology Center
National Science and Technology Development Agency
112 Thailand Science Park
Klong Luang
Pathumthani, 12120
Thailand
Phone: 66 25646900 ext 2471
Email: pasin.israsena@nectec.or.th

Abstract

Background: The aging population is one of the major challenges affecting societies worldwide. As the proportion of older people grows dramatically, so does the number of age-related illnesses such as dementia-related illnesses. Preventive care should be emphasized as an effective tool to combat and manage this situation.

Objective: The aim of this pilot project was to study the benefits of using neurofeedback-based brain training games for enhancing cognitive performance in the elderly population. In particular, aiming for practicality, the training games were designed to operate with a low-cost consumer-grade single-channel electroencephalogram (EEG) headset that should make the service scalable and more accessible for wider adoption such as for home use.

Methods: Our training system, which consisted of five brain exercise games using neurofeedback, was serviced at 5 hospitals in Thailand. Participants were screened for cognitive levels using the Thai Mental State Examination and Montreal Cognitive Assessment. Those who passed the criteria were further assessed with the Cambridge Neuropsychological Test Automated Battery (CANTAB) computerized cognitive assessment battery. The physiological state of the brain was also assessed using 16-channel EEG. After 20 sessions of training, cognitive performance and EEG were assessed again to compare pretraining and posttraining results.

Results: Thirty-five participants completed the training. CANTAB results showed positive and significant effects in the visual memory (delayed matching to sample [percent correct] *P*=.04), attention (median latency *P*=.009), and visual recognition (spatial working memory [between errors] *P*=.03) domains. EEG also showed improvement in upper alpha activity in a resting state (open-eyed) measured from the occipital area (*P*=.04), which similarly indicated improvement in the cognitive domain (attention).

Conclusions: Outcomes of this study show the potential use of practical neurofeedback-based training games for brain exercise to enhance cognitive performance in the elderly population.

(JMIR Serious Games 2021;9(2):e26872) doi:10.2196/26872

KEYWORDS

neurofeedback; serious gaming; serious game; brain exercise; cognition training; EEG; aging; cognition; cognitive; brain game

Introduction

Background

The global population aged above 60 years is expected to increase from 900 million in 2015 to 1.4 billion by 2030 and to 2.1 billion by 2050 [1]. Approximately two-thirds of these elderly people live in low-income and middle-income countries [2]. This growing number of elderly will lead to an increase in the incidence of aging-related disorders, including cognitive-related diseases such as dementia, Alzheimer disease, and mild cognitive impairment (MCI), which are increasingly becoming more common [3]. Dementia, the most frequent brain disorder, is differentiated according to the etiology of either vascular dementia or neurodegenerative dementia (such as Alzheimer disease) [4,5]. Dementia is characterized by a devastating reduction in cognitive abilities, including control of behavior, learning, memory, attention/sleep, language, intelligence, perception, as well as functional independence and social relationships. More than 47 million people are living with dementia [6], approximately 60% of whom live in low- and middle-income countries. Driven by population aging, this number is expected to triple by 2050, as the incidence of dementia rises sharply at ages above 75 years. The economic impact has been estimated at US $600 billion per year worldwide [7], including a significant burden of unpaid family caregiving.

In contrast to Alzheimer disease and dementia, which are conditions that are diagnosed at a state that is difficult to treat, MCI—representing an intermediate state between cognitive decline associated with normal aging and dementia [8,9]—is more suitable for targeted preventive measures along with active prevention in the healthy elderly population, as an elderly individual with MCI still has well-preserved functional abilities [10,11]. The global prevalence of MCI in the population aged above 60 years reaches up to 38.60% [12]. MCI is a risk factor for dementia [13] and is associated with a 6-fold increased risk of Alzheimer disease [14]. Indeed, as pharmaceutical-based treatments have yet to be very successful in a late-stage condition such as Alzheimer disease, and may come with certain side effects [15-18], researchers have been exploring the usefulness of alternative strategies to prevent, prolong decline, or even enhance cognitive performance with different methods such as cognitive training [19-22] or neurofeedback training (NFT) [23-25]. Because of the aging trend in developing countries mentioned above, in addition to their demonstrated effectiveness, these methods should be low cost so that they can be beneficial to the majority of the target group.

NFT as a Solution for Cognitive Decline

NFT is a noninvasive technique to alter brain activity [23-25]. NFT does not actively interfere with the brain but is rather a process during which subjects learn to influence their brain wave pattern by receiving feedback of their brain activity (visual, audio, or other modalities). The controllability of brain rhythms from users often also elicits neural plasticity in the brain that can affect their behaviors and cognitive functions. The use of NFT in medical and therapeutic contexts has continuously attracted interest in research and practice [23]. In particular, in the field that addresses mental problems such as attention deficit hyperactivity disorder (ADHD) or depression [26-31], NFT has been considered a potentially viable tool.

There are several types of NFT based on different brain imaging technologies such hemoencephalographic, low-resolution electromagnetic tomography, functional magnetic resonance imaging, or functional near-infrared spectroscopy. However, the most common type of NFT usually provides brain state information to the user via electroencephalogram (EEG) signals (also called EEG biofeedback). Recorded at the scalp, the EEG is produced by synchronous postsynaptic potentials from thousands to millions of neurons. When amplified, digitized, and plotted, the raw EEG signal appears as a complex oscillatory pattern. Raw EEG signals are observed via frequency bands of interest (ie, delta, 1-4 Hz; theta, 4-8 Hz; alpha, 8-12 Hz; beta, 12-32 Hz; and gamma, 32-60 Hz), usually in the forms of their amplitudes or power spectra.

Brain wave characteristics have been shown to change dramatically depending on the task at hand. Among these, alpha is a particularly interesting oscillation, which is the predominant rhythm in the human brain in a resting state, especially when the eyes are closed [32]. Originally, alpha NFT was considered as simple relaxation training. However, renewed interest has arisen to clarify the relation between the alpha frequency band and cognitive performance [33-35] via indicators such as alpha peak frequency [36,37] and individual upper alpha amplitude/power [38-40]. In resting-state recordings, the alpha peak is clearly visible between 7.5 and 12.5 Hz. Higher alpha peak frequencies (eg, 12 Hz in comparison to 10 Hz) have been shown to correlate positively with high memory performance [36,37] and IQ [41]. Moreover, in terms of alpha amplitude or power, a positive correlation between individual upper alpha amplitude and IQ has been reported (eg, [42,43]). Additionally, as suggested by Klimesch [38], high alpha power during a resting state and low alpha power during the execution of a task were associated with good performance in semantic long-term memory tasks. Despite these promising results, NFT has not yet been translated into mainstream practice. This could be due to limitations of study designs such as lack of a methodological design, therapeutic effect, and validity, as reviewed previously [44,45].

Consumer-Grade EEG

Until recently, most EEG equipment has been designed for either research or medical purposes. These devices were rather expensive and not easy to operate. An EEG headset usually has multichannel EEG wet sensors held together with a head cap in predefined positions according to the international 10-20 system [46]. Therefore, considerable time is needed for each setup before use and for subsequent cleanup, as these wet sensors needed to be lubricated before each use. The emergence of low-cost consumer-grade EEG headsets offers an easy-to-administer and practical solution via the use of dry or semidry electrodes, along with wireless connectivity, thereby opening the technology for mass use. More importantly, most of the consumer-grade EEGs also have an option for continuous raw EEG reading, as opposed to the batch records usually offered in medical-grade EEGs. This online EEG reading feature is a necessity for systems such as NFT.

There are several consumer-grade EEGs currently available on the market, which differ in terms of the numbers and types of electrodes, positioning, and design. Prominent examples include the Neurosky, Emotiv, or Muse systems. Researchers have been evaluating and comparing these devices with research- or medical-grade solutions [47-52]. In general, the research- or clinical-grade EEGs are more accurate as they employ multichannel wet sensors. For consumer-grade EEG, studies have confirmed that useful EEG readings are possible. These devices are easy to use, but involve dry electrodes with no impedance matching, making them more susceptible to artifacts. They also generally come with a lower number of channels.

Consumer-grade EEG-based systems for applications such as NFT for children with ADHD have already been discussed (eg, [29,30]). We have also published primary research results of NFT using Emotiv [53]. Experience from the trial indicated that although the multichannel Emotiv system provided good results, it was quite time-consuming to administer, as saline still needed to be applied for each sensor. The headset design was also found to be rather bulky for long-term serious use, especially if it was to be shared as part of a multiuser training station.

Alternatively, a single-channel device would provide an even cheaper solution that could expand deployment sites. In particular, the Neurosky design appeared to offer a solid headset design with a single dry electrode placed on the forehead. Although Maskeliunas et al [51] compared the Neruosky device with the multichannel Emotiv device and found the reading quality to be somewhat inferior, it was hypothesized that for NFT application, Neurosky could still be sufficient, especially as the position of the sensor is outside the frontal lobe area, where the cognitive state of interest such as attention occurs. Indeed, in contrast to the results of Maskeliunas et al [51], some authors have gone as far as stating that a single-channel device was preferred [45]. This was supported by the finding of a negative correlation between the number of bands used to compose the feedback signal and the success of training; that is, more complicated protocols that promoted or inhibited several bands worsened the final results.

The aim of this study was therefore to implement a system of brain training games with NFT on a single channel using a consumer-grade EEG headset, and to evaluate its efficacy for cognitive training in elderly participants. Physiological evidence in terms of EEG was also investigated to determine whether the brain rhythms of interest change as expected after training.

Methods

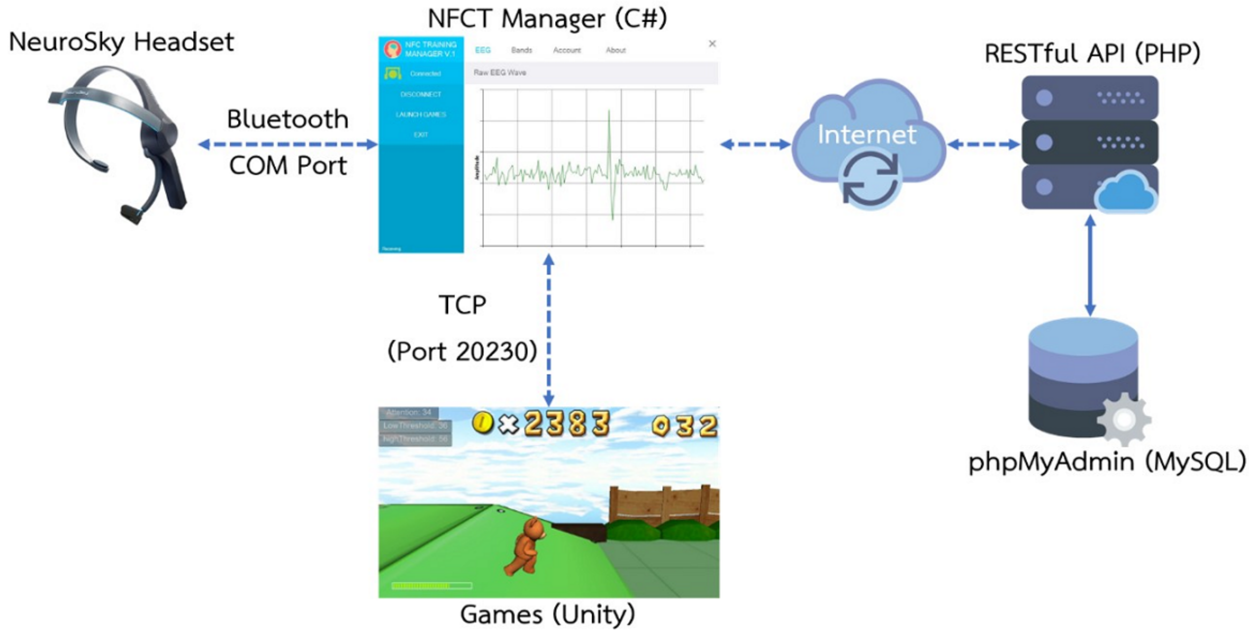
System Overview

Overall, the architecture of the NFT gaming system was designed to support multiuser application in multiple locations. At each location, the PC-based games ran locally on the computer, with the Neurosky EEG headset connected via Bluetooth to provide brain signal feedback. Training data were collected and sent to the remote database server over the internet (Figure 1).

In terms of software, the games were developed using the Unity game development platform [54], and were managed by the Neurofeedback Cognitive Training Manager program written in C#. Training data were delivered to the server via the RESTful application programming interface (API) web service. The server was programmed using PHP with the MySQL database. Parameters such as EEG readings, scoring, and usage time/characteristics of each user could be collected for analysis via the system's web service API (Figure 2).

Figure 1. System overview. API: application programming interface.

Figure 2. System configuration. NFCT: Neurofeedback Cognitive Training; API: application programming interface.



NFT Control and EEG Signal Processing

Selection of Feedback Signals

Different strategies for selecting which brain rhythms to be fed back for training had been considered, as discussed previously [45,55]. In general, there are two classical directions in NFT, focusing on either low frequencies (alpha or theta) to strengthen relaxation and focus or on high frequencies (low beta, beta, and theta) for reinforcing activation, organizing, and inhibiting distractibility. More specifically, the alpha protocol (ie, using the alpha wave as the feedback signal) has historically been used for relaxation training. Beta was considered to be associated with mental performance, and was used to improve focus or attention; alpha/theta has been used for stress reduction; gamma is a high-frequency band associated with cognitive processing and memory; and theta (including theta/beta) has been considered for anxiety, depression, ADHD, and emotional disorder treatments.

For this study, we decided on the type of feedback signals for our cognitive training scheme based on our observation that the strongest evidence in the cognitive training domain has been related to attention [26,29,30,53]. Attention was defined as the ability to focus cognitive resources on one relevant aspect of the environment while ignoring other irrelevant aspects. A significant portion of the published attention-based NFT literature suggests the use of alpha and beta rhythms [45,55-58]. For example, Li et al [58] provided evidence that the beta band (during a task) contained considerable information about attention, indicating the possibility of recognizing a subject's attention level by studying the EEG data. In addition, according to Klimesch [36], the alpha band showed task-related desynchronization; that is, it increased during resting states (especially when the eyes were closed) and decreased during performance of a cognitive task (eg, mental calculations). Therefore, it was reasonable to try to mimic the phenomena observed by the good performers by means of NFT (ie, to enhance beta power and reduce alpha power during active mental tasks) to enhance cognitive performance. Accordingly, our system targeted training of attention as the first cognitive function for improvement. Correlations between attention and other cognitive functions such as memory have previously been discussed. For example, memory showed limited capacity, and thus attention determined what would be encoded [59]. We hypothesized that improvement in higher attention could also lead to improvement in other cognitive domains.

In our previous work [53], the feedback was defined as the attention level calculated as:

$$\text{Original attention feedback} = K_\beta \times P_\beta / K_\alpha \times P_\alpha$$

Where K_β and K_α are empirical scaling constants, P_β is the power spectral density (PSD) of the beta band (12-32 Hz), and P_α is the PSD of the alpha band (8-12 Hz).

With recent studies showing that theta (4-8 Hz) training that could affect memory performance [38], with beta low (12-16 Hz) specifically found to affect attention [60], we revised the attention feedback level as follows:

$$\text{New attention feedback} = (K_\beta \times P_{\beta L})/(K_{\alpha L} \times P_{\alpha L} + K_\theta \times P_\theta)$$

Where $P_{\beta L}$ is the PSD of the beta low band (12-16 Hz), $P_{\alpha L}$ is the PSD of the alpha low band (8-10 Hz), and P_θ is the PSD of the theta band (4-8 Hz).

As the individual upper alpha (IUA) has been shown to affect cognitive power [38], the value of cognitive level monitored was then calculated from:

$$\text{Cognitive level} = (K_\beta \times P_{\beta L} + K_{\alpha h} \times P_{\alpha h})/(K_{\alpha L} \times P_{\alpha L} + K_\theta \times P_\theta)$$

Where $P_{\alpha h}$ is the PSD of the alpha high band (10-12 Hz).

Adaptive Threshold Control

In our system, the brain feedback signal (attention) was used as a game-controlling signal in an adaptive manner, with two additional parameters introduced: the upper and lower thresholds. Before each session, the system would measure the user's level of attention under two mental states: being attentive and relaxed. The difference between the two values was used to calculate the upper and lower threshold levels. Figure 3 shows an example of the condition when the lower and upper thresholds were adjusted by one third of the difference from the relaxed and attentive states, respectively. During the game play, if the user's attention level was higher than the threshold level, they were rewarded such as by having the controlled character moving faster. By contrast, if the attention level was lower than the lower threshold, the character would slow down. Attention levels between the two thresholds had no effects on the game being played. As the brain signal is highly volatile and is known to exhibit session-to-session variation, the value of the thresholds would continuously adapt to suit the present situation (Figure 4). For example, if the user became fatigued and the attention level dropped down for too long, the thresholds would adapt so that the user could hit the upper threshold more easily. This was important so that the user could feel that the game was responsive and therefore retain their engagement.

Figure 3. Setting upper and lower thresholds.

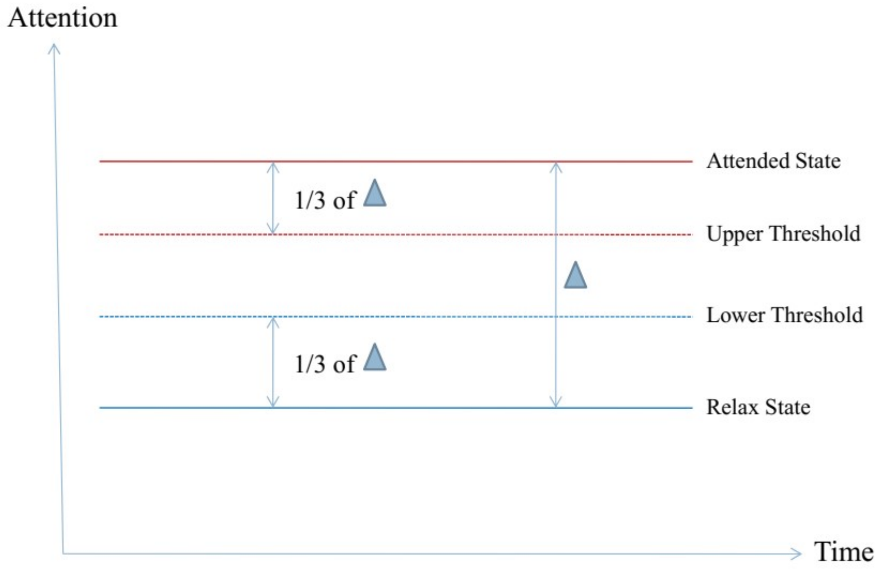
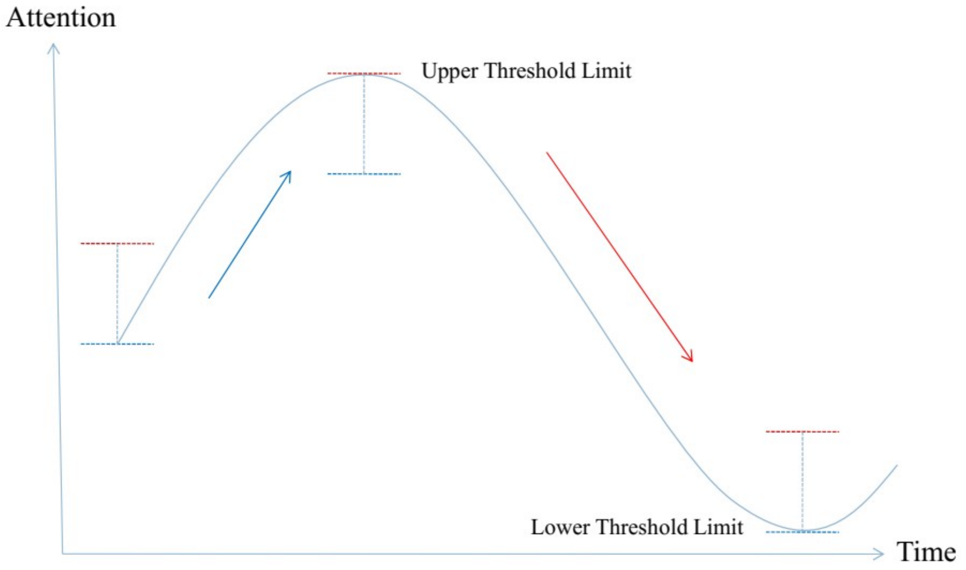


Figure 4. Adaptive threshold.



Game Description

In terms of game design, there were five games in our NFT system (Figure 5). The main concept used for all 5 games was based on the same idea that the main character's moving characteristic is controlled by the attention level feedback. The different scenarios between the games were mainly aimed at maintaining the user's engagement. In "Run Run," the first game, the bear runs faster when the player focuses on the game (ie, attention level above the upper threshold) and runs slower when the user loses focus (ie, attention below the lower threshold). The bear could also jump to collect coins if he ran fast enough, which was designed to provide the user more incentive to focus. For the second game, termed "Sunshine Day," if the player focused on the girl character to cheer her up, the girl in the garden would look happy and the character would jump to indicate that she's in a good mood. In "Cast Away," the third game, the player's attention level controls the cooking fire. There are three types of food, each of which requires a different level of cooking heat to be cooked. If the fire is strong, the food will be cooked quickly, resulting in high game scores. The fourth game was called "Paper Plane." If the attention value is high, the plane moves fast and is able to fly away from obstacles. Competing against paper planes that are computer-controlled was also provided for motivation. The last game was "321 Shoot." In this game, the average attention level was calculated every 3 seconds. If it was above the set threshold, then the basketball player scored. After every 10 points, the threshold was increased, and the basketball player would stand further away from the hoop.

Figure 5. Screenshots of games: (A) Run Run, (B) Sunshine Day, (C) Cast Away, (D) Paper Plane, (E) 321 Shoot.

Trial Protocol

Design and Setting

The trial was designed as a multisite pilot study. Forty participants were targeted for the intervention by undergoing training at the following 5 participating hospitals and health centers in Thailand: King Narai Hospital; Queen Savang Vadhana Memorial Hospital; Thai Redcross Health Station 2, Sukuman Health Center; Thai Redcross Health Station 5, Sawangkanives; and Ban Bang Khae Social Welfare Development Center for Older Persons.

Intervention results were compared with those obtained from the control group, who received care-as-usual treatment.

Participants and Eligibility

Subjects were recruited on a voluntary basis from the participating sites. All subjects who had received a diagnosis of normal aging elderly and MCI were invited to participate. Diagnoses were made by psychiatrists according to standardized clinical criteria (Thai Mental State Examination [TMSE] and Montreal Cognitive Assessment [MoCA]-Thai) and the criteria for MCI established by Petersen et al [11].

The inclusion criteria were 55-80 years old, with adequate verbal expression, visual and hearing abilities; at least 4 years of education; ≥24 on the TMSE and ≤24 on the MoCA-Thai for MCI patients, or ≥24 on the TMSE and >24 on the MoCA-Thai for normal aging elderly; and not currently enrolled in another research study.

The exclusion criteria were presence or history of a confounding central neurologic disease (eg, brain tumor, stroke, epilepsy); currently under acetyl cholinesterase inhibitor treatment (eg, donepezil, galantamine, and rivastigmine); and presence of substance abuse disorder or substance dependence.

Experimental Protocol

At the pretest period, all test subjects were assessed for cognitive functions via Cambridge Neuropsychological Test Automated Battery (CANTAB) tests. The data, together with EEG records, were collected as baseline data. The intervention period started no later than 1 week after the assessment. The intervention group trained via NFT for attention for 30 minutes per session with 2 sessions per week for a period of 10 weeks (total 20 sessions). The subjects were then evaluated for cognitive functions outcome at week 10. Posttraining EEG data were also collected for analysis.

The study was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand (593/57). All participants provided written informed consent, and the study was performed in accordance with the Declaration of Helsinki.

Cognitive Performance (CANTAB)

Delayed Matching to Sample

Delayed matching to sample (DMS) was used to assess both simultaneous visual matching ability and short-term visual recognition memory for nonverbalizable patterns. Each user was shown a complex visual pattern (abstract and nonverbal), followed by four similar patterns after a brief delay (0, 4, or 12 seconds). The pattern that exactly matched the sample had to be selected (Table 1).

Table 1. Parameters used for delayed matching to sample assessment.

Parameters	Meaning
%correct	Percentage of occasions upon which the subject selected the correct stimulus
%correct (simultaneous)	Percentage of occasions upon which the subject selected the correct stimulus in trials when the target stimulus and the three distractors were presented after the stimulus had been hidden, with delays of 0 ms
%correct (all delays)	The percentage of occasions upon which the subject selected the correct stimulus in trials when the target stimulus and the three distractors were presented after the stimulus had been hidden, with delays of 0 ms, 4000 ms, and 12,000 ms
Median correct latency	Median latency considering all occasions where the subject selected the correct stimulus on their first response in trials when the target stimulus and the three distractors were presented after the stimulus had been hidden after a specified delay. Latency was measured in milliseconds; a lower score was better
Prob error given error	Probability of an error occurring when the previous trial was responded to incorrectly

Motor Screening Task

The motor screening task (MOT) was used to assess whether sensorimotor deficits or lack of comprehension would limit the collection of valid data from the participant. For this test, colored crosses were presented in different locations on the screen, one at a time. The participant had to select the cross on the screen as quickly and accurately as possible (Table 2).

Table 2. Parameters in the motor screening task assessment.

Parameters	Meaning
Mean latency	Arithmetic mean of latency, defined as the time taken for the subject to touch the cross after it appeared. Latency was measured in milliseconds; a lower mean was better
Median latency	Arithmetic median of latency
Mean error	Measure of the subject's pointing accuracy, according to the mean distance between the center of the cross and the location the subject touched on the screen, for the 10 crosses presented to which the subject correctly responded; a lower error was better

Pattern Recognition Memory

Pattern recognition memory was applied as a test of visual pattern recognition memory. The participant was presented with a series of visual patterns, one at a time, in the center of the screen. These patterns were designed so that they could not easily be given verbal labels. In the recognition phase, the participant was required to choose between a pattern they had already seen and a novel pattern in a 2-choice forced discrimination paradigm. In addition, the test patterns in this phase were presented in reverse order to the original order of presentation. This was then repeated, immediately or after a delay, with new patterns (Table 3).

Table 3. Parameters for the pattern recognition memory assessment.

Parameters	Meaning
%correct	Number of correct responses, expressed as a percentage
Mean correct latency	Arithmetic mean of the correct responses; a lower value was better
Median correct latency	Arithmetic median of the correct responses; a lower value was better
#correct	Number of correct answers given by the subject

Rapid Visual Information Processing

The rapid visual information processing (RVP) assessment was designed to measure sustained attention. At a rate of 100 digits per minute, digits from 2 to 9 appeared in a pseudorandom order inside a white box shown at the center of the screen. Participants were requested to detect target sequences of digits (eg, 3-5-6). The button in the center of the screen had to be selected as quickly as possible when the target sequence was noted. The level of difficulty varied with either one- or three-target sequences that the participant had to watch for at the same time (Table 4).

Table 4. Parameters for rapid visual information processing assessment.

Parameters	Meaning
A'	How effective the subject was at detecting target sequences
Probability of hit	The probability of the subject responding correctly; a higher value was better
Total false alarms	Number of times the subject responded outside the response window of a target sequence; a lower value was better
Mean latency	Mean time taken to respond (milliseconds)
Median latency	Measure of the median time taken to respond
Probability of false alarm	Probability of the subject responding inappropriately, equal to total false alarms/(total false alarms + total correct rejections); a lower value was better

Spatial Span

A spatial span test was used to assess visuospatial working memory capacity. White squares were shown on the screen, some of which briefly changed color in a variable sequence. The boxes that changed color in the same order in which they were displayed by the computer (for the forward variant) or in the reverse order (for the backward variant) had to be selected. The number of boxes in the sequence increased from two at the start of the test to nine at the end of the sequence, and the colors were varied throughout the test (Table 5).

Table 5. Parameters in the spatial span assessment.

Parameters	Meaning
Span length	The longest sequence successfully recalled by the subject
Mean time to first response (span length 3)	The time the subject took to initiate problems of the span length specified by the span length option (3). The time was measured from the end of the presentation phase (the moment the final box closed) until the subject touched the screen. Attempts undertaken on spans that the subject did not pass were included in this calculation. A lower value was better.
Total errors	Number of times the subject selected an incorrect box; a lower value was better
Mean time to last response (span length 3)	Time the subject took to complete problems of the span length (3). The time was measured from the end of the presentation phase (the moment the final box closes) to the time of the subject’s final response on a given attempt. Attempts undertaken on spans that the subject did not pass were included in this calculation. A lower value was better.

Spatial Working Memory

The spatial working memory (SWM) test required the retention and manipulation of visuospatial information, and provided a measure of strategy as well as working memory errors. First, several colored squares (boxes) were shown on the screen. By selecting the boxes and using a process of elimination, the participant had to find one yellow “token” in each of a given number of boxes and use them to fill up an empty column on the righthand side of the screen. The number of boxes could be gradually increased until a maximum of 12 boxes were shown for the participants to search. The color and position of the boxes used also changed from trial to trial to discourage the use of stereotyped search strategies (Table 6).

Table 6. Parameters for spatial working memory (SWM) assessment.

Parameters	Meaning
Between errors	Results for trials containing the number of boxes specified by n (4, 6, or 8) only; a lower value was better
Strategy	Specifies the lower and upper numbers of boxes to calculate the SWM strategy measure; a lower value was better
Median time to first response	Mean time between the problem being presented to the subject and the subject first touching the screen to open a box, for problems with the specified number of boxes; a lower value was better
Median time to last response	Mean time of the subject's last response for a problem, calculated from the time between the problem being presented to the subject and the subject's last screen touch to open a box to locate the final token for the problem; a lower value was better

Physiological Brain Characteristics

Before the first training session, each participant recoded their brain EEG signals using a research-grade multichannel g.Nautilus headset from g.tec medical engineering. The g.Nautilus device is a wireless EEG sensing and amplification system with a 24-bit resolution at a 250/500 Hz sampling rate. The version used was configured to work with g.SAHARA hybrid EEG electrodes for dry or wet recordings. After asking the participant to relax, the EEG signals were recorded at 16 positions (Fp1, Fp2, F3, Fz, F4, T7, C3, Cz, C4, T8, P2, Pz, P4, PO7, PO8, and Pz) when the eyes were both open and closed. The entire procedure was repeated after the last training session (session 20).

Results

Participants

A total of 104 individuals applied for the intervention program. Seventeen were excluded from the study based on their health conditions. A further 22 candidates did not meet the TMSE and MoCA criteria. Of the remaining candidates, 43 individuals eventually participated in the program and 35 completed the training (8 dropouts).

CANTAB Results

Table 7 shows the baseline CANTAB values of the control and intervention groups. Analysis results of CANTAB readings between pre- and posttraining using paired *t* tests are shown in Table 8.

Statistically significant improvements were found in terms of DMS (%correct, %correct all delays), median latency, and SWM (between errors), meaning that there were benefits from the training in terms of visual memory (DMS), attention (MOT), and visual recognition (SWM). Improvements were also found in the control group in terms of RVP (probability of hit, probability of false alarm) and MOT (median latency, mean error).

The results indicated that after brain exercising based on the proposed NFT system, we did find improvements in both memory and attention domains as we had hoped. DMS was the test of visual matching to sample, assessing mainly visual memory, and depended heavily on a forced-choice decision that may be dependent on frontal as well as temporal lobe functions. By contrast, the SWM score depended on the ability to retain spatial information to manipulate remembered items in working memory. Therefore, the improved SWM score could indicate improvements in the frontal lobe as well as in executive function.

Table 7. Baseline information of the participants.

Characteristic	Control (n=30)	Neurofeedback training (n=35)
Age (years), mean (SD)	69.86 (8.95)	69.48 (6.15)
Female, n (%)	28 (93)	31 (89)
TSME[a], mean (SD)	27.50 (1.70)	27.09 (2.05)
MoCA[b], mean (SD)	20.83 (2.53)	21.00 (2.03)

[a]TMSE: Thai Mental State Examination.

[b]MoCA: Montreal Cognitive Assessment.

Table 8. Cambridge Neuropsychological Test Automated Battery results.

Parameter[a]	Control			Neurofeedback training		
	t (*df*=29)	SD (x – y)	*P* value	*t* (*df*=34)	SD (x – y)	*P* value
DMS[b]						
%correct	–0.5177	8.8165	.61	–3.0566	9.2627	.004
%correct (simultaneous)	1.2289	10.4	.23	–1.5082	14.5695	.14
%correct (all delays)	–0.9581	11.4336	.35	–2.6199	11.6131	.01
Median correct latency	–1.9927	1.02E+03	.06	0.3146	1.18E+03	.76
Prob error given error	–0.4167	0.2088	.68	0.7981	0.1566	.43
MOT[c]						
Mean latency	1.8352	244.0492	.08	1.0251	275.8327	.31
Median latency	3.6956	166.6858	<.001	2.7713	213.8717	.009
Mean error	–2.5181	2.7958	.02	–0.9254	5.1444	.36
PRM[d]						
%correct	–0.1606	9.472	.87	–1.1758	14.9745	.25
Mean correct latency	–0.5861	596.2864	.56	0.8479	1.18E+03	.40
Median correct latency	–1.6612	248.2209	.11	1.4163	502.0288	.17
#correct	–0.1575	2.3183	.88	–1.1758	3.5939	.25
RVP[e]						
A'	–2.041	0.0535	.05	–0.3157	0.0802	.75
Probability of hit	–2.3338	0.1766	.03	0.5545	0.1806	.58
Total false alarms	0.3692	1.978	.71	0.8318	7.5184	.41
Mean latency	–0.1099	90.2135	.91	1.6845	265.2661	.10
Median latency	0.2864	110.6188	.78	1.9901	289.9302	.05
Probability of false alarm	2.2401	0.0083	.03	0.8618	0.0325	.39
SSP[f]						
Span length	1.27	1.0063	.21	–1.0964	1.0792	.28
Mean time to first response (span length 3)	–0.9127	622.248	.37	1.2085	1.50E+03	.24
Total errors	0.1436	6.3576	.89	0.2642	7.0368	.79
Mean time to last response (span length 3)	1.5412	1.32E+03	.13	1.0786	2.43E+03	.29
SWM[g]						
Between errors	–0.2158	9.3051	.83	2.2582	12.2757	.03
Strategy	0.5781	2.8424	.57	0.0838	4.0362	.93
Median time to first response	–1.5109	654.9549	.14	0.7888	991.3451	.44
Median time to last response	–0.6447	5.89E+03	.52	1.1679	7.14E+03	.25

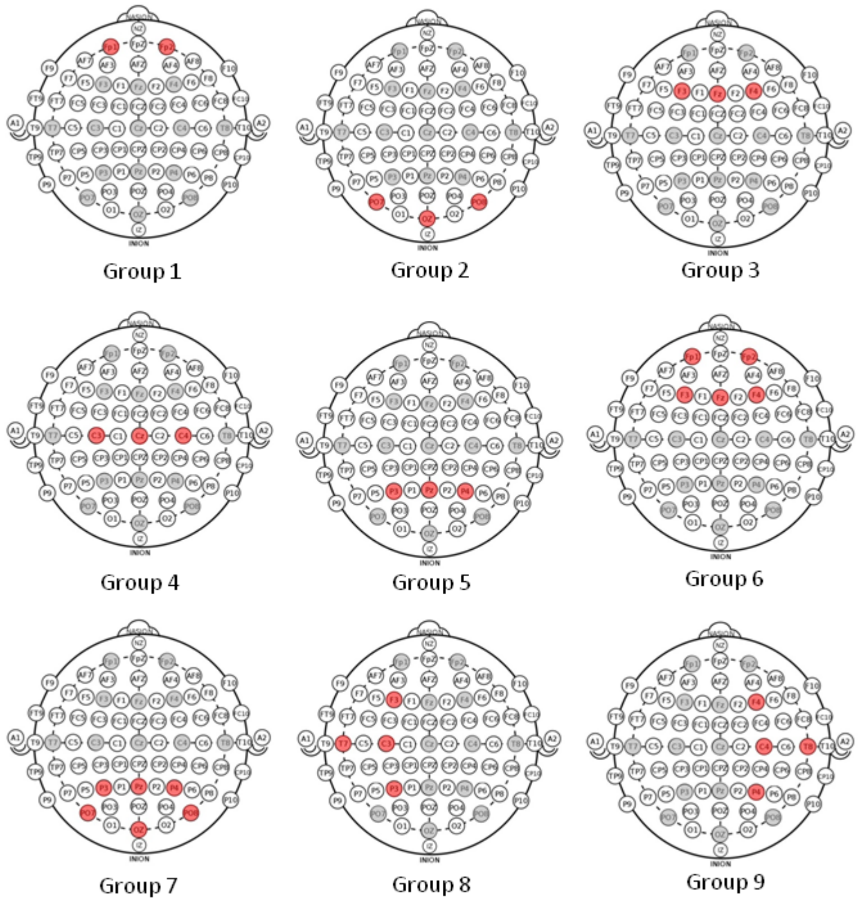
[a]See Tables 1-6 for a description of the parameters.
[b]DMS: delayed matching to sample.
[c]MOT: motor screening.
[d]PRM: pattern recognition memory.
[e]RVP: rapid visual information processing.
[f]SSP: spatial span.
[g]SWM: spatial working memory.

EEG Results

To analyze the EEG data, the channels were first grouped according to their cortical areas, as shown in Figure 6. This was necessary because of the known associations between brain patterns and where they would occur in the brain areas: the prefrontal (Group 1), occipital (Group 2), frontal (Group 3), central (Group 4), and parietal (Group 5) areas. Signals from the prefrontal and frontal areas (Group 6) and from the parietal and occipital areas (Group 7) were also analyzed together, respectively. EEGs from the temporal left and right sides were grouped as Group 8 and Group 9, respectively.

The PSD of each group was then calculated as the average of each channel's PSD within the group, with the paired *t* test calculated from pretraining and posttraining group PSD values. Significant changes in alpha high were found during the open-eyed relaxed condition in Group 2 and Group 7 (*P*=.04 for both). This agreed well with the finding of Klimesch [38] that an increase in alpha high corresponded to improvement in cognitive performance. This observation in Group 2 and Group 7, which covered the occipital areas, further supported the validity of the results as these are the areas where alpha was expected to be more evident.

Figure 6. Groupings of electroencephalogram channels.



Discussion

Principal Results

The goal of this study was to investigate whether brain training games with consumer-grade single-channel EEG sensor neurofeedback could be effective in enhancing the cognitive functions of elderly users. The results seemed to confirm its usefulness, especially in terms of memory improvements in both DMS and SWM. Compared to our prior works, SWM improvement was indeed also observed in NFT in children [30] and in our previous trial with older subjects [53]. The difference in terms of the setup between this trial and the previous trial [53] was that this trial used single-channel EEG instead of multichannel EEG (Emotiv Epoc). A single-channel EEG was more practical and easier to administer, but signal reading quality could suffer as a result. Another main difference is that this trial also employed the new feedback signal and adaptive threshold concepts. All of these factors could have contributed to the slightly different results obtained between the two studies.

For the attention domain, a significant change was found on MOT compared to significant changes found for RVP in previous studies [30,53]. This difference could also be due to the difference in population and setup. Notably, the MOT change was also found in the control group. As MOT was designed primarily for screening purposes, we could speculate that it might have a sensitivity that made it less suitable for attention assessment.

In general, the results of this study are in line with the current literature given a similar focus on two main areas of potential benefits: attention [61-63] and working memory [64,65]. Evidently, the published results to date are variable. For example, Gadea et al [66] reported no significant changes on theta downregulation but achieved significant changes in attention. By contrast, other studies [67-69] reported successful voluntary modulation of brain activity (change in the EEG signal) but no effects on the behavior (eg, symptom reduction). Both effects were shown in other studies [70-74]. A recent study concerning NFT using low-cost EEG found alpha enhancement but no effect on cognitive performance [75]. There are many factors that may have contributed to these mixed results. Experimental variables (factors) such as protocol, duration of the training, location of electrodes, or signal modality could have been set up differently, leading to these mixed results. Indeed, the large majority of neurofeedback studies appeared to have at least one major methodological limitation such as lack of randomization, nonblind designs, and use of waiting list control conditions. Controls without sham groups in a quasiexperimental design used in many trials means that evidence for the efficacy of the independent variable (ie, feedback training) could not be fully validated. Unfortunately, use of a sham control for the positive effects is ethically challenging, since the use of placebo (which also includes sham groups) instead of clinical treatments may lead to a deterioration of symptoms. Thus, the well-controlled EEG-based neurofeedback studies performed to date could only be carried out on treatment-resistant or healthy subjects. In nonblind designs, factors such as the relationship with the therapist are also relevant, especially when treating children.

Another issue regarding the effectiveness of NFT is sustainability validated through long-term studies. There have only been a few such studies, mostly targeted toward ADHD treatment. For example Lee et al [62] and Rijken et al [76] evaluated the long-term effect of NFT through follow-up. Compared to nonactive control treatments, NFT appeared to have more durable treatment effects for at least 6 months following treatment [77]. The reasons for the lack of such studies may be related to practical issues such as difficulty to reconcile schedules with the participants or withdrawal of subjects.

Despite these limitations, it is still evident from literature reviews that there is a clear scientific movement toward the use of neurofeedback as a tool to improve cognition and behavior besides rehabilitation. Moreover, it is interesting to note a shift in the direction of research in this field, characterized by a move from the classical clinical standard of evaluating the effectiveness of the method (standardized double-blind, randomized experiment) to the use of other assessment methods that seem more appropriate for neuropsychological treatments. The primary reason is related to contradictions between the positive outcomes of single case studies and the ineffectiveness of studies with large numbers of subjects [78]. This trend eventually may lead to the development of individualized treatment protocols, where effect of the treatment is assessed within the single case.

With respect to the details of the design, for simplicity, we used alpha high with a fixed frequency range, given recent studies indicating that it was possible to enhance cognitive capacities of healthy individuals by means of IUA NFT, where the IUA was located between the individual alpha peak (IAP; between 7.5 and 12.5 Hz) and IAP+2 Hz [38]. Changes from upper alpha to IUA could have potential to improve the effectiveness further.

The training time of 30 minutes per session provided a user with considerable training time, while being short enough to not cause too much fatigue. During game development and testing, we found that an individual game duration of 2-3 minutes was appropriate. Game players usually spent the first 30 seconds up to 1 minute to ramp up their attention levels. These levels were then maintained or slightly improved throughout the game. After 2 minutes, the level of attention tended to drop, and several players felt fatigue and lost attention completely after 3 minutes. As a result, all 5 games were programmed to run for 2-3 minutes. Players were also asked to take a break after every game.

Each NFT set comprised a low-cost one-channel EEG headset and a PC with internet and Bluetooth connection. Once set up, the user could simply put on the headset and start a game. This could be a low-cost addition to community centers for the elderly around the country. These NFT games could be played leisurely to supplement existing activities such as group music and craft works.

Finally, as Staufenbiel et al [79] and Wang and Hsieh [80] pointed out, the literature remains sparse in relation to assessing the impact of neurofeedback in the elderly population. Our work, showing positive benefits on both EEG and cognitive functions, hopefully contributes as an additional step toward the confirmation of benefits for NFT in the elderly.

Limitations

This study has some limitations that warrant future research. First, the sample size could be considered to be small. According to Begemann et al [81], to detect a medium effect size of 0.5, a minimal sample size of 64 per group would be needed. This criterion was not met by our study or most of the related works reported to date. Underpowered studies could carry the risk of both false-positive and false-negative findings, and are potentially more likely to be affected by publication bias, selective data analysis, and selective reporting of outcomes [82]. Moreover, as in the majority of similar research, our study design was quasiexperimental. This could also raise a question of how much the success observed in the modulation of the brain activity or behavior was due to actual training as opposed to nonspecific factors that could significantly contribute to the results. Although we have taken precautionary measures to limit these potential biases, a full randomized controlled trial with a higher number of samples would be one way to further solidify the findings. A long-term study would also be beneficial. To increase the chance for successful completion, the study could be designed with a shorter but more intensive intervention to encompass the follow-up phase with fewer possible withdrawals.

Conclusions

In this work, we have demonstrated the potential benefit of game-based brain exercise using neurofeedback. The system implemented was based on a practical single-channel EEG using new feedback and gaming techniques. Results from the 5-site pilot study have demonstrated improvement in the visual memory (DMS), attention (MOT), and visual recognition (SWM) domains. EEG analyses also showed improvement in upper alpha at a resting state (open-eyed) in the frontal area, which similarly indicated improvement in the cognitive domain.

Acknowledgments

The authors would like to acknowledge with gratitude the funding from the National Science and Technology Development Agency (NSTDA). The support we received from the staff of the NECTEC-CU Center of Technology for Cognitive Care in Elderly has also been very much appreciated. We are also thankful for local support by our partner sites, especially Sukunya Rojhprasert, PhD and Chanon Sirikul, MD from King Narai Hospital; Komwudh Koncha, MD from Queen Savang Vadhana Memorial Hospital; Mrs. Saowaluck Thanasirorat and Miss Phasita Hunarak from Thai Redcross Health Station 2; Miss Sasithon Thiengtham from Ban Bang Khae; and Mrs. Kulchalee Sukunwaranil from Thai Redcross Health Station 5.

Conflicts of Interest

None declared.

References

1. Multisectoral action for a life course approach to healthy ageing: draft global strategy and plan of action on ageing and health. 69th World Health Assembly, Geneva (A69/17). World Health Organization. 2016 Apr 22. URL: http://apps.who.int/gb/ebwha/pdf_files/WHA69/A69_17-en.pdf [accessed 2021-06-05]
2. Rao D, Luo X, Tang M, Shen Y, Huang R, Yu J, et al. Prevalence of mild cognitive impairment and its subtypes in community-dwelling residents aged 65 years or older in Guangzhou, China. Arch Gerontol Geriatr 2018;75:70-75. [doi: 10.1016/j.archger.2017.11.003] [Medline: 29197258]
3. Klimova B, Maresova P. Computer-based training programs for older people with mild cognitive impairment and/or dementia. Front Hum Neurosci 2017;11:262. [doi: 10.3389/fnhum.2017.00262] [Medline: 28559806]
4. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimers Dement 2013 Jan;9(1):63-75. [doi: 10.1016/j.jalz.2012.11.007] [Medline: 23305823]
5. Bang J, Spina S, Miller BL. Frontotemporal dementia. Lancet 2015 Oct 24;386(10004):1672-1682 [FREE Full text] [doi: 10.1016/S0140-6736(15)00461-4] [Medline: 26595641]
6. Dementia Fact Sheet. World Health Organization. 2015. URL: http://www.who.int/mediacentre/factsheets/fs362/en/ [accessed 2021-06-05]
7. Wimo A, Winblad B, Jönsson L. The worldwide societal costs of dementia: Estimates for 2009. Alzheimers Dement 2010 Mar;6(2):98-103. [doi: 10.1016/j.jalz.2010.01.010] [Medline: 20298969]
8. Reisberg B, Ferris S, de Leon MJ, Crook T. Global Deterioration Scale (GDS). Psychopharmacol Bull 1988;24(4):661-663. [Medline: 3249768]
9. Fischer P, Jungwirth S, Zehetmayer S, Weissgram S, Hoenigschnabl S, Gelpi E, et al. Conversion from subtypes of mild cognitive impairment to Alzheimer dementia. Neurology 2007 Jan 23;68(4):288-291. [doi: 10.1212/01.wnl.0000252358.03285.9d] [Medline: 17242334]
10. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L, et al. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004 Sep;256(3):240-246 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01380.x] [Medline: 15324367]
11. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
12. Alkhunizan M, Alkhenizan A, Basudan L. Prevalence of mild cognitive impairment and dementia in Saudi Arabia: a community-based study. Dement Geriatr Cogn Dis Extra 2018;8(1):98-103 [FREE Full text] [doi: 10.1159/000487231] [Medline: 29706986]
13. Mirza SS, Ikram MA, Bos D, Mihaescu R, Hofman A, Tiemeier H. Mild cognitive impairment and risk of depression and anxiety: A population-based study. Alzheimers Dement 2017 Feb;13(2):130-139. [doi: 10.1016/j.jalz.2016.06.2361] [Medline: 27520773]
14. Bae JB, Kim YJ, Han JW, Kim TH, Park JH, Lee SB, et al. Incidence of and risk factors for Alzheimer's disease and mild cognitive impairment in Korean elderly. Dement Geriatr Cogn Disord 2015;39(1-2):105-115. [doi: 10.1159/000366555] [Medline: 25401488]
15. Schneider LS, Dagerman KS, Insel P. Risk of death with atypical antipsychotic drug treatment for dementia: meta-analysis of randomized placebo-controlled trials. JAMA 2005 Oct 19;294(15):1934-1943. [doi: 10.1001/jama.294.15.1934] [Medline: 16234500]
16. Smith TGH. Antipsychotics in dementia--mortality risks and strategies to reduce prescribing. Evid Based Ment Health 2011 May;14(2):35-36. [doi: 10.1136/ebmh.14.2.35] [Medline: 21502142]

17. Vigen CL, Mack WJ, Keefe RS, Sano M, Sultzer DL, Stroup TS, et al. Cognitive effects of atypical antipsychotic medications in patients with Alzheimer's disease: outcomes from CATIE-AD. Am J Psychiatry 2011 Aug;168(8):831-839 [FREE Full text] [doi: 10.1176/appi.ajp.2011.08121844] [Medline: 21572163]
18. Jalbert JJ, Daiello LA, Eaton CB, Miller SC, Lapane KL. Antipsychotic use and the risk of diabetes in nursing home residents with dementia. Am J Geriatr Pharmacother 2011 Jun;9(3):153-163. [doi: 10.1016/j.amjopharm.2011.04.006] [Medline: 21596626]
19. Coyle H, Traynor V, Solowij N. Computerized and virtual reality cognitive training for individuals at high risk of cognitive decline: systematic review of the literature. Am J Geriatr Psychiatry 2015 Apr;23(4):335-359. [doi: 10.1016/j.jagp.2014.04.009] [Medline: 24998488]
20. Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. Am J Psychiatry 2017 Apr 01;174(4):329-340. [doi: 10.1176/appi.ajp.2016.16030360] [Medline: 27838936]
21. Reijnders J, van Heugten C, van Boxtel M. Cognitive interventions in healthy older adults and people with mild cognitive impairment: a systematic review. Ageing Res Rev 2013 Jan;12(1):263-275. [doi: 10.1016/j.arr.2012.07.003] [Medline: 22841936]
22. Simon SS, Yokomizo JE, Bottino CM. Cognitive intervention in amnestic mild cognitive impairment: a systematic review. Neurosci Biobehav Rev 2012 Apr;36(4):1163-1178. [doi: 10.1016/j.neubiorev.2012.01.007] [Medline: 22322184]
23. Omejc N, Rojc B, Battaglini PP, Marusic U. Review of the therapeutic neurofeedback method using electroencephalography: EEG Neurofeedback. Bosn J Basic Med Sci 2019 Aug 20;19(3):213-220. [doi: 10.17305/bjbms.2018.3785] [Medline: 30465705]
24. Budzynski T. From EEG to neurofeedback. In: Invans JR, Arbarbanel A, editors. Introduction to Quantitative EEG Neurofeedback. 1st ed. San Diego: Academic Press Inc; 1999.
25. Neurofeedback Overview. International Society for Neuroregulation & Research (ISNR). URL: https://www.isnr.org/neurofeedback-introduction [accessed 2021-06-05]
26. Arns M, de Ridder S, Strehl U, Breteler M, Coenen A. Efficacy of neurofeedback treatment in ADHD: the effects on inattention, impulsivity and hyperactivity: a meta-analysis. Clin EEG Neurosci 2009 Jul;40(3):180-189. [doi: 10.1177/155005940904000311] [Medline: 19715181]
27. Schoenberg PLA, David AS. Biofeedback for psychiatric disorders: a systematic review. Appl Psychophysiol Biofeedback 2014 Jun;39(2):109-135. [doi: 10.1007/s10484-014-9246-9] [Medline: 24806535]
28. Sonuga-Barke EJS, Brandeis D, Cortese S, Daley D, Ferrin M, Holtmann M, European ADHD Guidelines Group. Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. Am J Psychiatry 2013 Mar;170(3):275-289. [doi: 10.1176/appi.ajp.2012.12070991] [Medline: 23360949]
29. Sudnawa KK, Chirdkiatgumchai V, Ruangdaraganon N, Khongkhatithum C, Udomsubpayakul U, Jirayucharoensak S, et al. Effectiveness of neurofeedback versus medication for attention-deficit/hyperactivity disorder. Pediatr Int 2018 Sep;60(9):828-834. [doi: 10.1111/ped.13641] [Medline: 29931709]
30. Israsena P, Hemrungrojn S, Sukwattanasinit N, Maes M. Development and evaluation of an interactive electro-encephalogram-based neurofeedback system for training attention and attention defects in children. J Med Imaging Hlth Inform 2015 Sep 01;5(5):1045-1052. [doi: 10.1166/jmihi.2015.1492]
31. Hammond DC. Neurofeedback with anxiety and affective disorders. Child Adolesc Psychiatr Clin N Am 2005 Jan;14(1):105-123. [doi: 10.1016/j.chc.2004.07.008] [Medline: 15564054]
32. Wan L, Huang H, Schwab N, Tanner J, Rajan A, Lam NB, et al. From eyes-closed to eyes-open: Role of cholinergic projections in EC-to-EO alpha reactivity revealed by combining EEG and MRI. Hum Brain Mapp 2019 Feb 01;40(2):566-577 [FREE Full text] [doi: 10.1002/hbm.24395] [Medline: 30251753]
33. Reis J, Portugal A, Pereira M, Dias N. Alpha and theta intensive neurofeedback protocol for age related cognitive deficits. 2015 Presented at: 7th International IEEE/EMBS Conference on Neural Engineering (NER); April 22-24, 2015; Montpellier, France p. 715. [doi: 10.1109/NER.2015.7146723]
34. Navarro Gil M, Escolano Marco C, Montero-Marín J, Minguez Zafra J, Shonin E, García Campayo J. Efficacy of neurofeedback on the increase of mindfulness-related capacities in healthy individuals: a controlled trial. Mindfulness 2017 Jul 14;9(1):303-311. [doi: 10.1007/s12671-017-0775-1]
35. Hsueh J, Chen T, Chen J, Shaw F. Neurofeedback training of EEG alpha rhythm enhances episodic and working memory. Hum Brain Mapp 2016 Jul;37(7):2662-2675 [FREE Full text] [doi: 10.1002/hbm.23201] [Medline: 27038114]
36. Klimesch W. EEG-alpha rhythms and memory processes. Int J Psychophysiol 1997 Jun;26(1-3):319-340. [doi: 10.1016/s0167-8760(97)00773-3] [Medline: 9203012]
37. Angelakis E, Lubar JF, Stathopoulou S, Kounios J. Peak alpha frequency: an electroencephalographic measure of cognitive preparedness. Clin Neurophysiol 2004 Apr;115(4):887-897. [doi: 10.1016/j.clinph.2003.11.034] [Medline: 15003770]
38. Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Brain Res Rev 1999 Apr;29(2-3):169-195. [doi: 10.1016/s0165-0173(98)00056-3] [Medline: 10209231]

39. Nan W, Rodrigues JP, Ma J, Qu X, Wan F, Mak P, et al. Individual alpha neurofeedback training effect on short term memory. Int J Psychophysiol 2012 Oct;86(1):83-87. [doi: 10.1016/j.ijpsycho.2012.07.182] [Medline: 22864258]
40. Escolano C, Aguilar M, Minguez J. EEG-based upper alpha neurofeedback training improves working memory performance. Annu Int Conf IEEE Eng Med Biol Soc 2011;2011:2327-2330. [doi: 10.1109/IEMBS.2011.6090651] [Medline: 22254807]
41. Posthuma D, Neale MC, Boomsma DI, de Geus EJ. Are smarter brains running faster? Heritability of alpha peak frequency, IQ, and their interrelation. Behav Genet 2001 Nov;31(6):567-579. [doi: 10.1023/a:1013345411774] [Medline: 11838534]
42. Stankova EP, Myshkin IY. Association between individual EEG characteristics and the level of intelligence. Moscow Univ Biol Sci Bull 2016 Dec 28;71(4):256-261. [doi: 10.3103/s0096392516040118]
43. Lias S, Murat Z, Sulaiman N, Taib M. IQ Index using alpha-beta correlation of EEG power spectrum density (PSD). : IEEE; 2010 Presented at: IEEE Symposium on Industrial Electronics and Applications (ISIEA); October 3-5, 2010; Penang, Malaysia p. 612-616. [doi: 10.1109/isiea.2010.5679391]
44. Begemann MJH, Florisse EJR, van Lutterveld R, Kooyman M, Sommer IE. Efficacy of EEG neurofeedback in psychiatry: A comprehensive overview and meta-analysis. Transl Brain Rhythmicity 2016;1(1):19-29. [doi: 10.15761/tbr.1000105]
45. Rogala J, Jurewicz K, Paluch K, Kublik E, Cetnarski R, Wróbel A. The do's and don'ts of neurofeedback training: a review of the controlled studies using healthy adults. Front Hum Neurosci 2016;10:301. [doi: 10.3389/fnhum.2016.00301] [Medline: 27378892]
46. American Electroencephalographic Society. American Electroencephalographic Society guidelines for standard electrode position nomenclature. J Clin Neurophysiol 1991 Apr;8(2):200-202. [Medline: 2050819]
47. Ratti E, Waninger S, Berka C, Ruffini G, Verma A. Comparison of medical and consumer wireless EEG systems for use in clinical trials. Front Hum Neurosci 2017;11:398. [doi: 10.3389/fnhum.2017.00398] [Medline: 28824402]
48. Sawangjai P, Hompoonsup S, Leelaarporn P, Kongwudhikunakorn S, Wilaiprasitporn T. Consumer grade EEG measuring sensors as research tools: a review. IEEE Sensors J 2020 Apr 15;20(8):3996-4024. [doi: 10.1109/jsen.2019.2962874]
49. Khng KH, Mane R. Beyond BCI—Validating a wireless, consumer-grade EEG headset against a medical-grade system for evaluating EEG effects of a test anxiety intervention in school. Adv Eng Informat 2020 Aug;45:101106. [doi: 10.1016/j.aei.2020.101106]
50. Taherian S, Selitskiy D, Pau J, Claire Davies T. Are we there yet? Evaluating commercial grade brain-computer interface for control of computer applications by individuals with cerebral palsy. Disabil Rehabil Assist Technol 2017 Feb;12(2):165-174. [doi: 10.3109/17483107.2015.1111943] [Medline: 26699697]
51. Maskeliunas R, Damasevicius R, Martisius I, Vasiljevas M. Consumer-grade EEG devices: are they usable for control tasks? PeerJ 2016;4:e1746. [doi: 10.7717/peerj.1746] [Medline: 27014511]
52. Grummett TS, Leibbrandt RE, Lewis TW, DeLosAngeles D, Powers DMW, Willoughby JO, et al. Measurement of neural signals from inexpensive, wireless and dry EEG systems. Physiol Meas 2015 Jul;36(7):1469-1484. [doi: 10.1088/0967-3334/36/7/1469] [Medline: 26020164]
53. Jirayucharoensak S, Israsena P, Pan-Ngum S, Hemrungrojn S, Maes M. A game-based neurofeedback training system to enhance cognitive performance in healthy elderly subjects and in patients with amnestic mild cognitive impairment. Clin Interv Aging 2019;14:347-360. [doi: 10.2147/CIA.S189047] [Medline: 30863028]
54. Unity Technologies. URL: https://unity.com/ [accessed 2021-06-05]
55. Marzbani H, Marateb HR, Mansourian M. Neurofeedback: a comprehensive review on system design, methodology and clinical applications. Basic Clin Neurosci 2016 Apr;7(2):143-158 [FREE Full text] [doi: 10.15412/J.BCN.03070208] [Medline: 27303609]
56. Liang S, Lin C, Wu R, Chen Y, Huang T, Jung T. Monitoring driver's alertness based on the driving performance estimation and the EEG power spectrum analysis. 2005 Presented at: IEEE Engineering in Medicine and Biology 27th Annual Conference; January 17-18, 2006; Shanghai, China p. 5738-5741. [doi: 10.1109/iembs.2005.1615791]
57. Hamadicharef B, Zhang H, Guan C, Wang C, Phua K, Tee K, et al. Learning eeg-based spectral-spatial patterns for attention level measurement. 2009 Presented at: IEEE International Symposium on Circuits and Systems; May 24-27, 2009; Taipei, Taiwan p. 1465-1468. [doi: 10.1109/iscas.2009.5118043]
58. Li Y, Li X, Ratcliffe M, Liu L, Qi Y, Liu Q. A real-time EEG-based BCI system for attention recognition in ubiquitous environment. 2011 Presented at: 2011 International Workshop on Ubiquitous Affective Awareness and Intelligent Interaction; September 2011; Beijing, China p. 33-40. [doi: 10.1145/2030092.2030099]
59. Chun MM, Turk-Browne NB. Interactions between attention and memory. Curr Opin Neurobiol 2007 Apr;17(2):177-184. [doi: 10.1016/j.conb.2007.03.005] [Medline: 17379501]
60. Egner T, Gruzelier JH. EEG biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. Clin Neurophysiol 2004 Jan;115(1):131-139. [doi: 10.1016/s1388-2457(03)00353-5] [Medline: 14706480]
61. Zandi Mehran Y, Firoozabadi M, Rostami R. Improvement of neurofeedback therapy for improved attention through facilitation of brain activity using local sinusoidal extremely low frequency magnetic field exposure. Clin EEG Neurosci 2015 Apr;46(2):100-112. [doi: 10.1177/1550059414524403] [Medline: 24939868]

62. Lee T, Quek SY, Goh SJA, Phillips R, Guan C, Cheung YB, et al. A pilot randomized controlled trial using EEG-based brain-computer interface training for a Chinese-speaking group of healthy elderly. Clin Interv Aging 2015;10:217-227. [doi: 10.2147/CIA.S73955] [Medline: 25624754]
63. Yeo SN, Lee TS, Sng WT, Heo MQ, Bautista D, Cheung YB, et al. Effectiveness of a personalized brain-computer interface system for cognitive training in healthy elderly: a randomized controlled trial. J Alzheimers Dis 2018;66(1):127-138. [doi: 10.3233/JAD-180450] [Medline: 30248056]
64. Reis J, Portugal AM, Fernandes L, Afonso N, Pereira M, Sousa N, et al. An alpha and theta intensive and short neurofeedback protocol for healthy aging working-memory training. Front Aging Neurosci 2016;8:157. [doi: 10.3389/fnagi.2016.00157] [Medline: 27458369]
65. Campos da Paz VK, Garcia A, Campos da Paz Neto A, Tomaz C. SMR neurofeedback training facilitates working memory performance in healthy older adults: a behavioral and EEG study. Front Behav Neurosci 2018;12:321. [doi: 10.3389/fnbeh.2018.00321] [Medline: 30618671]
66. Gadea M, Aliño M, Garijo E, Espert R, Salvador A. Testing the benefits of neurofeedback on selective attention measured through dichotic listening. Appl Psychophysiol Biofeedback 2016 Jun;41(2):157-164. [doi: 10.1007/s10484-015-9323-8] [Medline: 26683198]
67. Naas A, Rodrigues J, Knirsch J, Sonderegger A. Neurofeedback training with a low-priced EEG device leads to faster alpha enhancement but shows no effect on cognitive performance: A single-blind, sham-feedback study. PLoS One 2019;14(9):e0211668 [FREE Full text] [doi: 10.1371/journal.pone.0211668] [Medline: 31483789]
68. Gonçalves Ó, Carvalho S, Mendes AJ, Leite J, Boggio PS. Neuromodulating attention and mind-wandering processes with a single session real time EEG. Appl Psychophysiol Biofeedback 2018 Jun;43(2):143-151. [doi: 10.1007/s10484-018-9394-4] [Medline: 29797155]
69. Sherwood MS, Parker JG, Diller EE, Ganapathy S, Bennett KB, Esquivel CR, et al. Self-directed down-regulation of auditory cortex activity mediated by real-time fMRI neurofeedback augments attentional processes, resting cerebral perfusion, and auditory activation. Neuroimage 2019 Jul 15;195:475-489 [FREE Full text] [doi: 10.1016/j.neuroimage.2019.03.078] [Medline: 30954710]
70. Zoefel B, Huster RJ, Herrmann CS. Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. Neuroimage 2011 Jan 15;54(2):1427-1431. [doi: 10.1016/j.neuroimage.2010.08.078] [Medline: 20850552]
71. Hoedlmoser K, Pecherstorfer T, Gruber G, Anderer P, Doppelmayr M, Klimesch W, et al. Instrumental conditioning of human sensorimotor rhythm (12-15 Hz) and its impact on sleep as well as declarative learning. Sleep 2008 Oct;31(10):1401-1408 [FREE Full text] [Medline: 18853937]
72. Gevensleben H, Holl B, Albrecht B, Vogel C, Schlamp D, Kratz O, et al. Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. J Child Psychol Psychiatry 2009 Jul;50(7):780-789. [doi: 10.1111/j.1469-7610.2008.02033.x] [Medline: 19207632]
73. Steiner NJ, Frenette EC, Rene KM, Brennan RT, Perrin EC. In-school neurofeedback training for ADHD: sustained improvements from a randomized control trial. Pediatrics 2014 Mar;133(3):483-492. [doi: 10.1542/peds.2013-2059] [Medline: 24534402]
74. Strehl U, Leins U, Goth G, Klinger C, Hinterberger T, Birbaumer N. Self-regulation of slow cortical potentials: a new treatment for children with attention-deficit/hyperactivity disorder. Pediatrics 2006 Nov;118(5):e1530-e1540. [doi: 10.1542/peds.2005-2478] [Medline: 17060480]
75. Naas A, Rodrigues J, Knirsch J, Sonderegger A. Neurofeedback training with a low-priced EEG device leads to faster alpha enhancement but shows no effect on cognitive performance: A single-blind, sham-feedback study. PLoS One 2019;14(9):e0211668 [FREE Full text] [doi: 10.1371/journal.pone.0211668] [Medline: 31483789]
76. Rijken NH, Soer R, de Maar E, Prins H, Teeuw WB, Peuscher J, et al. Increasing performance of professional soccer players and elite track and field athletes with peak performance training and biofeedback: a pilot study. Appl Psychophysiol Biofeedback 2016 Dec;41(4):421-430 [FREE Full text] [doi: 10.1007/s10484-016-9344-y] [Medline: 27761664]
77. Van Doren J, Arns M, Heinrich H, Vollebregt MA, Strehl U, K Loo S. Sustained effects of neurofeedback in ADHD: a systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2019 Mar;28(3):293-305 [FREE Full text] [doi: 10.1007/s00787-018-1121-4] [Medline: 29445867]
78. Rossiter T. The effectiveness of neurofeedback and stimulant drugs in treating AD/HD: Part I. Review of methodological issues. Appl Psychophysiol Biofeedback 2004 Jun;29(2):95-112. [doi: 10.1023/b:apbi.0000026636.13180.b6]
79. Staufenbiel SM, Brouwer A, Keizer AW, van Wouwe NC. Effect of beta and gamma neurofeedback on memory and intelligence in the elderly. Biol Psychol 2014 Jan;95:74-85. [doi: 10.1016/j.biopsycho.2013.05.020] [Medline: 23751914]
80. Wang J, Hsieh S. Neurofeedback training improves attention and working memory performance. Clin Neurophysiol 2013 Dec;124(12):2406-2420. [doi: 10.1016/j.clinph.2013.05.020] [Medline: 23827814]
81. Begemann MJH, Florisse EJR, van Lutterveld R, Kooyman M, Sommer IE. Efficacy of EEG neurofeedback in psychiatry: A comprehensive overview and meta-analysis. Transl Brain Rhythmicity 2016;1(1):19-29. [doi: 10.15761/tbr.1000105]
82. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013 May;14(5):365-376. [doi: 10.1038/nrn3475] [Medline: 23571845]

Abbreviations

ADHD: attention deficit hyperactivity disorder
API: application programming interface
CANTAB: Cambridge Neuropsychological Test Automated Battery
DMS: delayed matching to sample
EEG: electroencephalogram
IAP: individual alpha peak
IUA: individual upper alpha
MCI: mild cognitive impairment
MoCA: The Montreal Cognitive Assessment
MOT: motor screening task
NFT: neurofeedback training
PSD: power spectral density
RVP: rapid visual information processing
SWM: spatial working memory
TMSE: Thai Mental State Examination

Edited by N Zary; submitted 31.12.20; peer-reviewed by V Kotrajaras, C Basch, B Woo; comments to author 12.02.21; revised version received 27.03.21; accepted 17.04.21; published 15.06.21.

Please cite as:
Israsena P, Jirayucharoensak S, Hemrungrojn S, Pan-Ngum S
Brain Exercising Games With Consumer-Grade Single-Channel Electroencephalogram Neurofeedback: Pre-Post Intervention Study
JMIR Serious Games 2021;9(2):e26872
URL: https://games.jmir.org/2021/2/e26872
doi:10.2196/26872
PMID:34128816

Original Paper

How Gameful Experience Affects Public Knowledge, Attitudes, and Practices Regarding COVID-19 Among the Taiwanese Public: Cross-sectional Study

Li-Hsun Peng[1], PhD; Ming-Han Bai[2], MSc

[1]Department of Creative Design, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan

[2]Graduate School of Design, National Yunlin University of Science and Technology, Douliou, Yunlin, Taiwan

Corresponding Author:
Ming-Han Bai, MSc
Graduate School of Design
National Yunlin University of Science and Technology
123 University Road
Section 3
Douliou, Yunlin, 640301
Taiwan
Phone: 886 961055900
Email: minghan.bai@gmail.com

Abstract

Background: In 2019, with the COVID-19 pandemic sweeping across the globe, public health systems worldwide faced severe challenges. Amid the pandemic, one simulation game, *Plague Inc.*, has received substantial attention. This game has indirectly drawn greater public attention to public health issues by simulating pathogen transmission and disease symptoms.

Objective: Against this backdrop, this research investigates whether the gameful experience of *Plague Inc.* has indirectly affected public knowledge, attitudes, and practices (KAP) regarding COVID-19.

Methods: An online survey was conducted through social networking services in Taiwan from May 6-28, 2020.

Results: A total of 486 subjects participated in this study, of which 276 (56.8%) had played *Plague Inc.* This study had several findings. First, participants who had played *Plague Inc.* demonstrated higher levels of knowledge (P=.03, median 7, IQR 7-8) and attitudes (P=.007, median 8, IQR 7-8) than participants who had not played *Plague Inc.* (knowledge: median 7, IQR 6-8; attitude: median 7, IQR 6-8). Second, there was a significant correlation between creative thinking (ρ=.127, P=.04) and dominance (ρ=.122, P=.04) in attitude. Finally, there was a significant correlation between creative thinking (ρ=.126, P<.001) and dominance (ρ=.119, P=.049) in practice.

Conclusions: Serious games highlighting the theme of pathogen transmission may enhance public knowledge and attitudes regarding COVID-19. Furthermore, the creative thinking and dominance involved in gameful experiences may act as critical factors in public attitudes and practices regarding COVID-19. These findings should be further verified through experimental research in the future.

(JMIR Serious Games 2021;9(2):e26216) doi:10.2196/26216

KEYWORDS

COVID-19; knowledge; attitude; practice; serious game; gameful experience

Introduction

At the end of 2019, an outbreak of pneumonia of unknown cause was detected in Wuhan, China, and it was quickly discovered to be caused by a novel coronavirus, SARS-CoV-2 [1]. On January 21, 2020, Taiwan reported the first confirmed COVID-19 case overseas [2]. As of March 21, 2021, more than 270 million diagnosed cases and 1 million deaths had been reported globally [3]. Although vaccines are considered one of the means of prevention of infectious diseases such as COVID-19, it takes time to verify their safety before they can be brought to market [4].

In the face of the severe threat of COVID-19, implementing controls on health care services and communities has been deemed an effective strategy [5]. In addition, a general population with higher levels of knowledge regarding pathogens is considered more likely to adopt the correct preventive measures and thereby reduce the prevalence rate of infections [6,7]. Therefore, many governments in the world have proactively provided the latest valid COVID-19 information and guidelines on epidemic prevention to their citizens using mainstream media or social media [8].

Promoting public attention toward health and medical care using game feedback has been proposed, as it helps the public understand possible threats to health and medical challenges when facing a difficult situation related to public health and safety [9]. The Centers for Disease Control and Prevention in the United States and the University of Derby in the United Kingdom have also attempted to impart knowledge about virus transmission and microbiology using educational games [10,11].

Teaching medical knowledge through the form of games has been found to be as effective as traditional text-based teaching methods [12]. In addition to teaching accurate information in a guided and context-based manner, incorporating knowledge pertaining to food and drug safety into video games has been proven to effectively raise the level of relevant knowledge among youngsters [13,14]. Moreover, teaching cardiopulmonary resuscitation through serious games has also been shown to result in a higher retention rate of knowledge and skills [15,16]. Teaching hand hygiene and diarrhea prevention methods through educational games has been confirmed to enhance users' awareness and practice of public health habits [17,18]. Although serious games have potential for use in medical education and promote proactiveness in learners, it is necessary to strike a balance between instructional theories, educational content, and game interaction to effectively enhance learning performance [19].

Plague Inc. is a pathogen simulation game that attracts widespread public attention whenever a pandemic breaks out. This game introduces the public to epidemiology in an unconventional way, allowing users to obtain knowledge of the transmission and characteristics of pathogens through gaming [20]. As a form of new media, *Plague Inc.* employs simple graphics but has diverse charts and functions (Figure 1) [21]. The public may acquire scientific knowledge through intuitive simulation games. In addition, students may also learn about and analyze pathogen transmission and disease symptoms through gameful experiences [22,23].

However, the behaviors that players exhibit while playing games are not just intended for entertainment and recreation. Players' personality traits are believed to be projected onto their behavior and choices during a game [24,25]. In particular, educational issue–based games not only integrate common gaming elements such as points, badges, and rankings into their design but also offer experiences that enhance extrinsic and intrinsic motivations for players [26-29]. Gamification technology has been found to have a positive effect on health and well-being [30]. Moreover, it has also been deemed to have an important mediating role in health behaviors [31]. Gamification technology has proven to be an effective tool for clinical nursing skill training. However, users' feelings and attitudes toward the gamification content should be considered [32]. In addition, both gamification technologies and gameful experience can result in psychological or behavioral changes [33].

Figure 1. Game guild levels of Plague Inc.

Objective

This research aims to investigate and explore the effects of gameful experience of *Plague Inc.* on public knowledge, attitudes, and practices (KAP) regarding COVID-19 through online surveys.

Methods

Study Design

This research used a cross-sectional study design to explore the impact of Taiwanese gameful experiences on KAP regarding COVID-19.

From May 6-28, 2020, we collected data through an online survey, the link to which was published on social networking services (such as Facebook, PTT Bulletin Board System, internet game forums).

The target population was residents of Taiwan, the population of which was 23,591,920 [34] in April 2020. We used the sample size calculator provided by Rao Soft Inc and selected a 95% confidence level, 5% margin of error, and 23,591,920 population size; the calculated recommended sample size was 385.

This study was conducted in line with the Declaration of Helsinki. All participants were anonymous and provided informed consent. They clearly understood the research content in the online questionnaire [35]. The survey information sheet contained an introduction, background, goals, anonymity statement, and questionnaire descriptions. The average survey completion time was 5 minutes.

Survey Instrument

The online survey consisted of four parts: demographics, KAP of COVID-19, game experience of *Plague Inc.*, and a gameful experience scale.

The first part includes demographic variables such as gender, age range, education, living arrangement, and occupation.

The second part pertains to the revised COVID-19 scale. The scale contains 27 questions on the three dimensions of COVID-19—namely, knowledge, attitudes, and practices. Each dimension has 9 questions, with 0 as the lowest score and 9 as the highest. Respondents are required to select “Yes,” “No,” or “Uncertain” for each question. For each correctly answered question, one point is added to the total score of the dimension that contains the question. Moreover, no point was awarded for questions with an incorrect or “Uncertain” answer.

The original scale was used to discuss the KAP regarding Middle East respiratory syndrome coronavirus (MERS-CoV) during the annual Hajj pilgrimage [36], and it was adjusted and translated to Chinese. In addition, five experts in clinical medical treatment and public health were invited to examine the questionnaire content to ensure the questionnaire’s validity. The questionnaire is shown in Table 1.

The third part is about gameful experience of *Plague Inc.*, including “ever played *Plague Inc.*,” “time since last played,” and “total playtime ranges.” In this part, screenshots were used to describe the game’s content (Figure 2).

The fourth part is the gameful experience scale consisting of 27 questions, divided into 6 dimensions: enjoyment, absorption, creative thinking, activation, absence of negative affect, and dominance [37]. The scale is a 5-point Likert scale, with options from 1 (strongly disagree) to 5 (strongly agree).

This scale was also used in nursing-related research and was considered to demonstrate excellent reliability and validity. It can be used to evaluate nursing students’ gaming experiences during training [38]. Three practitioners from the gaming industry were invited to give suggestions on modifying the questionnaire after translation.

Table 1. Questionnaire of knowledge (K), attitudes (A), and practices (P) regarding COVID-19.

Questions	Correct response (%)
K1. The incubation period of COVID-19 is from 2 to 14 days.	81.3
K2. Stomach cramps, loss of appetite, nausea, and vomiting are the main symptoms of COVID-19.	38.9
K3. People with comorbidities are more likely to be infected.	77.0
K4. COVID-19 spreads through close contact with an infected person, such as caring for and/or living with them.	94.7
K5. The primary source of COVID-19 is a plant.	89.1
K6. Antibiotics are the first line of treatment.	52.5
K7. Vaccinations for COVID-19 are available.	91.4
K8. COVID-19 can be fatal.	96.9
K9. Isolation of patients with COVID-19 is essential to ensure effective implementation of infection control measures.	99.0
A1. Do you agree that there is no risk of contracting COVID-19 when going out and traveling?	83.7
A2. Do you agree that taking antibiotics can treat COVID-19?	48.8
A3. Do you agree that pain medication and/or fever medications cannot relieve COVID-19 symptoms?	38.7
A4. Do you agree that washing your hands with soap and water for at least 30 seconds can prevent disease transmission?	95.7
A5. Do you agree that avoiding undercooked meat or food prepared under unsanitary conditions can prevent the transmission of disease?	87.4
A6. Do you agree that avoiding contact with live animals such as bats can prevent the spread of disease?	86.6
A7. Do you agree that avoiding contact with ill people can prevent the spread of disease?	97.3
A8. Do you agree that before attending public gatherings or traveling, one must have the necessary information about COVID-19?	95.9
A9. Do you agree that reporting COVID-19 symptoms to local health authorities is essential to prevent further disease transmission?	98.8
P1. At a public gathering, did you cover your mouth when sneezing/coughing?	97.3
P2. At a public gathering, did you wash your hands with soap and water after sneezing/coughing?	59.5
P3. At a public gathering, did you wash your hands before preparing or eating foods?	80.7
P4. At a public gathering, did you wash your hands after contact with possibly contaminated surfaces or materials?	85.0
P5. At a public gathering, did you avoid close contact with people when you were sick?	91.6
P6. At a public gathering, did you wear a face mask in heavily crowded areas?	96.7
P7. At a public gathering, did you share cups or eating utensils with other people?	78.8
P8. Did you consult a health care worker during a public gathering if you had a fever, cough, or difficulty breathing?	92.4
P9. At a public gathering, did you avoid direct hand contact with your eyes, nose, and mouth?	78.0

Figure 2. Screenshot of *Plague Inc.*

Game

Plague Inc. is a strategic simulation game that simulates how various pathogens are spread throughout the world. It was released on the App Store and Google Play in 2012. In 2014, *Plague Inc: Evolved* was launched on Steam, a platform for PC games.

At the beginning of each game, players can repeatedly spread new pathogens to different countries. Using DNA points in the game, players try to achieve the goal of eliminating all of mankind by employing different modes of transmission, symptoms, and abilities, and influencing the infectivity, severity, and lethality of the viruses.

Players may indirectly control the speed at which infectious pathogens spread by selecting different routes of transmission, as shown in Figure 3. In particular, the game simulates how pathogens spread via airplanes and boats, increasing the number of infected countries and individuals.

Players may change the infectivity and lethality of pathogens by selecting different symptoms, as shown in Figure 4. They may keep the death toll from ascending too rapidly by balancing infectivity and lethality, while allowing pathogens to continue spreading through infected individuals.

Plague Inc. simulates countries with different climates and latitudes. In addition, as the game progresses, vaccines will appear to prevent players from winning. Therefore, players have to use their abilities to influence the pathogen's temperature and drug resistance, as shown in Figure 5.

Figure 3. Screenshot of Plague Inc.: disease transmission.

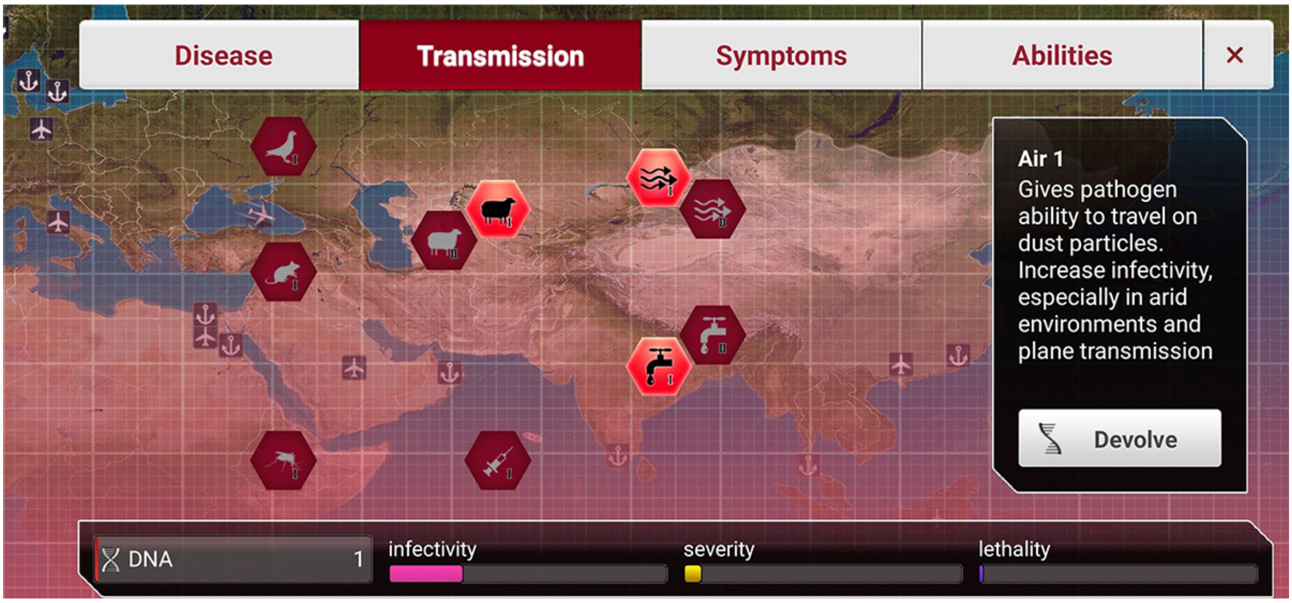


Figure 4. Screenshot of Plague Inc.: disease symptoms.

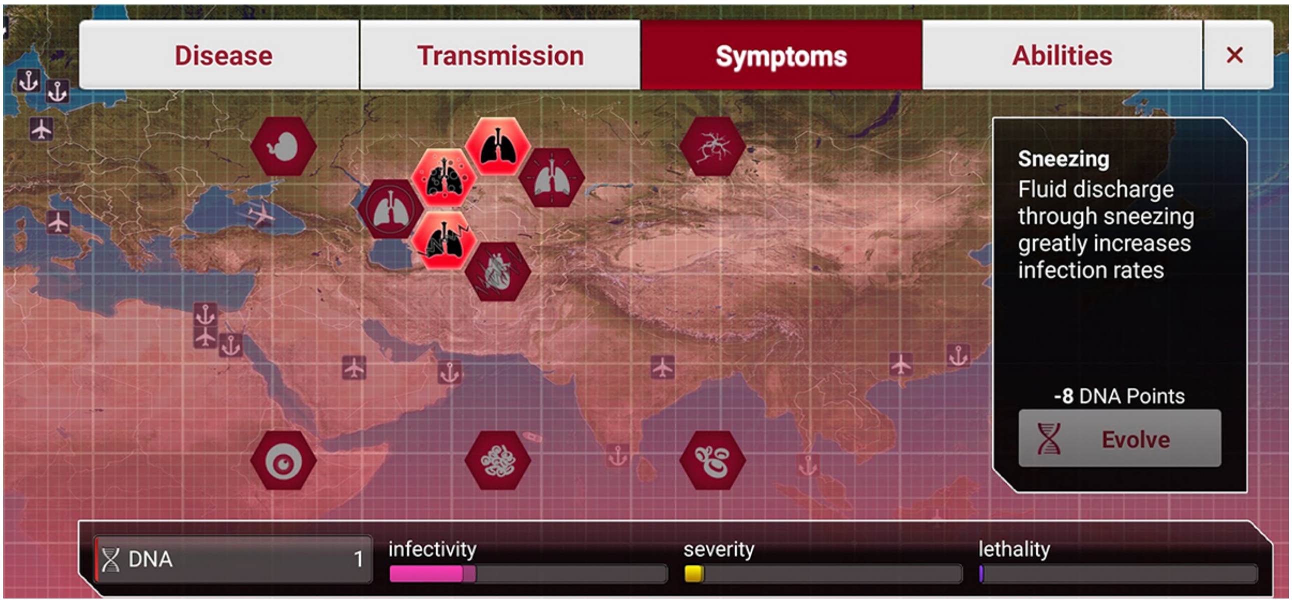
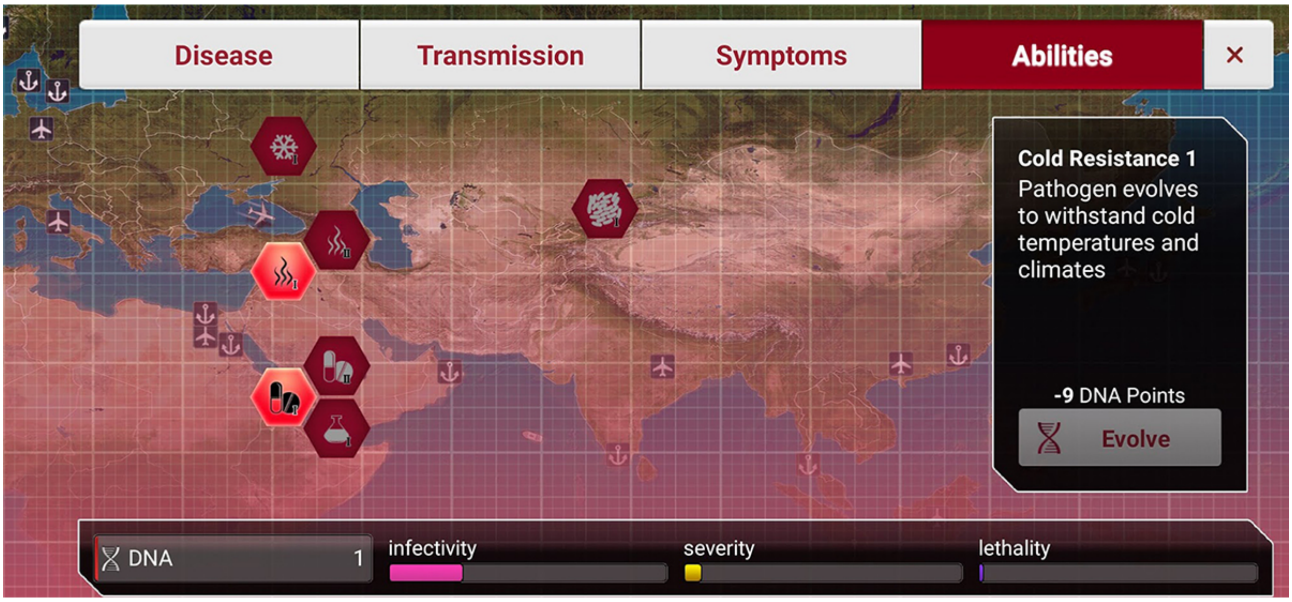


Figure 5. Screenshot of Plague Inc.: disease abilities.



Data Analysis

Based on the normality test, this study used nonparametric statistics, the Mann–Whitney *U* test, and the Kruskal–Wallis test to verify the KAP of COVID-19 to understand whether the KAP of COVID-19 was affected by gaming experience. In addition, we used the chi-square test to analyze whether there is a significant difference between whether participants have played *Plague Inc.* and their demographic characteristics and the Spearman rank correlation test to analyze the correlation between gameful experience and COVID-19 KAP.

In this study, the Cronbach α test was used to detect the internal consistency of the KAP of COVID-19 questionnaire and the gameful experience scale. For the KAP of COVID-19 questionnaire, Cronbach α=.82, while Cronbach α=.994 for the gameful experience scale; the results had acceptable internal consistency [39].

Results

A total of 486 participants were recruited for this study; demographic characteristics are shown in Table 2. Participants included 200 (41.2%) women and 286 (58.8%) men. The majority of participants were aged 20-29 years (n=323, 66.5%), had a university degree (n=305, 62.8%), were living with others (n=365, 75.1%), and were part of the labor force (n=335, 68.9%).

In total, 276 (56.8%) participants indicated they have played *Plague Inc.* and 210 (43.2%) have never played the game (Table 3).

Table 2. Demographic characteristics (N=486).

Variables	Participants, n (%)
Gender	
Male	286 (58.8)
Female	200 (41.2)
Age group (years)	
20-29	323 (66.5)
30-39	127 (26.1)
40-49	32 (6.6)
≥50	4 (0.8)
Education	
High school or below	20 (4.1)
University	305 (62.8)
Graduate school	161 (33.1)
Living arrangement	
Alone	121 (24.9)
With others	365 (75.1)
Occupation	
Labor force	335 (68.9)
Non–labor force	103 (21.2)
Unemployed	48 (9.9)

Table 3. Descriptive statistics regarding participants' game experience (N=486).

Variables	Participants, n (%)
Ever played	
Yes	276 (56.8)
No	210 (43.2)
Time since last played	
<1 week	26 (5.3)
<1 month	63 (13)
<6 months	83 (17.1)
<1 year	25 (5.1)
≥1 year	79 (16.3)
Total playtime	
<1 hour	63 (13)
1-10 hours	127 (26.1)
≥10 hours	86 (17.7)

The chi-square test was used to examine whether there is a significant difference in the demographic characteristics of participants who have played *Plague Inc.* and those who have not. The analysis results are shown in Table 4.

The analysis results show significant differences in gender (P<.001) and age (P=.006) between participants who have and have not played *Plague Inc.* However, no significant difference is observed in education, living arrangement, or occupation.

In terms of gender, males who have played *Plague Inc.* account for a higher percentage (196/286, 68.5%) than those who have not (90/286, 31.5%), whereas more females have not played *Plague Inc.* (never played: 120/200, 60%; ever played: 80/200, 40%).

As for the different age groups, among participants aged 20-29 years, more have played *Plague Inc.* (195/323, 60.4%) than not (128/323, 39.6%); among those aged 30-39 years, more have played (69/127, 54.3%) than not (58/127, 45.7%); among those aged 40-49 years, more have not played (20/32, 62.5%) than have played (12/32, 37.5%); and finally, among participants aged ≥50 years, none have played Plague Inc. before (4/4, 100%).

Table 4. Chi-square test for demographic characteristics between those who have or have not played Plague Inc. (N=486).

Variables	Participants, n	Ever played *Plague Inc.*, n (%)		Chi-square test	
		Yes (n=276)	No (n=210)	χ^2 (*df*)	*P* value
Gender				39 (1)	<.001
Male	286	196 (68.5)	90 (31.5)		
Female	200	80 (40)	120 (60)		
Age group (years)				12.1 (3)	.006
20-29	323	195 (60.4)	128 (39.6)		
30-39	127	69 (54.3)	58 (45.7)		
40-49	32	12 (37.5)	20 (62.5)		
≥50	4	0 (0)	4 (100)		
Education				1.2 (2)	.54
High school or below	20	11 (55)	9 (45)		
University	305	179 (58.7)	126 (41.3)		
Graduate school	161	86 (53.4)	75 (46.6)		
Living arrangement				1.8 (1)	.18
Alone	121	75 (62)	46 (38)		
With others	365	201 (55.1)	164 (44.9)		
Occupation				3 (2)	.22
Labor force	335	197 (58.8)	138 (41.2)		
Non–labor force	103	57 (55.3)	46 (44.7)		
Unemployed	48	22 (45.8)	26 (54.2)		

The Mann-Whitney *U* and Kruskal-Wallis tests were used to assess any differences in KAP of COVID-19 between participants who had ever or never played *Plague Inc.*, while the Kruskal-Wallis test was used to determine whether differences in KAP of COVID-19 were related to differences in time since last played and total playtime, as shown in Table 5.

There was a statistically significant difference between knowledge and attitude regarding COVID-19 for the item of whether a participant had ever played the game *Plague Inc.*

The scores on questions related to knowledge of COVID-19 were lower among never-played participants (median 7, IQR 6-8) than among ever-played participants (median 7, IQR 7-8).

The scores on the questions about attitude regarding COVID-19 were lower among never-played participants (median 7, IQR 6-8) than among ever-played participants (median 8, IQR 7-8). The results indicated that the time since last played had a significant effect for the item of attitude regarding COVID-19. The scores on the questions about attitude regarding COVID-19 were lower among never-played participants (median 7, IQR 6-8) than among those who had not played in over one year (median 8, IQR 7-9).

The results indicated that there was a significant difference in scores on the questions about attitude regarding COVID-19 among participants with different total playtimes. The scores on questions about attitude regarding COVID-19 were lower among never-played participants (median 7, IQR 6-8) than among those with a total playtime of 1-10 hours (median 8, IQR 7-8).

Spearman rank correlation was used to examine the relationship between KAP of COVID-19 and gameful experience (Table 6). The results showed that there was a significant correlation between creative thinking (ρ=.127, *P*=.04) and dominance (ρ=.122, *P*=.04) in attitude; in addition, there was a significant correlation between creative thinking (ρ=.126, *P*<.001) and dominance (ρ=.119, *P*=.049) in practice.

Table 5. Mann-Whitney U test and Kruskal-Wallis test for determining the effect of game experience on knowledge, attitudes, and practices regarding COVID-19 (N=276).

Variables	Knowledge score			Attitude score			Practice score		
	Median (IQR)	Mean (SD)	*P* value	Median (IQR)	Mean (SD)	*P* value	Median (IQR)	Mean (SD)	*P* value
Ever played									
Yes	7 (7-8)	7.31 (1.23)	.03	8 (7-8)	7.46 (1.15)	.007	8 (7-9)	7.54 (1.51)	.39
No	7 (6-8)	7.07 (1.2)		7 (6-8)	7.16 (1.27)		8 (7-9)	7.67 (1.46)	
Last played									
<1 week	7.5 (7-8)	7.58 (0.99)	.26	8 (7-8)	7.62 (0.94)	.02	8.5 (7-9)	8.12 (1.11)	.18
<1 month	7 (7-8)	7.3 (1.24)		7 (7-8)	7.32 (1.13)		8 (7-9)	7.75 (1.43)	
<6 months	7 (6-8)	7.25 (1.4)		8 (7-8)	7.4 (1.24)		8 (6-9)	7.48 (1.48)	
<1 year	8 (7-8)	7.4 (1)		7 (6-8)	7.12 (1.24)		8 (6.5-9)	7.4 (1.71)	
≥1 year	7 (6-8)	7.25 (1.16)		8 (7-9)	7.68 (1.08)		8 (6-9)	7.3 (1.62)	
Total playtime									
<1 hour	7 (6-8)	7.27 (1.31)	.10	7 (7-8)	7.32 (1.1)	.02	8 (7-9)	7.57 (1.55)	.47
1-10 hours	7 (6-8)	7.23 (1.31)		8 (7-8)	7.53 (1.23)		8 (7-9)	7.63 (1.49)	
≥10 hours	8 (7-8)	7.45 (1.01)		7 (7-8)	7.45 (1.06)		8 (6.75-9)	7.4 (1.51)	

Table 6. Spearman rank correlation of gameful experience with knowledge, attitudes, and practices regarding COVID-19 (N=324).

Variables	1	2	3	4	5	6	7	8	9
1. Knowledge	—[a]	—	—	—	—	—	—	—	—
2. Attitude	.458[b]	—	—	—	—	—	—	—	—
3. Practice	.087	.166[b]	—	—	—	—	—	—	—
4. Enjoyment	.020	.034	.049	—	—	—	—	—	—
5. Absorption	–.028	–.082	.031	.115	—	—	—	—	—
6. Creative thinking	.029	.127[c]	.126[c]	.361[b]	.351[b]	—	—	—	—
7. Activation	.076	.031	.052	.259[b]	.508[b]	.396[b]	—	—	—
8. Absence of negative affect	.067	–.087	–.115	–.279[b]	.012	–.147[c]	.087	—	—
9. Dominance	.003	.122[c]	.119[c]	.242[b]	.422[b]	.478[b]	.453[b]	–.141[c]	—

[a]Not applicable.

[b]*P*<.01.

[c]*P*<.05.

Discussion

Principal Findings

This study determined the demographics of members of the public who have played *Plague Inc.* The results show that the research participants who have played *Plague Inc.* are mainly males aged 20-39 years. Studies in Spain and Turkey also found that gender indirectly affects game type preferences [40,41]. Research in the United States discovered that although the majority of middle-aged and older adults do not play video games, age does not affect their preference for strategic simulation games [42].

Participants who have played *Plague Inc.* displayed higher levels of knowledge and attitudes toward COVID-19. Health education implemented through gameful experience has been proven to effectively enhance knowledge of leptospirosis [43]. In addition, integrating gamified media into medical education courses has been shown to effectively improve knowledge and skill performance among nursing students [44-46]. Moreover, designing applications to integrate with gameful experience to explore learning performance in professional nursing education has also been proven to generate a higher level of knowledge when compared to traditional teaching methods [15].

In particular, research participants who had played *Plague Inc.* one year before the outbreak of the COVID-19 pandemic had higher attitude scores for COVID-19 than those who had not. Designing medical clinical instructional content as a serious game and incorporating gameful experience into training may effectively enhance medical knowledge and confidence among nursing students. Moreover, students may repeatedly practice in the simulated environment, and thereby develop a self-guided learning strategy [47,48]. China's research on COVID-19–related serious games also confirmed that educational games result in better performance in learning retention than online lectures [49].

Participants who have played *Plague Inc.* indicated higher creative thinking and dominance during their gameful experience. Their attitudes and practices regarding COVID-19 have a significant positive correlation. In addition, the creative thinking involved in gameful experience is considered to be one of the significant factors that affect learning performance [50]. Gamified learning materials have been proven to stimulate creative thinking in learners and improve their attitudes and soft skills for school subjects [51]. Moreover, simulated situations can also cultivate creative thinking and indirectly improve attitudes toward public health [52,53].

The game environment of *Plague Inc.* allows players to repeatedly experiment with different pathogens and features. During this process, players gain a better understanding of how to enhance the infectivity and lethality of a pathogen. The challenge experiences and strategies accumulated through playing the game may indirectly improve the players' corresponding attitudes and practices toward pathogens. Previous research has demonstrated that the reason that effective learning performance is achieved through serious games is due to the concept of experience-based learning. In addition, with game situations designed for players to repeatedly try and challenge, such games may motivate players to reflect on the content and strategize accordingly [54].

Limitations

This study has several limitations. First, the sample number may be insufficient to represent the general public of Taiwan. Second, as the survey was conducted on social networking websites, the surveys were mainly completed by males aged 20-39 years. This study lacks a broader scope of participants. Third, the research survey was conducted in May 2020; the COVID-19 pandemic was already easing in Taiwan by then. Moreover, the government and medical units had also implemented advocacy and educational messaging through multiple channels. As such, the public had developed a basic understanding of COVID-19. Fourth, as the game investigated in this study is a strategic simulation game, research participants may indirectly represent a specific group. Finally, *Plague Inc.* was not designed specifically for COVID-19. In addition, the goal of the game is to infect and kill all human beings. A new version of the game, *The Cure*, was launched on November 11, 2020, which allows players to promote antipandemic prevention and isolation measures by roleplaying as the World Health Organization; however, the public may still be unable to acquire precise knowledge of specific viruses via this game.

Future Research

This research focuses on the relationship between public KAP regarding COVID-19 and the gameful experience of *Plague Inc.* However, there may be more factors that affect public KAP regarding COVID-19 other than the gameful experience of *Plague Inc.* Therefore, further research is required to explore possible interactions with other dimensions [31].

Conclusions

This study has demonstrated that *Plague Inc.* provides pathogen simulation scenarios that allow players to continuously seek challenges. Members of the public who have played *Plague Inc.* exhibit higher levels of knowledge and attitudes regarding COVID-19. Moreover, the creative thinking and dominance developed in the gameful experience have been found to be correlated with attitudes and practices regarding COVID-19. Reinforcing creative thinking and dominance through serious games may be a valid approach. However, further verification is required.

Finally, it is imperative to heighten public awareness of the symptoms and transmission routes of COVID-19 amid the unprecedented crisis of the current pandemic. This study suggests that new media, such as games, should be employed to help the public accumulate learning experiences regarding pathogens and further establish public KAP toward pathogens. In this manner, the public may become more heedful of public health issues in the future.

Acknowledgments

We sincerely thank all the health care workers fighting COVID-19 and all participants of this study.

Authors' Contributions

LHP dealt with conceptualization, funding acquisition, methodology, project administration, supervision, and writing (review and editing). MHB dealt with data curation, formal analysis, investigation, resources, validation, and writing the original draft. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Lu H, Stratton CW, Tang Y. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. J Med Virol 2020 Apr;92(4):401-402 [FREE Full text] [doi: 10.1002/jmv.25678] [Medline: 31950516]

2. Cheng S, Chang Y, Fan Chiang Y, Chien Y, Cheng M, Yang C, et al. First case of Coronavirus Disease 2019 (COVID-19) pneumonia in Taiwan. J Formos Med Assoc 2020 Mar;119(3):747-751 [FREE Full text] [doi: 10.1016/j.jfma.2020.02.007] [Medline: 32113824]
3. COVID-19 Weekly Epidemiological Update. World Health Organization. 2021 Mar 21. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20210323_weekly_epi_update_32.pdf [accessed 2021-03-25]
4. Kaur SP, Gupta V. COVID-19 Vaccine: A comprehensive status report. Virus Res 2020 Oct 15;288:198114 [FREE Full text] [doi: 10.1016/j.virusres.2020.198114] [Medline: 32800805]
5. Yen M, Chiu A, Schwartz J, King C, Lin Y, Chang S, et al. From SARS in 2003 to H1N1 in 2009: lessons learned from Taiwan in preparation for the next pandemic. J Hosp Infect 2014 Aug;87(4):185-193 [FREE Full text] [doi: 10.1016/j.jhin.2014.05.005] [Medline: 24996515]
6. Aburto NJ, Pevzner E, Lopez-Ridaura R, Rojas R, Lopez-Gatell H, Lazcano E, et al. Knowledge and adoption of community mitigation efforts in Mexico during the 2009 H1N1 pandemic. Am J Prev Med 2010 Nov;39(5):395-402. [doi: 10.1016/j.amepre.2010.07.011] [Medline: 20965376]
7. Lin L, Jung M, McCloud RF, Viswanath K. Media Use and Communication Inequalities in a Public Health Emergency: A Case Study of 2009–2010 Pandemic Influenza a Virus Subtype H1N1. Public Health Rep 2014 Nov 01;129(6_suppl4):49-60. [doi: 10.1177/00333549141296s408]
8. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. J Public Health Policy 2020 Dec 21;41(4):410-420 [FREE Full text] [doi: 10.1057/s41271-020-00247-w] [Medline: 32826935]
9. Brox E, Fernandez-Luque L, Tøllefsen T. Healthy Gaming – Video Game Design to promote Health. Appl Clin Inform 2017 Dec 16;02(02):128-142. [doi: 10.4338/aci-2010-10-r-0060]
10. Devlin S, Cowling P, Kudenko D, Goumagias N, Nucciareli A, Cabras I. Game Intelligence. In: 2014 IEEE Conference on Computational Intelligence and Games. 2014 Presented at: 2014 IEEE Conference on Computational Intelligence and Games; August 26-29, 2014; Dortmund, Germany. [doi: 10.1109/cig.2014.6932917]
11. Robinson LA, Turner IJ, Sweet MJ. The use of gamification in the teaching of disease epidemics and pandemics. FEMS Microbiol Lett 2018 Jun 01;365(11):1 [FREE Full text] [doi: 10.1093/femsle/fny111] [Medline: 29718203]
12. 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]
13. Crovato S, Pinto A, Giardullo P, Mascarello G, Neresini F, Ravarotto L. Food safety and young consumers: Testing a serious game as a risk communication tool. Food Control 2016 Apr;62:134-141. [doi: 10.1016/j.foodcont.2015.10.009]
14. Abraham O, Feathers A, Grieve L, Babichenko D. Developing and piloting a serious game to educate children about over - the - counter medication safety. J Pharm Health Serv Res 2019 Feb 21;10(2):235-241. [doi: 10.1111/jphs.12292]
15. Gutiérrez-Puertas L, García-Viola A, Márquez-Hernández VV, Garrido-Molina JM, Granados-Gámez G, Aguilera-Manrique G. Guess it (SVUAL): An app designed to help nursing students acquire and retain knowledge about basic and advanced life support techniques. Nurse Educ Pract 2021 Jan;50:102961. [doi: 10.1016/j.nepr.2020.102961] [Medline: 33421681]
16. Semeraro F, Frisoli A, Loconsole C, Mastronicola N, Stroppa F, Ristagno G, et al. Kids (learn how to) save lives in the school with the serious game Relive. Resuscitation 2017 Jul;116:27-32. [doi: 10.1016/j.resuscitation.2017.04.038] [Medline: 28476478]
17. Arbianingsih, Utario Y, Rustina Y, Krianto T, Ayubi D. Arbi Care application increases preschool children's hand-washing self-efficacy among preschool children. Enfermería Clínica 2018 Feb;28:27-30. [doi: 10.1016/s1130-8621(18)30031-7]
18. Higgins A, Hannan M. Improved hand hygiene technique and compliance in healthcare workers using gaming technology. J Hosp Infect 2013 May;84(1):32-37. [doi: 10.1016/j.jhin.2013.02.004] [Medline: 23498360]
19. Agudelo-Londoño S, Gorbanev I, Delgadillo V, Muñoz Ó, Cortes A, González RA, et al. Development and Evaluation of a Serious Game for Teaching ICD-10 Diagnosis Coding to Medical Students. Games Health J 2019 Oct 01;8(5):349-356. [doi: 10.1089/g4h.2018.0101] [Medline: 31166817]
20. Mitchell S, Hamilton SN. Playing at apocalypse: Reading Plague Inc. in pandemic culture. Convergence 2017 Jan 17;24(6):587-606. [doi: 10.1177/1354856516687235]
21. Lorenzo S. H5N1 for Angry Birds: Plague Inc., Mobile Games, and the Biopolitics of Outbreak Narratives. Science Fiction Studies 2016;43(1):85. [doi: 10.5621/sciefictstud.43.1.0085]
22. Cheng SL, Egues A, Cohen-Brown G. Visualizing Medicine: Mapping Connections with Plague Inc. to Learn in the Interdisciplinary Classroom. In: Lansiquot RD, MacDonald SP, editors. Interdisciplinary Place-Based Learning in Urban Education: Exploring Virtual Worlds. Cham: Springer International Publishing; 2018:111-132.
23. Nascimento FGM, Benedetti TR, Santos AR. Use of the Pest Game Inc.: a possibility for Science Education in the days of COVID-19. BJD 2020;6(5):25909-25928. [doi: 10.34117/bjdv6n5-156]
24. Bai MH. Exploring the Relationship of Player Behavior and Personality Traits Using Simulation Game. Tainan, Taiwan: National University of Tainan; 2018.
25. Worth NC, Book AS. Dimensions of video game behavior and their relationships with personality. Computers in Human Behavior 2015 Sep;50:132-140. [doi: 10.1016/j.chb.2015.03.056]

26. Chou YK. Actionable Gamification: Beyond Points, Badges and Leaderboards. Scotts Valley, CA, USA: CreateSpace Independent Publishing Platform; 2015.
27. Inoue A. Gamification - changes the business. Tokyo, Japan: NHK Shuppan; 2011.
28. Kyewski E, Krämer NC. To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. Computers & Education 2018 Mar;118:25-37. [doi: 10.1016/j.compedu.2017.11.006]
29. Feng Y, Jonathan Ye H, Yu Y, Yang C, Cui T. Gamification artifacts and crowdsourcing participation: Examining the mediating role of intrinsic motivations. Computers in Human Behavior 2018 Apr;81:124-136. [doi: 10.1016/j.chb.2017.12.018]
30. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: A systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
31. Mulcahy RF, Zainuddin N, Russell-Bennett R. Transformative value and the role of involvement in gamification and serious games for well-being. JOSM 2020 Aug 21;32(2):218-245. [doi: 10.1108/josm-05-2019-0137]
32. Gutiérrez-Puertas L, Márquez-Hernández VV, Román-López P, Rodríguez-Arrastia MJ, Ropero-Padilla C, Molina-Torres G. Escape Rooms as a Clinical Evaluation Method for Nursing Students. Clinical Simulation in Nursing 2020 Dec;49:73-80. [doi: 10.1016/j.ecns.2020.05.010]
33. Landers RN, Tondello GF, Kappen DL, Collmus AB, Mekler ED, Nacke LE. Defining gameful experience as a psychological state caused by gameplay: Replacing the term 'Gamefulness' with three distinct constructs. International Journal of Human-Computer Studies 2019 Jul;127:81-94. [doi: 10.1016/j.ijhcs.2018.08.003]
34. Latest indicators of total population. National Statistics RoCT. URL: https://eng.stat.gov.tw/point.asp?index=9 [accessed 2020-04-29]
35. World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull World Health Organ 2001;79(4):373-374 [FREE Full text] [Medline: 11357217]
36. Alhomoud F, Alhomoud F. 'Your Health Essential for Your Hajj': Muslim pilgrims' knowledge, attitudes and practices regarding Middle East respiratory syndrome coronavirus (MERS-CoV) during Hajj season. J Infect Chemother 2017 May;23(5):286-292 [FREE Full text] [doi: 10.1016/j.jiac.2017.01.006] [Medline: 28216182]
37. Eppmann R, Bekk M, Klein K. Gameful Experience in Gamification: Construction and Validation of a Gameful Experience Scale [GAMEX]. Journal of Interactive Marketing 2018 Aug;43:98-115. [doi: 10.1016/j.intmar.2018.03.002]
38. Márquez-Hernández VV, Garrido-Molina JM, Gutiérrez-Puertas L, García-Viola A, Aguilera-Manrique G, Granados-Gámez G. How to measure gamification experiences in nursing? Adaptation and validation of the Gameful Experience Scale [GAMEX]. Nurse Educ Today 2019 Oct;81:34-38. [doi: 10.1016/j.nedt.2019.07.005] [Medline: 31319349]
39. Taber KS. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Res Sci Educ 2017 Jun 7;48(6):1273-1296. [doi: 10.1007/s11165-016-9602-2]
40. Dindar M. An empirical study on gender, video game play, academic success and complex problem solving skills. Computers & Education 2018 Oct;125:39-52. [doi: 10.1016/j.compedu.2018.05.018]
41. Gómez-Gonzalvo F, Molina P, Devís-Devís J. Which are the patterns of video game use in Spanish school adolescents? Gender as a key factor. Entertainment Computing 2020 May;34:100366. [doi: 10.1016/j.entcom.2020.100366]
42. Salmon JP, Dolan SM, Drake RS, Wilson GC, Klein RM, Eskes GA. A survey of video game preferences in adults: Building better games for older adults. Entertainment Computing 2017 Jun;21:45-64. [doi: 10.1016/j.entcom.2017.04.006]
43. Azhari NN, Abdul Manaf R, Ng SW, Shakeeb Arsalaan Bajunid SFB, Mohd Gobil AR, Saad WZ, et al. Gamification, a Successful Method to Foster Leptospirosis Knowledge among University Students: A Pilot Study. Int J Environ Res Public Health 2019 Jun 14;16(12):2108 [FREE Full text] [doi: 10.3390/ijerph16122108] [Medline: 31207881]
44. Brull S, Finlayson S, Kostelec T, MacDonald R, Krenzischeck D. Using Gamification to Improve Productivity and Increase Knowledge Retention During Orientation. J Nurs Adm 2017 Sep;47(9):448-453. [doi: 10.1097/NNA.0000000000000512] [Medline: 28834805]
45. Roche CC, Wingo NP, Westfall AO, Azuero A, Dempsey DM, Willig JH. Educational Analytics: A New Frontier for Gamification? Comput Inform Nurs 2018 Sep;36(9):458-465 [FREE Full text] [doi: 10.1097/CIN.0000000000000455] [Medline: 29985195]
46. Wingo NP, Roche CC, Baker N, Dunn D, Jennings M, Pair L, et al. 'Playing for Bragging Rights': A Qualitative Study of Students' Perceptions of Gamification. J Nurs Educ 2019 Feb 01;58(2):79-85. [doi: 10.3928/01484834-20190122-04] [Medline: 30721307]
47. Tsopra R, Courtine M, Sedki K, Eap D, Cabal M, Cohen S, et al. AntibioGame®: A serious game for teaching medical students about antibiotic use. Int J Med Inform 2020 Apr;136:104074 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104074] [Medline: 31926355]
48. Tan AJQ, Lee CCS, Lin PY, Cooper S, Lau LST, Chua WL, et al. Designing and evaluating the effectiveness of a serious game for safe administration of blood transfusion: A randomized controlled trial. Nurse Educ Today 2017 Aug;55:38-44. [doi: 10.1016/j.nedt.2017.04.027] [Medline: 28521248]

49. Hu H, Xiao Y, Li H. The Effectiveness of a Serious Game Versus Online Lectures for Improving Medical Students' Coronavirus Disease 2019 Knowledge. Games Health J 2021 Jan 11:1. [doi: 10.1089/g4h.2020.0140] [Medline: 33434099]
50. Parra-González ME, Segura-Robles A, Romero-García C. Analysis of creative thinking and levels of student activation after a gamification experience. Educar 2020 Jul 20;56(2):475. [doi: 10.5565/rev/educar.1104]
51. Cahyana U, Paristiowati M, Hasyrin S, Hasanah M. Development of students' creative thinking skills in chemistry using Mobile Game-Based Learning (M-GBL) with integrated creative problem-solving. In: Rahmawati Y, Taylor PC, editors. Empowering Science and Mathematics for Global Competitiveness. London: CRC Press; 2019:7.
52. Barzilai S, Blau I. Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. Computers & Education 2014 Jan;70:65-79. [doi: 10.1016/j.compedu.2013.08.003]
53. Fitzpatrick VE, Mayer C, Sherman BR. Undergraduate Public Health Capstone Course: Teaching Evidence-Based Public Health. Front Public Health 2016 Apr 18;4:70 [FREE Full text] [doi: 10.3389/fpubh.2016.00070] [Medline: 27148516]
54. Westera W. Why and How Serious Games can Become Far More Effective Accommodating Productive Learning Experiences, Learner Motivation and the Monitoring of Learning Gains. Journal of Educational Technology & Society 2019;22(1):59-69.

Abbreviations

KAP: knowledge, attitudes, and practices
MERS-CoV: Middle East respiratory syndrome coronavirus

Edited by N Zary; submitted 02.12.20; peer-reviewed by D Lee, O Ntsweng, K Wiley; comments to author 24.12.20; revised version received 27.01.21; accepted 13.03.21; published 06.04.21.

Please cite as:
Peng LH, Bai MH
How Gameful Experience Affects Public Knowledge, Attitudes, and Practices Regarding COVID-19 Among the Taiwanese Public: Cross-sectional Study
JMIR Serious Games 2021;9(2):e26216
URL: https://games.jmir.org/2021/2/e26216
doi:10.2196/26216
PMID:33737262

Original Paper

Current Competencies of Game Facilitators and Their Potential Optimization in Higher Education: Multimethod Study

Jannicke Baalsrud Hauge[1,2*], PhD; Heinrich Söbke[3*], PhD; Thomas Bröker[4*], PhD; Theodore Lim[5*], PhD; Angelo Marco Luccini[6*], Dott Ing; Maksims Kornevs[7*], PhD; Sebastiaan Meijer[7*], PhD

[1]KTH Royal Institute of Technology, Södertalje, Sweden
[2]BIBA – Bremer Institut für Produktion und Logistik GmbH, Bremen, Germany
[3]Bauhaus-Universität Weimar, Weimar, Germany
[4]Nuremberg Institute of Technology, Nuremberg, Germany
[5]Heriot Watt University, Edinburgh, United Kingdom
[6]Succubus Interactive, Nantes, France
[7]KTH Royal Institute of Technology, Stockholm, Sweden
[*]all authors contributed equally

Corresponding Author:
Heinrich Söbke, PhD
Bauhaus-Universität Weimar
Goetheplatz 7/8
Weimar, 99423
Germany
Phone: 49 151 1445 2482
Email: heinrich.soebke@uni-weimar.de

Abstract

Background: Serious games can be a powerful learning tool in higher education. However, the literature indicates that the learning outcome in a serious game depends on the facilitators' competencies. Although professional facilitators in commercial game-based training have undergone specific instruction, facilitators in higher education cannot rely on such formal instruction, as game facilitation is only an occasional part of their teaching activities.

Objective: This study aimed to address the actual competencies of occasional game facilitators and their perceived competency deficits.

Methods: Having many years of experience as professional and occasional facilitators, we (n=7) defined requirements for the occasional game facilitator using individual reflection and focus discussion. Based on these results, guided interviews were conducted with additional occasional game facilitators (n=4) to check and extend the requirements. Finally, a group of occasional game facilitators (n=30) answered an online questionnaire based on the results of the requirement analysis and existing competency models.

Results: Our review produced the following questions: Which competencies are needed by facilitators and what are their training needs? What do current training courses for occasional game facilitators in higher education look like? How do the competencies of occasional game facilitators differ from other competencies required in higher education? The key findings of our analysis are that a mix of managerial and technical competencies is required for facilitating serious games in higher educational contexts. Further, there is a limited or no general competence model for game facilitators, and casual game facilitators rarely undergo any specific, formal training.

Conclusions: The results identified the competencies that game facilitators require and a demand for specific formal training. Thus, the study contributes to the further development of a competency model for game facilitators and enhances the efficiency of serious games.

(JMIR Serious Games 2021;9(2):e25481) doi:10.2196/25481

KEYWORDS
facilitation; higher education; competency; simulation; gaming; educational games

Introduction

Games in education offer alternative means to balance the learning of a subject, along with its absorption, and application. Games used for nonentertainment purposes are often termed *serious games* [1-3]. However, there are also several other definitions and terms, including *simulation games*, *educational games*, or *digital educational games* [2], with the games being be single player or multiplayer games. Furthermore, the usage in an educational setting differs, but within engineering education, these are often facilitated and based on experiential learning models. For instance, it might be assumed that a gameplay cycle reflects the 4 phases of Kolb's experiential learning cycle [4]: a recursive cycle of experiencing, reflecting, thinking, and acting to help learners increase their learning power.

Experiential learning, as well as gameplay, requires learners to take responsibility for their own learning. However, teachers should not overestimate the students' level of achievement in conceptual understanding [5]. A particular experience of an individual is often idiosyncratic to their perception of that experience and outside the control of the teacher. This dichotomy could encourage teachers to provide answers to students. However, the key is to avoid answers and instead develop the learners' capability to find answers on their own [4,6-8].

Purposeful learning processes based on serious games are often known as game-based learning (GBL). The learning outcomes supported with GBL are diverse. For example, meta-skills can be trained in simulation games [9,10], affective learning outcomes are achieved [11], engagement of the learner is increased [12], system interdependencies are demonstrated [13,14], and pure factual knowledge is taught [15,16], all based on GBL. De Gloria et al [17] distinguish between factual knowledge and skills as learning outcomes supported by GBL. Meanwhile, Egenfeldt-Nielsen [18] discusses a wide range of GBL scenarios and the learning outcomes achieved. In a systematic review on the use of serious games, Boyle et al [19] identify the categories of science, technology, engineering and mathematics; health; business and economics; and languages as domains for GBL. From the fundamental perspective of game studies, Klabbers [20] discusses the issues of knowledge building in games, while Shaffer has coined the term *epistemic games* [21]. Games can be seen as environments in which experiences are gained, which are processed by the learners, thus leading to learning processes. The theory of situated learning [22] is one of the approaches to understanding the nature of GBL: games may be seen as a context in which experiences may be gained. Another approach is that of sense-making [23]: the experiences generated by games must be interpreted and opened up by the learners themselves. Without an emphasis on any of these and other approaches, it becomes obvious that GBL poses different demands on the teacher compared to traditional teaching concepts, such as traditional classroom settings. Accordingly, the role of a teacher in GBL must differ from that of a traditional classroom setting. Here, the teacher as a facilitator can improve the permeance of learning by activating the affordances of GBL. Chapman et al [24] have outlined several characteristics to define experiential learning through games. Of these, the mixture of content and process (a balance between experiential activities and the underlying content or theory), the role of reflection (students gaining insight into themselves and their interactions), and meaningful relationships (getting students to see their learning in the context of the entire world) have direct contextual alignment to GBL. If properly facilitated, game-based learning has the potential to allow students to experience all 4 phases of Kolb's experiential learning cycle.

Facilitation should not be mistaken as a resource bank to help develop students' confidence in the learning process. However, facilitators must set the right ambience where students feel engaged, valued, trusted, and respected [25]. They must create a space where students can express different viewpoints without a fear of saying the wrong thing, they must be sensitive to cultural experiences, and they must adapt the pace of learning while ensuring the learners understand and absorb the subject [24]. Students need to understand why they are doing something; otherwise, they may not reach the intended learning outcomes.

This paper asks where and what facilitation competency is necessary and to what extent competency models have been applied or administered by those who engage or use games in education. This kind of implementation (of facilitation competency and of competency model) might be somewhere in between facilitating collaborative knowledge building, and being aware of patterns in students' learning behavior (as an individual and in a group) towards problem-solving. To better understand this relation, the mode of research could be reshaped to question the role of reflection with respect to the content through a facilitator's ability to recognize students' knowledge acquisition under guided and unguided independent learning; to apply knowledge building mechanisms into the GBL process; and to differentiate between GBL and gamified learning, and other blended learning (including conventional) approaches. The overarching research examines a facilitator's familiarity with competency models in GBL and whether this knowledge has any impact on pedagogy.

Consequently, an attempt has been made to assimilate the external and internal factors that are relevant to GBL, including the experience of the facilitator, the types of games used, the structure of the course, the level and size of cohorts, the equipment and technology used, and the environment. The autoethnography approach taken draws from our experiences and combines this with an advocacy and participatory approach aimed at understanding and acquiring new knowledge regarding the facilitation of GBL. The primary purpose of our study is to impart practical actions for future course design. What follows is a quantitative analysis to identify facilitator roles and to describe characteristics of facilitation (including models) that may be associated with successful GBL.

Our objective was to provide insights into the competency models for facilitators before conducting a self-observation and reflexive investigation into facilitating GBL. We then present and discuss the qualitative and quantitative analysis before concluding with some remarks on the facilitation process and associated competency models for GBL.

Methods

Our study attempted to assimilate the external and internal factors that apply to game facilitation, including the experience of the facilitator, the types of games used, the structure of the course, the level and size of cohorts, the equipment used, and the environment. Using an autoethnographic and participatory approach, we drew from our experiences to understand and acquire new knowledge on facilitating GBL, with a primary purpose to impart practical actions for future course design. A quantitative analysis followed to identify facilitator roles and to describe characteristics of facilitation (including models) that may be associated with successful GBL. The research questions to be answered were the following:

Which competencies are particularly needed by facilitators, and what are the training requirements for facilitators? What do current training courses for occasional game facilitators in higher education look like? How do the competencies of occasional game facilitators differ from other competencies required in higher education?

The study offers perspectives into competency models for facilitators from the literature (see Findings From the Literature Review subsection). We then conducted a self-observation and reflexive investigation on facilitating GBL (n=7; see Authors' Experience Review subsection). The results were complemented by guided interviews with game facilitators from the field of engineering education (n=4; see External Experience Review subsection). The outcomes were compiled into a questionnaire with closed and open questions. The closed questions included an assessment of existing competency models for game facilitation–related activities. The questions determined the degree to which the competencies included in the models were considered important for game facilitation. Open questions aimed at contributing to a supplementation of the competency models. The questionnaire was distributed via social media. The qualitative and quantitative analyses of the answers (N=30) are presented (Questionnaire subsection) and discussed (see Discussion section) in this paper. We conclude with remarks on the facilitation process and associated competency models for GBL (see Summary and Future Work subsection).

Results

Findings From Literature Review

Competencies and their models have superseded the long-established job analysis and their resulting task descriptions in human resource management. For today's fast-changing world and its complex situations, job analysis turned out to be an inefficient methodology [26]. Competency models define a structured collection of competencies that organizations can align to their objectives and strategies [27]. They support organizations to handle current and future situations [28] and "are used to hire, train, evaluate and promote employees on the same attributes" [27]. The primary perception of competency as the ability to deal with typical situations (as discussed by Mudra [29]) has meanwhile grown into the ability to deal with complex situations [28,30]. Boulter et al [31], cited in [32], suggested a process of 6 stages to define competency models: defining performance criteria, choosing a sample of people, collecting data, analyzing the data, validating the found results, and applying the model in practice.

From the rising complexity of situations and problems, the role of the facilitator has emerged to support organizations in the process of transition and transformation [33]. Expert consultants diagnose problems and prescribe a therapy, while facilitators are process consultants. Facilitators organize and release information in a methodical and interactive manner instead of focusing on the group's output [34]. In this way, they improve a group's effectiveness [35]: "the facilitator supports and guides, reassures and encourages […] and never 'teaches' the meaning of what is happening" [25].

Nelson and McFadzean [34] developed a conceptual competency model for facilitators and defined 7 of competencies: understanding context; technical, rational, interpersonal, task, and human process competencies; and other personal characteristics. These authors further differentiated these competencies into 3 levels, ranging from an initial to an expert facilitator. Stewart [33] generated a facilitator competency model from a qualitative study of groups in facilitation contexts and validated it through a survey. It comprises 5 areas of competencies: interpersonal competency (as communication and further skills), management process competency, knowledge competency, and personal characteristics.

In higher education, the role of a facilitator is just one role of several alternatives for teaching. In a literature review, Hoidn and Kärkkäinen [36] summarized instructional effectiveness and process–outcome research on competencies for effective higher education teaching. They concluded that effective teachers start their series of lessons with direct instruction and decrease control with the growing mastery of the students "to allow for independent and fluent performance by the students themselves". However, they conclude that an ultimate consensus on effective teaching is difficult because of the manifold contexts of teaching and identify enthusiasm as a key element of instructional effectiveness. Blašková et al [37] proposed a competency model for university teachers based on theoretical analysis and a survey among students. The competency model they proposed comprises professional, educational, motivational, communicational, personal, science and research, and publication competency. Each competency is connected to indicators of positive and negative behavior. Both Blašková et al [37] and Hoidn and Kärkkäinen [36] emphasize the role of motivational competency.

In a similar concept involving the range between controlled lessons and independent performance of students, Leigh and Spindler [25] adopted the idea of closed and open games to develop competencies of simulation and game facilitators. The facilitators' competencies range according to their degree of emotional detachment and the number of possible learning goals they allow. An observation from this study indicated that a facilitator's preference on open or closed games was a product of learning preferences and personality type. Kortmann and Peters [38] proposed a competency model for game facilitation, identifying the differences between facilitating a simulation game and leading generic groups. They point out that the facilitation of games should require more competencies, as games can serve more aims. By interviews with a group of facilitation experts, they validated the model. However, the authors expressed concerns that competencies at a subconscious level might have been missed.

Kortmann and Peters [38] identified similar competencies between facilitating game and leading generic group sessions, but they also indicated some dissimilarities, suggesting that game facilitators may require additional competencies. Taken together, the literature suggests that participants of games play within a model that requires game facilitators to have additional abilities. Game facilitators have to be aware that games are a means of learning about real-life situations and thus need to switch and adapt quickly to different roles and styles during gameplay. To facilitate a session and delegate control to the game environment, they have to comprehend the game as an immersive instrument and understand its elements and mechanics. Finally, facilitators have to translate between the game environment and reality so that the participants transfer their experiences to the real world and back [38].

Authors' Experience Review

This section synthesizes our experiences as game facilitators (the author group). We reflected individually and independently upon a set of questions organized in 3 blocks: (1) experience with game facilitation, (2) knowledge on competency models, and (3) required competencies and the biggest discrepancy perceived between required and present competency.

The following information was collected to understand our experience with game facilitation: types of games used, number of games used, description of the educational settings for using the games, reasons for using games in teaching, changes in our perspective towards teaching, changes of our attitudes towards teaching, changes over time in the facilitation, and if we applied GBL in different settings.

The analysis shows that the group had extensive experience with facilitation. All but one participant had facilitated several games, a few had done so for decades, and all had done so for at least 4 years. Most of the group facilitated games within engineering or management studies at undergraduate and postgraduate level. Both fields have a strong tradition of using experiential learning [39,40]. In addition, 5 of the participants also facilitated games on an executive and vocational level.

We found that our group had played a large variety of games, had designed and developed games, and had used games developed by others. The games comprised multiplayer games, and both board or haptic games and digital games. The topics addressed in the games could be divided into managerial competency development and engineering topics, like bridge construction, sustainable manufacturing, and product development. Specific games addressing well-known problems, such as the bullwhip effect and capacity constraints (ie, where the player learns about consequences of a decision and strategies for solving a problem) were applied, as well as games addressing communication, cooperation, and team skills. A difference between graduate courses and undergraduate courses became apparent. At the undergraduate level, the focus is on factual and procedural knowledge reflection. At the master level, it is on developing new knowledge and higher-order thinking. This is in line with how universities structure their undergraduate and postgraduate studies and with current taxonomies of learning objectives [41].

Although we have many years of experience in facilitating games, few of us had any formal pedagogical background when we began our careers as game facilitators. None of us could be classified as a professional facilitator. Even if we facilitated regularly, it was never the primary task of our work obligations. Only MK and SM had game facilitation and GBL as a formal part of previous education at the master's level, and a few of us (SM, MK, and JBH) had aspects of GBL as a part of their PhD studies.

We were expected to be familiar with competency models—either explicitly or implicitly; information collected was from a general, teaching, and facilitation perspective. We asked ourselves whether we ever felt the need for a formal competency model during our game facilitation activities.

Our overall knowledge about competency models was inhomogeneous; 3 of us (AML, JBH, TB) neither knew of any models nor missed their absence, while 4 (HS, SM, TL, MK) had a general understanding or awareness of competency models. Various competency models were listed, but all were more relevant for higher education and vocational training [42-45]. One author (AML) pointed out that most studies on facilitation only cover small groups (fewer than 100 participants). We did not specifically feel that a formal competency model was lacking. Except for 2 (MK and JBH) people, we never felt a need for a formal competency model. All agreed that a formal competency model might be useful when changing and adapting guidelines for game facilitation. Two authors commented that facilitation of highly customized and user-specific games is not supported by existing competency models. The collected data indicate that the roles of the facilitator change depending on whether the facilitator handles the whole teaching unit or only manages the game facilitation.

Regarding which competencies the authors felt were required, 2 further questions were asked: (1) Which competencies are important for the successful use of games from your point of view? (2) For which competencies do you see the biggest discrepancies between requirements and actual competencies? Where is there a need for training?

Most of us explained that we had little pedagogical knowledge when we started facilitating games (and teaching in general). The general consensus was to have the following competencies: "active listening", as with "reactive "and "proactive abilities" to act on group reactions using strategies like "team management," "participation techniques," "consensus techniques," "community management empathy," "conflict resolution," and "flexibility"; "ability to assess pure facilitation techniques" and to "integrate experiential learning principles" of "moderation," "mentoring," and "instructional capabilities under GBL settings"; and an "understanding of the toolsets" that can be implemented.

These led us to identify the following challenges: digital skills, leadership capabilities, cognitive science, and motivational skills.

Another major issue which was discussed was the preference for the facilitator to ask closed questions. In higher education it is important to have learners synthesize and create new knowledge. Many students lack this ability; however, the games are often designed for this purpose. Facilitators need to have skills to encourage perspective taking and reasoning to foster this process. Regarding the facilitation of large groups, there are many digital tools that can support such facilitation. One hypothesis, based on our reflection in discussions, is that increased digital competencies among teachers and students could ease the use of these tools. As this has to be considered as a relevant topic, more information on tools and methods can be found in Multimedia Appendix 1.

Our collective experience as authors is, as stated earlier, in engineering and management in higher education, mostly at the university level. This is limited when analyzing the overall field of facilitation in higher education. Hence, additional data were collected from a wider audience through structured interviews and an online survey.

External Experience Review

This section analyses the experiences of 4 interviewees. The results were compared to the results of the author group (using the same questions) and the survey outcome. The 4 participants teach at different faculties, but all within engineering. Their experience in game facilitation ranged between 3 to over 10 years. All had been heavily involved in the development process of least one game. This differed from the author group, whose members had been additionally involved in the development of commercial off-the-shelf games. The topics the interviewees taught were urban mobility planning, traffic simulation, health care logistics, and games for ideation and innovation.

One facilitator used the same game twice in a course. The game was integrated into an existing curriculum and used at the beginning and at the end of the course in a workshop setting with undergraduate and graduate students, so that the students could experience the learning progress throughout the course. Another facilitator used the same game throughout the semester with graduate engineering students, deployed through a blended learning environment. The curriculum was based on (and constructed around) the game scenarios. The third facilitator had used the game in around 25 sessions. It was embedded in a course, similar to the experience of the first facilitator and was played in one room with a physical and a digital component. The fourth facilitator used a game in a workshop setting (half a day) for undergraduate students in civil engineering. It was a blended learning concept (as in the second case), comprising briefing, playing, and lecture. In all 4 cases, the usage of games was initiated because of research activities and all chairs having extensive experience in GBL. All 4 facilitators had at least 2 years of experience in facilitating games. They agreed that games can be motivating and deliver a different way of teaching. One emphasized the interactive and active learning activities in GBL.

Regarding the perspective on facilitation, the 2 participants with more extensive experience in facilitating explained more and paid more attention to the introductory setting than the participants with less experience. One interviewee reported that the game can lead to frustration, but also to high engagement. However, overall, interviewee knowledge about competency models was limited: 2 participants professed to having no knowledge of the models, 1 had a general understanding, and 1 knew about the different competency models but not related to teaching. None had ever used a competency model for facilitation. However, 3 stated to have overlooked or at least missed it. Before gathering their own experience with game facilitation, they would have liked to be more aware of the following: connection between game design and facilitation process; how to observe and what to observe; how to assess and conclude the game process; how to know what aspects or knowledge needs to be assessed; how to understand the players' game decisions, "soft data'" as behavior of players, the level of communication with others etc; and how to observe learners in relation to providing feedback.

The final interview questions focused on the competencies the interviewees felt were required and on the gap. All interviewees mentioned the importance of knowing the subject, the tools used, and the technical environment as a required competency. One pointed out that it is imperative to also understand the methods behind GBL and that connecting gameplay and the intended learning outcome is a key success factor. Two of the interviewees mentioned the need for motivational competency. Three saw the necessity for a facilitator to integrate the game in a larger context from a didactical perspective. One interviewee mentioned the ability to respect and regard the competency of the students.

The answers concerning the gap focused on the lack of formal pedagogical courses on game didactics and how to integrate games into the curriculum. The interviewees also saw a deficiency in competency for facilitating groups of extreme sizes (ie, fewer than 4 or over 100 students) and how to deal with a group's inhomogeneity in relation to the game runtime and different levels of knowledge. Further, 1 interviewee mentioned a dearth of methods for nurturing the reflection competency among the students during the game session and in the debriefing phase.

The results of the qualitatively analysis indicated a large uniformity related to facilitation and on the perception of required competencies. In order to obtain quantifiable data and a broader data source, an online survey was designed.

Questionnaire

This section describes the development of the questionnaire and its results. The questionnaire supplemented and refined the qualitative data with quantitative data on facilitation and competencies. The questionnaire was drafted by JBH, HS, and TB. It was followed by a pretest and a subsequent discussion among all authors to validate it. The questionnaire comprised 5 parts: (1) demographic data, (2) general questions on game facilitation, (3) questions on the most challenging game facilitation, (4) a section on the importance of a competency model, and (5) personal training received. The questionnaire concluded with an open question for comment on any other important topic.

The questionnaire was distributed via social media in the authors' personal networks and professional communities between June 1, 2020, and June 13, 2020, with 30 participants taking part in the survey. In the remainder of this section, the results specifically relevant for facilitating are described. Further results are included in Multimedia Appendix 2.

Demographics

The participants' teaching experience in higher education ranged from 2-30 years, with an average of 12.5 years (SD 7.02). There was an average of 9.3 years of game facilitation experience (SD 6.73; range 1-30). In addition, the participants stated that they had facilitated an average of 12.2 (SD 20.32) games. Here, the span was wide, ranging from 1 to 100, with 1 value (999) being considered an outlier. As most participants had facilitated more than one game, this percentage differs from the percentage of educational scenarios. In terms of positions, 37% (11/30) of participants, the largest group, were professors and senior lecturers, 23% (7/30) were senior academics, 23% (7/30) were other academics, 13% (4/30) were PhD students, and the remaining 3% (1/30) included other positions.

When asked about the rationale for the use of games (with multiple selections being permitted), 83.% (25/30) of participants indicated their own personal initiative, 67% (20/30) responded that a game was the best fit for the intended learning outcomes, 57% (17/30) used games to contribute to multifaceted teaching methods, 30% (9/30) of the participants stated that the game was available at the respective institution, and only 13.3% (4/30) of the participants used games because games were part of the curriculum. Overall, it can be stated that the games in higher education are not systematically anchored and are rather used because of the personal initiative of the lecturers.

Game facilitation comprises various roles. Thus, the participants were asked to rank 6 predefined roles taken during facilitation. For determining the statistical parameters, nonranked roles were given the value 7. Figure 1 shows that the role of the facilitator (average rank 2.3) and the role of the moderator (average rank 2.6) were considered the most important. The role of the instructor, which was ranked third (average rank 3.6), was followed by the mentor role (average rank 4.1). The presenter role (average rank 4.8) and the referee role (average rank 4.9) occupied the last 2 ranks. Overall, the results show a coherent picture of the importance of the different roles of a game facilitator. Table 1 shows the distribution of game types facilitated by participants.

Figure 1. Ranked roles in game facilitating (7-point Likert scale; N=30).

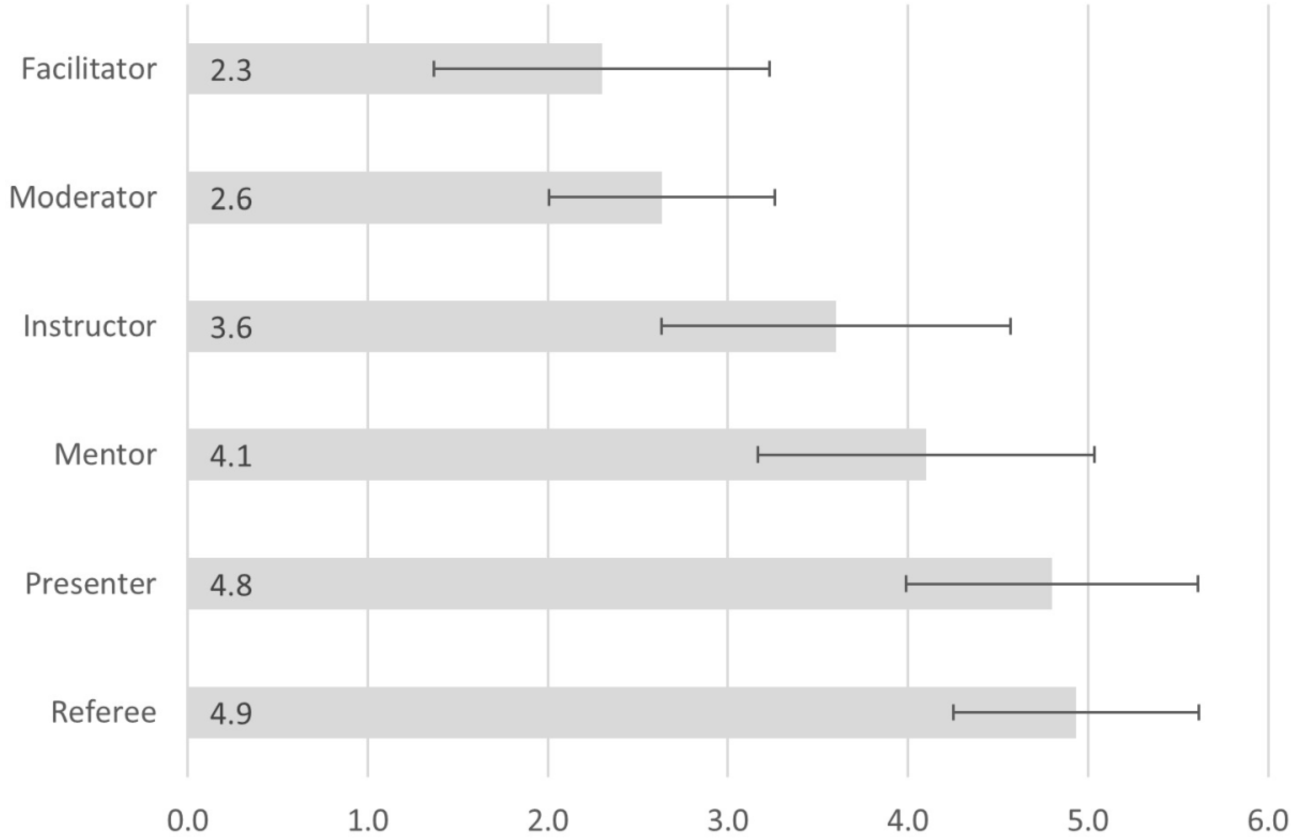


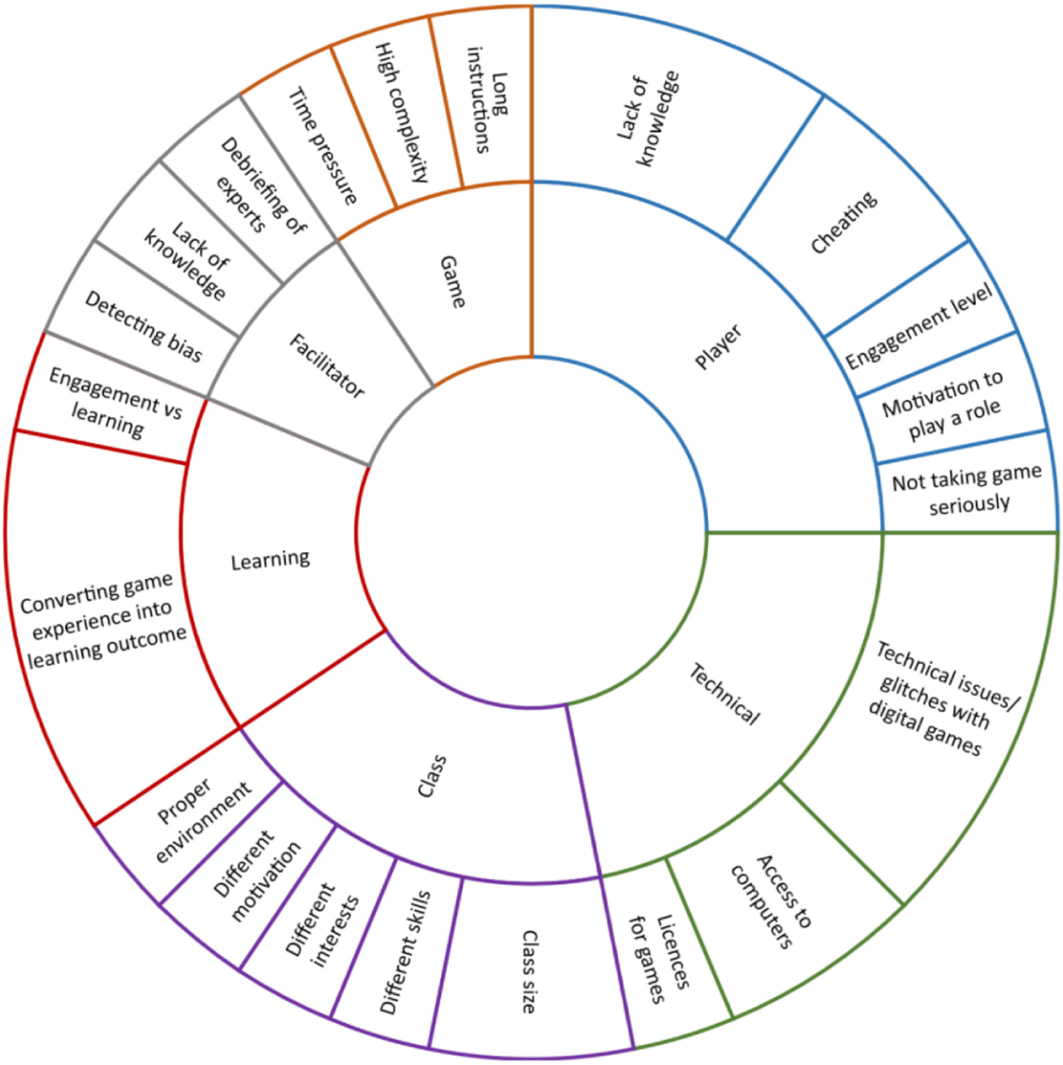
Table 1. Game types facilitated (multiple selections per respondent; N=30).

Game type	Frequency, n (%)
Specially developed serious games	22 (73)
Analogue games (eg, board games)	18 (60)
Analogue simulation games	14 (47)
Blended serious games (ie online and offline and /or digital and analogue)	14 (47)
Off-the-shelf games	13 (43)
Commercial entertainment games	12 (40)
“Modded” commercial entertainment games (ie, technically adapted)	6 (20)
Other	3 (10)

Challenges in Game Facilitation

With an open-ended question, the participants were requested to describe their most challenging game facilitation and to specify the challenges they faced. Using theme analysis, we identified 32 challenges from 18 of the answers. With each challenge being unique, they were clustered into groups based on common themes and shared characteristics. For example, some participants mentioned that players do not have sufficient knowledge or that players cheat. Although these examples are different, they still describe the characteristics or behavior of players. Hence, to ease the understanding of all challenges in game facilitation, they were grouped into 6 clusters: challenges related to (1) individual players and their actions, (2) technical aspects of using games, (3) class and collective aspects of players, (4) learning aspects of games, (5) games themselves, and (6) facilitators. Figure 2 shows these groups and challenges. The inner circle shows the clustered challenges and the outer circle the challenges themselves.

Figure 2. Cluster diagram of challenges in game facilitation as elicited by open-ended questions (N=18).



Challenges related to players include lack of knowledge to make all required decisions in a game. An example here might be a game for urban planning where students lack knowledge associated to economic aspects related to planning decisions. Furthermore, some students may not consider a game as a serious pedagogical tool. Convincing students to engage in game activities is challenging even without the added difficulty of students not fulfilling their role in the game.

Some challenges are technical in the context of digital games. Issues with software, device compatibility, or bugs or glitches in the game itself can be highly disruptive. Another challenge is student accessibility to computers or the internet to participate in a game.

Cohort sizes cause additional challenges on implementing games for learning, as discussed in the External Experience Review subsection and in Multimedia Appendix 1. A further issue is maintaining motivation for all players by ensuring the right environment during the entire game process. Mixed groups, in terms of skills, interests, and motivation can also be a challenge to the game facilitator.

As the purpose of games for learning is to ensure certain learning outcomes, converting the experiences of playing a game into these outcomes remains another major challenge. A concern mentioned by participants is that many students do not recognize the link between games and the subject of the overall program. Therefore, finding a proper balance between engagement and learning is a necessity. Immersion is important, but too high a level may hinder the learning effects.

Difficulties related to gameplay have implications for learning. The learning curve with respect to the game's usage and interaction can add additional complexity into a game learning experience. However, overly simple games may be boring and insufficient to achieving learning outcomes, while overcomplicated games may be too hard, not engaging, and may distract from intended learning outcomes.

Finally, some challenges are related to the facilitators themselves. Situations where a facilitator does not detect cognitive biases in individual and collaborative activities would consequently lead to an incorrect assessment of the game experience. Facilitating a game where experts are present can be a challenge too, because of the demand to define the problem and solutions. The latter challenges relate directly to game facilitation and reflect the competencies required of game facilitators. The challenges in the other categories in Figure 1 were reported by game facilitators to be of lower priority for game facilitation (Table 2). Many of the challenges may be not relevant during the facilitation itself, but become more pertinent when preparing the GBL activity and the infrastructure required. Thus, further expertise might be necessary, including, for example, when managing technical problems, such as licenses or access to computers, or when grappling with game-related challenges, such as high complexity or long instructions. Overall, the challenges outlined above suggest that successful use of GBL depends on good game facilitation, but also on many other aspects. The following section examines what training game facilitators receive to master these challenges.

Table 2. Prioritization of competencies (5-point Likert scale; N=30).

Competency	Stewart[a] (mean)	Study (mean)	Difference
Interpersonal competencies (communication skills)			
Verbal	4.9	4.5	–0.4
Nonverbal	4.6	3.4	–1.2
Written	4.2	3.2	–1.0
Questioning	4.8	4.2	–0.6
Active listening	4.8	4.2	–0.6
Perceptive listening	4.6	4.1	–0.5
Empathy	4.3	4.1	–0.2
Summarizing/paraphrasing	4.6	3.9	–0.7
Sensitivity to group	4.6	4.3	–0.3
Interpersonal competencies (further skills)			
Sensitivity to underlying emotions	4.5	3.9	–0.6
Cultural awareness	4.5	3.8	–0.7
Encouragement of participation	4.4	4.3	–0.1
Negotiation skills	4.5	3.5	–1.0
Flexibility	4.8	4.3	–0.5
Conflict recognition	4.5	4.0	–0.5
Conflict resolution	4.3	3.7	–0.6
Conflict transformation	4.2	3.6	–0.6
Leadership	4.1	3.7	–0.4
Motivating others to achieve goals [b]	*4.0*	*4.1*	*0.1*
Motivating others to participate creatively	*4.2*	*4.4*	*0.2*
Recognizing/rewarding achievement	*3.5*	*4.1*	*0.6*
Model neutrality	4.6	3.8	–0.8
Building relationships	4.3	3.9	–0.4
Management process competencies			
Planning/organizing	*4.4*	*4.5*	*0.1*
Managing time	4.5	4.3	–0.2
Managing audiovisual aids	4.4	3.5	–0.9
Managing physical environment	4.4	3.4	–1.0
Assimilating information	4.1	3.8	–0.3
Coaching others	*3.8*	*3.9*	*0.1*
Managing feedback	4.4	4.3	–0.1
Managing contract	4.1	3.2	–0.9
Understanding context competencies			
Understanding organizational context	4.4	4.0	–0.4
Knowledge of theory and application of group facilitation	4.1	3.9	–0.2
Personal characteristics			
Adaptability	4.7	4.6	–0.1
Intellectual agility	4.5	4.3	–0.2
Trustworthiness	4.6	4.1	–0.5

Competency	Stewart[a] (mean)	Study (mean)	Difference
Results motivation	4.3	3.7	–0.6
Objectivity	4.5	3.9	–0.6
Emotional resilience	4.7	3.9	–0.8
Self-awareness	4.6	3.9	–0.7
Self-development	4.3	3.8	–0.5

[a]Values in this column are from the paper by Stewart [33].

[b]Italics indicate competencies which scored higher for game facilitation over group facilitation.

Facilitation Training

The participants surveyed were also asked about the facilitation training they received (Table 3): 87% (26/30) selected "Learning by doing", 57% (17/30) selected "Co-facilitating with colleagues", and 47% (14/30) selected "Work shadowing with colleagues". It is worth noting that 2 of the 3 responses ("Learning by doing" and "Co-facilitation with colleagues") cannot be considered structured training and only a third of the participants received formal training ("Formal course in university pedagogy"). "Supervision by experienced colleagues" was mentioned by only a quarter of the participants, with supplier-specific training accounting for the remaining portion of replies. Overall, the proportion of those with formal training was rather low. This low rate of formal training may reflect the overall low rate of structured pedagogical training in university teaching.

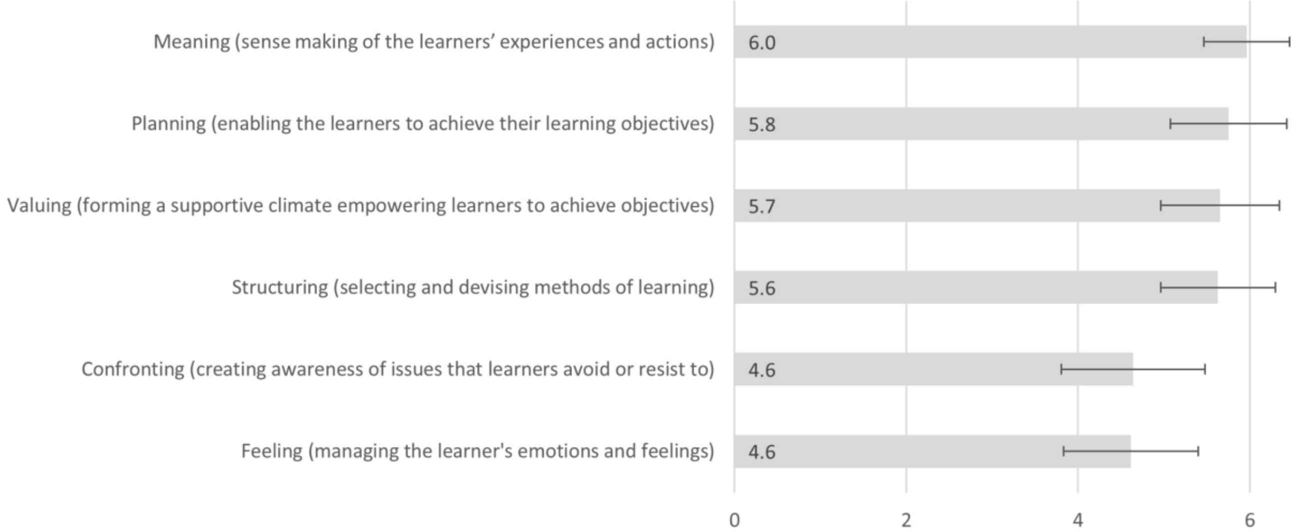
The taxonomy of Heron [46] was used to determine the extent to which training was helpful. Six dimensions of facilitation were defined in this taxonomy. On a 7-point Likert scale, participants were asked to rate for each dimension the degree to which the training received supported the respective dimension of facilitation (Figure 3). At the top of the scale with a value of 6.0 points was the dimension "Meaning", on which, as evidenced by the low SD, the participants largely agreed. The other dimensions with similarly high scores but larger SDs were "Planning" (mean score 5.8), "Valuing" (mean score 5.7), and "Structuring" (mean score 5.6). The dimensions "Confronting" and "Valuing" were rated on average 1 point (mean score 4.6) lower.

In summary, the high scores for all dimensions indicate that training was generally perceived as helpful. It remains to be clarified whether this assessment is because of the comparatively low amount of formal training, or whether the quality of facilitation might be strengthened by further targeted training.

Table 3. Types of training received (multiple selections per respondent; N=30).

Type of training	Frequency, n (%)
Learning by doing	26 (87)
Cofacilitating with colleagues	17 (57)
Work shadowing with colleagues	14 (47)
Formal course in university pedagogy	10 (33)
Supervision by experienced colleagues	8 (27)
Training course at the supplier or third party institution	7 (23)
None	2 (7)
Other	1 (3)

Figure 3. Helpfulness of trainings for categories according to Heron [46] (7-point Likert scale; N=27).



Competency Model

Stewart's [33] competency model for group facilitators comprises 41 competencies categorized in 5 groups. The competence model was selected because group facilitation is usually included in game facilitation. Thus, Stewart's model represents a valid initial foundation for further tailoring to game facilitation processes, as similarly concluded by Kortmann and Peters [38]. In this study, the participants were asked to rate the importance of each of the competencies for game facilitators on 5-point Likert scales (Table 2). Almost without exception, game facilitator competencies received lower scores for the importance of each competency.

Table 4 lists all competency group scores of the participants surveyed. The difference of 0.7 points for the group of communication skills as the most important competency group for group facilitation, is striking. For game facilitation covered by this study, all scores are in the narrow range of 3.9 to 4.0 and below the scores for group facilitation investigated by Stewart [33]. Potentially, games being in the focus of GBL activities act as compensating media, which require less competence from game facilitators than is necessary for group facilitators.

On average, the competencies of Stewart's model [33] received for game facilitation scored 0.5 points lower compared to group facilitation. To determine which competencies were important for game facilitation compared to group facilitation, only competencies which scored higher for game facilitation over group facilitation were selected (Table 2, italics). The most significant competence was recognizing/rewarding achievement" with a difference of 0.6, indicating a typical characteristic of games: recognition and rewards. In second with a difference of 0.2 was motivating others to participate creatively, which is associated with games where motivation plays a prominent role. Competences with a difference of 0.1 included motivating others to achieve goals. Finally, coaching others and planning/organizing were considered by the respondents more important for game facilitation over group facilitation.

Table 4. Difference of prioritization compared to Stewart [33] per competency group.

Competency group	Stewart[a] (mean)	Study (mean)	Difference
Interpersonal competencies (communication skills)	4.6	3.9	–0.7
Interpersonal competencies (further skills)	4.2	3.9	–0.2
Management process competencies	4.3	3.9	–0.4
Understanding context competencies	4.3	4.0	–0.3
Personal characteristics	4.5	4.0	–0.5

[a]Values in this column are from the paper by Stewart [33].

Discussion

Study Outline

Over the past decades, there has been a shift from teacher-centric toward learner-centric learning models [47], which has also nurtured the uptake of experiential learning, including GBL. This shift has also affected the teacher's role, but it remains unclear how this change has influenced the teacher's competencies regarding facilitation within higher education. Teachers within higher education often lack the formal pedagogical education that teachers within the K12 system have. In this study, therefore, the following 3 research questions were investigated: (1) Which competencies are particularly needed by the facilitator, and what are the training needs of the facilitator? (2) What do the relevant training courses for occasional game facilitators in higher education look like? (3) How do the competencies of occasional game facilitators differ from the competencies needed for other learning approaches, such as lectures, problem-based Learning, or online education?

Which Competencies Are Particularly Needed by the Facilitator, and What Are the Training Needs of the Facilitator?

Core competencies according to the interviews and authors' reflection were divided into 2 major groups: managerial and technical competencies. The more technical competencies comprise topics such as the knowledge of the gameplay, game content, its connection to the intended learning outcome, and the operation of any technical infrastructure. The more managerial competencies include active listening and reactive and proactive abilities to act on groups reactions, through the use of strategies like team management, leadership, motivational and participation techniques, consensus techniques, community management empathy, conflict resolution, and flexibility. This is underpinned by the survey (Table 2), which prioritized the competencies for game facilitation designated by Stewart [33]: verbal communication received a score 4.5 (out of 5), motivating others to participate creatively scored 4.4, and flexibility and encouraging participation (all interpersonal competencies) both scored 4.3. The personal characteristics adaptability and intellectual agility scored 4.6 and 4.3, respectively. Competencies like planning and organizing (score 4.5), and managing time and feedback (both scores 4.3) were also seen as core competencies, which is in line with the qualitative results.

The second part of this research asked about the need for training. The answer to this was complicated by the low proportion of people who received training (formal and informal). Among the interviewees, none had formal pedagogical training. This holds for two-thirds (67%, 20/30; Table 2) of the survey group, while we seem to have a higher percentage of formal course completion (10/30, 33.3%; Table 2). This might be because more than half of us hold faculty positions with requirements on pedagogical training for higher education. The qualitative results clearly state the need for training in the connection of the pedagogy and gameplay from a different perspective.

The actual training in the area of game facilitation for all was low, which can be seen also in the following wish list of the interviewees on the addressed training topics before first facilitation: connection between game design and facilitation process (Heron: meaning); how to observe and what to observe (Heron: meaning); how to assess and conclude the game process; how to know what aspects of knowledge need to be assessed (Heron: planning and structuring); how to holistically understand players' game decisions and “soft data” as representations of player behavior, and the level of communication with others (Heron: confronting and feeling); and how to observe players for feedback purposes.

According to Figure 3, training in general was perceived as helpful. Looking at the wish list, it becomes apparent that a training program focusing on facilitation would support the formal competency development and also fill a need identified by many of the respondents.

What Do the Relevant Training Courses for Occasional Game Facilitators in Higher Education Look Like?

The answer to this question was somewhat negative (Table 2). In the survey, 87% (26/30) indicated that their training consisted of learning by doing. In the author group, fewer than 20% (1/7) had received formal training before starting their game facilitation. Overall, the share of formal training remains low. Compared with K12 teachers or professional vocational trainers, this low rate of formal training may reflect the overall low rate of structured pedagogical training in university teaching. There is, however, a change in higher education. Although, scientific excellence used to play an overarching role in applying for academic faculty positions, there has been a shift toward also focusing on teaching experience and competency over the past decade. This may relate also to an increased focus on the process quality within higher education [48]. More and more countries have imposed formal requirements on pedagogical competencies, as a part of the appraisal procedure. Maturity models are widely used for assessing the maturity level within a specific area [49,50]. Despite not being frequently used, different maturity models also exist for higher education. For example, Zhou [51] has developed a capability maturity model of the e-learning process. Game facilitation is not primary about e-learning, as many games are haptic or board games, but Zhou includes the dimension of process capability in his model, which is transferable for game facilitation and higher education. It comprises the following levels (those transferrable from e-learning to GBL): delivery, delivers facilitated GBL units; planned, outlines clear and measurable objectives for GBL projects; defined, provides a defined process for development and support of GBL; managed, ensures the quality of the resources and the deliveries; and optimizing, continually improves in all aspects of the process. In matching the outcomes of the received training focusing on facilitation, it can be concluded that maturity is maximum at level 1of the model proposed by Zhou [51]. This might also be a reason why the uptake is so low. Staff interested in using games for education must undergo a time-consuming process of learning by doing. This leads to difficulties in delivering an acceptable quality of teaching during the first years occupying this role. When we consider the fact that most facilitators only use these methods once or twice a year, it is easy to see why there might be a problem.

As there is arguably hardly any training for facilitation of games provided and an increasing number of universities are offering programs on didactics in higher education, it would be valuable to know how the participants in this study rate the differences in competency needs.

How Do the Competencies Of Occasional Game Facilitators Differ From The Competencies Needed for Other Learning Approaches, Such as Lectures, Problem-Based Learning, or Online Education?

Table 2 and Table 4 illustrate this issue. With reference to Stewart's competency model, it can clearly be stated that the participants see large differences in quite a few areas. For communication skills, the overall tendency is that this is less important in the facilitation of games, where nonverbal and written communication showed the highest difference (1.2 and 1.0, respectively), but is nonetheless still relevant. This might be seen in the light of the answer given to the question of what type of games the participants use: a large majority indicated that they use games with in-built communication. For interpersonal skills, there are also large deviations, specifically on negotiation (difference of 1.0); however, these skills might also considered part of the gameplay, and thus not so relevant for the facilitation. According to Table 2, the differences for competencies on motivation related to goal achievements, creative participation, and reward and recognition are quite low; however, there is a higher need for these competencies. These are also competencies that were identified on the wish list in the qualitative part of the study. On the other hand, personal characteristic competencies seem to be less needed.

Summary and Future Work

It can be concluded that the maturity of game facilitation in higher education is low. There is a need for formal training courses, with competency models rarely being implemented in this field. Besides implementing training for game facilitators, further approaches are available for increasing the diffusion of serious games and the effectiveness of GBL. Figure 2 provides an overview of the challenges involved—such as motivating players—to be covered by approaches for increasing the diffusion of serious games. For example, giving learners a choice to take part in GBL activities or to engage in some other learning activities is likely to increase learner motivation in the chosen learning activity [15,52]. Likewise, the choice of the game used itself has a great impact on learning success: learners have preferences for games depending on learner traits, such as age, ethnicity, and gender [53,54], but certain game mechanics might be especially suited for GBL activities [55,56].

This multimethod study investigated the competencies essential for game facilitation in higher education and analyzed, with use of empirical data, the perceived gap between essential and existing competencies. This paper also discusses if there is a structured approach for competency development for the target group. The findings indicate that there is a limited or no general competence model for game facilitators and that casual game facilitators rarely undergo any specific, formal training. The lack of specific training is to be regarded as one reason for the lack of dissemination of games in higher education. The study provides the basis for a competence model for game facilitators that may serve as a prerequisite for the development of specific trainings. Future work includes the confirmation, consolidation, and refinement of the competence model presented, for example, by means of an extended survey for a larger group of participants. Based on the competence model defined, we plan to develop organizational policies for training. With an increased dissemination of GBL provided by the growing of game facilitation competencies, the effects on teaching in higher education should be explored. However, one approach that could replace the training of game facilitators is the digital support or even the replacement of game facilitators by virtual assistants as supported by improvements in artificial intelligence [57].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Facilitation with large groups.
[PDF File (Adobe PDF File), 48 KB - games_v9i2e25481_app1.pdf]

Multimedia Appendix 2
Further results from the questionnaire: educational scenarios.
[PDF File (Adobe PDF File), 42 KB - games_v9i2e25481_app2.pdf]

References

1. Breuer J, Bente G. Why so serious? on the relation of serious games and learning. J. Comput. Game Cult 2010;4:7-24.
2. Michael D, Che R. Serious Games: Games That Educate, Train, and Inform. Mason, OH, USA: Course Technology; 2005.
3. Ratan R, Ritterfeld U. Classifying serious games. In: Cody M, Vorderer P, Ritterfeld U, editors. Serious Games: Mechanisms and Effects. New York: Routledge; 2009:10-22.
4. Kolb D. Experiential Learning: Experience as the Source of Learning and Development. New Jersy: Pearson FT Press; 2014.
5. Urhahne D, Chao SH, Florineth ML, Luttenberger S, Paechter M. Academic self-concept, learning motivation, and test anxiety of the underestimated student. Br J Educ Psychol 2011 Mar;81(Pt 1):161-177. [doi: 10.1348/000709910X504500] [Medline: 21391968]
6. Laurillard D. The pedagogical challenges to collaborative technologies. Computer Supported Learning 2008 Dec 10;4(1):5-20. [doi: 10.1007/s11412-008-9056-2]
7. Abraham RR, Komattil R. Heutagogic approach to developing capable learners. Medical Teacher 2016 Dec 25;39(3):295-299. [doi: 10.1080/0142159x.2017.1270433]
8. González A, Jennings D, Manriquez L. Multi-faceted impact of a team game tournament on the ability of the learners to engaged and develop their own critical skill set. M. Int. J. Eng. Educ 20154;30:1213-1224.

9. Gentile M, Città G, Perna S, Signa A, Dal Grande V, Ottaviano S, et al. The role of disposition to critical thinking in digital game-based learning. IJSG 2019 Sep 20;6(3):51-63. [doi: 10.17083/ijsg.v6i3.316]
10. Steinkuehler C, Duncan S. Scientific habits of mind in virtual worlds. J Sci Educ Technol 2008 Sep 3;17(6):530-543. [doi: 10.1007/s10956-008-9120-8]
11. Bellotti F, Berta R, De Gloria A, Lavagnino E, Antonaci A, Dagnino F, et al. Serious games and the development of an entrepreneurial mindset in higher education engineering students. Entertainment Computing 2014 Dec;5(4):357-366. [doi: 10.1016/j.entcom.2014.07.003]
12. Pechenkina E, Laurence D, Oates G, Eldridge D, Hunter D. Using a gamified mobile app to increase student engagement, retention and academic achievement. Int J Educ Technol High Educ 2017 Aug 4;14(1):1-1. [doi: 10.1186/s41239-017-0069-7]
13. Dörner D. Über die Schwierigkeiten menschlichen Umgangs mit Komplexität. Psychol. Rundschau 1981;XXXI:163-179. [doi: 10.14361/transcript.9783839414354.71]
14. Arnold U, Söbke H, Reichelt M. SimCity in infrastructure management education. Education Sciences 2019 Aug 05;9(3):209. [doi: 10.3390/educsci9030209]
15. Söbke H. A case study of deep gamification in higher engineering education. : Springer; 2018 Presented at: Games and Learning Alliance 7th International Conference; 2018 Dec 5-7; Palermo, Italy.
16. Riemer V, Schrader C. Learning with quizzes, simulations, and adventures: Students' attitudes, perceptions and intentions to learn with different types of serious games. Computers & Education 2015 Oct;88:160-168. [doi: 10.1016/j.compedu.2015.05.003]
17. De Gloria A, Bellotti F, Berta R. Serious Games for education and training. IJSG 2014 Feb 03;1(1):1-1. [doi: 10.17083/ijsg.v1i1.11]
18. Egenfeldt-Nielsen S. Overview of research on the educational use of video games. Nord. J. Digit. Lit 2006;1:184-214 [FREE Full text]
19. Boyle EA, Hainey T, Connolly TM, Gray G, Earp J, Ott M, et al. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Computers & Education 2016 Mar;94:178-192. [doi: 10.1016/j.compedu.2015.11.003]
20. Klabbers H. The Magic Circle: Principles of Gaming & Simulation. Rotterdam: Sense Publishers; 2009.
21. Shaffer DW. Epistemic frames for epistemic games. Computers & Education 2006 Apr;46(3):223-234. [doi: 10.1016/j.compedu.2005.11.003]
22. Lave J. In: Wenger E, editor. Situated Learning: Legitimate Peripheral Participation. UK: Cambridge University Press; 1991.
23. Gioia DA, Mehra A, Weick KE. Sensemaking in organizations. The Academy of Management Review 1996 Oct;21(4):1226. [doi: 10.2307/259169]
24. Chapman S, McPhee P, Proudman B. What is experiential education? Journal of Experiential Education 2016 Jul 21;15(2):16-23. [doi: 10.1177/105382599201500203]
25. Leigh E, Spindler L. Congruent facilitation of simulations and games. Gaming, Simulations Soc. Res. Scope Perspect 2005:189-198. [doi: 10.1007/4-431-26797-2_20]
26. Bodner S. The evolution of job analysis: competency assessment comes of age. Int. J. Career Manag 2012;7:1-11.
27. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
28. Gessler M, Sebe-Opfermann A. Kompetenzmodelle. In: Müller-Vorrüggen M, Radel J, editors. Handbuch Personalentwicklung. Stuttgart, Germany: Schäffer-Poeschel Verlag; 2017:159-184.
29. Mudra P. Personalentwicklung - Integrative Gestaltung betrieblicher Lern- und Veränderungsprozesse. Munich: C.H. Beck; 2004.
30. Krumm S, Mertin I, Dries C. Kompetenzmodelle. Göttingen: Hogrefe Verlag; 2012.
31. Boulter N, Dalziel M, Hill J. Achieving the Perfect Fit. Houston, TX, USA: Gulf Publishing Company; 1998.
32. Duffy F. National framework of professional standards for change leadership in education. Int. J. Educ. Leadersh. Prep 2009:4-49.
33. Stewart J. Gigh-performing (and threshold) competencies for group facilitators. Journal of Change Management 2006 Dec;6(4):417-439. [doi: 10.1080/14697010601087115]
34. Nelson T, McFadzean E. Facilitating problem‐solving groups: facilitator competences. Leadership & Org Development J 1998 Apr;19(2):72-82. [doi: 10.1108/01437739810208647]
35. Baker L, Fraser C. Facilitator core competencies as defined by the International Association of Facilitators. IAF Handb. Gr. Facil 2005:459-471.
36. Hoidn S, Kärkkäinen K. Promoting skills for innovation in higher education: a literature review on the effectiveness of problem-based learning and of teaching behaviours. Pr. OECD Educ. Work. Pap 2014:177-182. [doi: 10.1787/5k3tsj67l226-en]
37. Blašková M, Blaško R, Kucharčíková A. Competences and competence model of university teachers. Procedia - Social and Behavioral Sciences 2014 Dec;159:457-467. [doi: 10.1016/j.sbspro.2014.12.407]

38. Kortmann R, Peters V. Demystifying the unseen helmsman: Towards a competency model for game facilitators. Delft University of Technology. 2017. URL: https://doi.org/10.4233/uuid:d6f58b7e-ac7d-4d72-b32d-67600dce1217 [accessed 2021-04-07]
39. Despeisse M. Games and simulations in industrial engineering education: a review of the cognitive and affective learning outcomes. : IEEE; 2018 Presented at: InWinter Simulation Conference (WSC); 2018 Dec 9-12; Gothenburg, Sweden p. 4046-4057. [doi: 10.1109/wsc.2018.8632285]
40. Grogan PT, Meijer SA. Gaming methods in engineering systems research. Syst Eng 2017 Nov 02;20(6):542-552. [doi: 10.1002/sys.21409]
41. Irvine J. A comparison of revised Bloom and Marzano's new taxonomy of learning. Res. High. Educ. J 2017;33:1-16.
42. Peth-Pierce R. A good beginning: sending America's children to school with the social and emotional competence they need to succeed. The child mental health foundations and agencies network. 2000. URL: https://files.eric.ed.gov/fulltext/ED445810.pdf [accessed 2021-04-07]
43. Valiathan P. Blended learning models. ASTD. 2002. URL: https://purnima-valiathan.com/wp-content/uploads/2015/09/Blended-Learning-Models-2002-ASTD.pdf [accessed 2021-04-07]
44. Söbke H, Hadlich C, Müller N, Hesse T, Henning C, Schneider S, et al. Social Game Fliplife: Digging for talent. 2012 Presented at: The 6th European Conference on Games Based Learning; 2012 Oct 4-5; Cork, Ireland.
45. Baalsrud Hauge J, Bellotti F, Nadolski R, Berta R, Carvalho MB. Deploying serious games for management in higher education: lessons learned and good practices. EAI Endorsed Transactions on Game-Based Learning 2014 Aug 26;1(3):e4. [doi: 10.4108/sg.1.3.e4]
46. Heron J. The Complete Facilitator's Handbook. London: Kogan Page; 1999.
47. Saulnier B, Landry J, Wagner T. From teaching to learning: learner-centered teaching assessment in information systems education. J. Inf. Syst. Educ 2008;19(2):169-174.
48. Cardoso S, Tavares O, Sin C. The quality of teaching staff: higher education institutions' compliance with the European Standards and Guidelines for Quality Assurance—the case of Portugal. Educ Asse Eval Acc 2015 Jan 22;27(3):205-222. [doi: 10.1007/s11092-015-9211-z]
49. DeCarolis A, Macchi M, Negri E, Terzi S. A Maturity Model for Assessing the Digital Readiness of Manufacturing Companies. In: Lödding H, Riedel R, Thoben KD, von Cieminski G, Kiritsis D (eds.) Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing. 2017 Presented at: IFIP WG 5.7 International Conference; 2017 Sept 3-7; Hamburg, Germany p. 13-20.
50. Machado CG, Winroth M, Carlsson D, Almström P, Centerholt V, Hallin M. Industry 4.0 readiness in manufacturing companies: challenges and enablers towards increased digitalization. Procedia CIRP 2019;81:1113-1118. [doi: 10.1016/j.procir.2019.03.262]
51. Zhou Y. Towards Capability Maturity Model of e-Learning Process. IIM 2012;04(04):95-98. [doi: 10.4236/iim.2012.44015]
52. Bowey JT, Friehs MA, Mandryk RL. ed or blue pill: Fostering identification and transportation through dialogue choices in RPGs. 2019 Presented at: The 14th International Conference on the Foundations of Digital Games; 2019 Aug; San Luis Obispo, California, USA.
53. Passmore C, Birk MV, Mandry RL. The privilege of immersion: racial and ethnic experiences, perceptions, and beliefs in digital gaming. 2018 Presented at: The 2018 Conference on Human Factors in Computing Systems; 2018 Apr 21-26; Montreal, Canada. [doi: 10.1145/3173574.3173957]
54. Birk M, Friehs MA, Mandryk RL. Age-based preferences and player experience: a crowdsourced cross-sectional study. 2017 Presented at: The annual Symposium on Computer-Human Interaction in Play; 2017 Oct 15-18; Amsterdam, Netherlands. [doi: 10.1145/3116595.3116608]
55. Alexandrovsky D, Friehs M, Birk MV, Yates RK, Mandryk RL. Game dynamics that support snacking, not feasting. 2019 Presented at: The Annual Symposium on Computer-Human Interaction in Play; 2019 Oct 22-25; Barcelona, Spain. [doi: 10.1145/3311350.3347151]
56. Söbke H, Baalsrud Hauge J, Stefan IA. Prime example ingress reframing the pervasive game design framework (PGDF). IJSG 2017 Jun 13;4(2):39-56. [doi: 10.17083/ijsg.v4i2.182]
57. Luccini A, Montara M, Catalano CE. Deliverable 3.2: SG thematic application fields report. In: Games and Learning Alliance. 2012 Presented at: 12th International Conference on Advanced Learning Technologies; 4-6 July 2012; Rome p. 1-1.

Abbreviations

GBL: game-based learning

Edited by N Zary; submitted 03.11.20; peer-reviewed by M de Wijse-van Heeswijk, E Leigh, M Friehs; comments to author 08.12.20; revised version received 28.01.21; accepted 19.03.21; published 05.05.21.

Please cite as:
Baalsrud Hauge J, Söbke H, Bröker T, Lim T, Luccini AM, Kornevs M, Meijer S
Current Competencies of Game Facilitators and Their Potential Optimization in Higher Education: Multimethod Study
JMIR Serious Games 2021;9(2):e25481
URL: https://games.jmir.org/2021/2/e25481
doi:10.2196/25481
PMID:33949956

Original Paper

Mobile Game Design Guide to Improve Gaming Experience for the Middle-Aged and Older Adult Population: User-Centered Design Approach

Seyeon Lee[1], MSc; Hyunyoung Oh[1], MSc; Chung-Kon Shi[1], PhD; Young Yim Doh[1], PhD

Graduate School of Culture Technology, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea

Corresponding Author:
Young Yim Doh, PhD
Graduate School of Culture Technology
Korea Advanced Institute of Science and Technology
291 Daehak-ro, Yuseong-gu
Daejeon, 34141
Republic of Korea
Phone: 82 42 350 2994
Fax: 82 42 350 2910
Email: yydoh@kaist.ac.kr

Abstract

Background: The number of older adult gamers who play mobile games is growing worldwide. Earlier studies have reported that digital games provide cognitive, physical, and socioemotional benefits for older adults. However, current mobile games that understand older adults' gameplay experience and reflect their needs are very scarce. Furthermore, studies that have analyzed older adults' game experience in a holistic manner are rare.

Objective: The purpose of this study was to suggest mobile game design guidelines for adults older than 50 years from a holistic gaming experience perspective. Adopting a human-centric approach, this study analyzes middle-aged and older adults' gameplay experience and suggests practical design guides to increase accessibility and satisfaction.

Methods: We organized a living laboratory project called the "Intergenerational Play Workshop." In this workshop, 40 middle-aged and older adults (mean age 66.75 years, age range 50-85 years) played commercial mobile games of various genres with young adult partners for 1 month (8 sessions). Using a convergent parallel mixed-method design, we conducted a qualitative analysis of dialogue, game diaries, and behavioral observations during the workshop and a quantitative analysis of the satisfaction level of the game elements for the mobile games that they played.

Results: This project was active from April 2019 to December 2021, and the data were collected at the workshops from July 1 to August 28, 2019. Based on the identified themes of positive and negative experiences from the qualitative data, we proposed 45 design guides under 3 categories: (1) cognitive and physical elements, (2) psychological and socioemotional elements, and (3) consumption contextual elements. Our empirical research could reaffirm the proposals from previous studies and provide new guidelines for improving the game design. In addition, we demonstrate how existing commercial games can be evaluated quantitatively by using the satisfaction level of each game's elements and overall satisfaction level.

Conclusions: The final guidelines were presented to game designers to easily find related information and enhance the overall understanding of the game experience of middle-aged and older adults.

(JMIR Serious Games 2021;9(2):e24449) doi:10.2196/24449

KEYWORDS

mobile games; older adults; middle-aged adults; design guideline; gaming experience

Introduction

Background

The number of mobile game players has increased annually to 1.1 billion in 2017 and 1.5 billion in 2020. It is expected to increase by 50 million to 100 million people every year and to reach 1.8 billion by 2024 [1]. The rise of smartphone penetration [2] could give older adults an opportunity to easily access digital games. According to the American Association of Retired Persons research in 2019, 73% of game players aged 50 years and older use mobile devices in the United States—this proportion has increased from 57% in 2016, which was only a few years earlier. In contrast, the proportion of game players using computers and laptops decreased from 56% in 2016 to 47% in 2019 [3]. The Korean mobile game market for the older adult population is developing actively with high smartphone penetration rates and high Wi-Fi accessibility in public places. In a survey conducted in 2020, 52.1% of Koreans in their 50s and 31.9% of Koreans aged 60 to 65 years mentioned that they played mobile games [4]. Previous studies have reported various benefits of digital gameplay for older adults, including lower cognitive decline [5-7], increased physical activity [8-11], and improved socioemotional status [12-14]. However, most digital games cater to the younger generations, making it difficult for many older adults to play them [15,16]. In addition, game designers may not be sensitive to the gaming experience of older adults because most of them belong to the younger generation. Therefore, it is necessary to improve the game design based on the understanding of older adults' needs, preference, and play experience. As the demand for mobile games is expected to increase among middle-aged and older adults in the future, this study aims to provide experience analysis and design guidelines focused on mobile games.

Design Guide for Age-Related Changes

Earlier studies have reported various factors that cause difficulties for older adults to play mobile games. In a survey of Canadian gamers aged over 55 years, the highest ranked difficulty was that the games were too complicated [17]. In addition, age-related limitations such as declining vision, hearing, cognition, and motor functions can cause difficulties in playing games [18]. Older adults who were skilled gamers did not experience many difficulties, but they complained of difficulty in changing game settings, small object size, time limit, understanding rules, and fast pace [16]. To improve the accessibility of digital technologies such as computers and the internet, many studies have suggested interface design guides for older adults, considering age-related changes such as vision, motor control, hearing, cognition, and so on [19-21]. Several attempts have been made to apply a similar approach to digital game design. Gamberini et al [22] studied age-related changes in terms of perception (vision, hearing, and touch/movement), attention, learning/memory, and cognitive tasks and proposed specific implications for game design. Fua et al [23] also studied the relationship between cognitive abilities and gameplay mechanics for older adults, especially within populations at risk of dementia. In addition, Ijsselsteijn et al [24] provided insights into digital game design based on usability guides for older adults. Gerling et al [25] created structural models of gameplay and proposed guides that consider age-related changes under the categories of game elements, such as players and resources, user interface, core mechanics, and outcomes. However, these studies did not analyze the game experience of older adults through an empirical study. De Barros et al [26] improved their mobile game app by reflecting on older adults' behavior and feedback, and they proposed recommendations under 3 categories, that is, navigation, interaction, and visual design.

Design Guide Related to Psychological and Socioemotional Aspects

Various surveys conducted in the United States, Canada, South Korea, and Flanders showed similar results regarding game preferences of older adults. They indicated that older adults strongly prefer puzzles, card/board games, and strategy games to other genres. These types of games are mostly casual and easy to learn and play, but are also challenging [4,12,27,28]. Other studies also showed that older adults prefer an intellectual challenge to fighting and racing games, which require skills related to speed and reflexes [29,30]. Another experimental study identified how the enjoyment of older adults differs depending on the genre of the games. It was found that puzzle games were easier to play and more enjoyable than simulation games or action games [31]. In addition, older adults' preference for the game genre did not differ between gamers and nongamers [32]. Among the 6 motives of game play presented from the uses and gratifications perspective [33], challenge was the highest among older adults, followed by arousal, diversion, fantasy, competition, and social interaction [27]. In addition, older adults played digital games for relaxation, fun, and to pass time [16]. For older adults, the emphasis was not only on fun associated with games but also on personal growth and benefits as motivations to play [34]. In fact, many older adults described playing games as a means of exercising their brain [35].

De Carvalho and Ishitani [36] conducted an empirical study to test existing mobile puzzle games in older adults and proposed guidelines related to motivational aspects for serious games. Cota et al [37] also found the main features that motivate the older adults by analyzing their evaluations of mobile games developed by the researchers. Both studies proposed design guides to encourage motivation, such as usability heuristics, clear and motivational feedback, clarifying the benefits of gaming, cognitive training, and providing proper difficulty levels.

Older adults play games mostly alone rather than with others [27,28]. However, baby boomers are community-oriented, and social interaction is an important motivation for them [30]. Moreover, experimental studies have shown that playing with a partner leads to a higher enjoyment than playing with a computer [38,39]. In particular, middle-aged adults (40-59 years) are more likely to play games with their children [28]. In the United States, 57% of parents play games with their children at least once a week [40]. De Schutter and Vanden Abeele [41] recommended a game design that considered the psychosocial context of older adult players, which include finding the right playing partner, vicarious play, time management, sharing high scores, and balancing teams. Additionally, McLaughlin et al [42] conducted a qualitative study on the experience of older adults in terms of costs (eg, initial frustration, stereotype threat, emotional arousal, usability challenges) and benefits (eg, self-esteem, physical activity, social interaction, well-being, fun, learning). Marston [43] also analyzed older adults' game experience empirically under a framework of 3 categories (older adults, technology, and content/interaction) and applied them to suggest design guidelines.

Research Questions in This Study

While many user experience studies have been conducted on digital apps for older adults, not many have considered "playability" for older adults on "mobile game" media. Furthermore, studies that have analyzed older adults' gaming experience in a holistic manner are rare. Some studies have investigated the elements of age-related changes and suggested guides [23-25]. However, they were not empirical studies. Some empirical studies have analyzed the experience of older adults; however, they lacked suggestions for various game elements under the structural framework [36,37,42]. In addition, previous guides did not include consumer behavior contexts related to mobile gameplay, such as devices, installation, in-game payment, and advertising. Moreover, some empirical studies suggested guides under the structural framework, but they tested only limited devices or genres [43]. Therefore, the purpose of this study was to suggest practical design guides based on empirical studies for game designers, including various mobile game elements, from the holistic perspective of middle-aged and older adults' gaming experience. Game elements can be considered as a set of building blocks or features shared by games [44]. We used "game elements" in a broad interpretation to connect user-perceived game elements to game design elements. This study was guided by the following 3 research questions (RQs):

RQ1. What are the positive and negative experiences of middle-aged and older adults when playing mobile games?

RQ2. What mobile game design guides should be considered to embrace middle-aged and older adults?

RQ3. How satisfied are middle-aged and older adults with the game elements of commercial mobile games?

We planned to identify several themes of the gaming experience, design guidelines from qualitative analyses (RQ1 and RQ2), and provide useful indicators by presenting the satisfaction levels of the game elements in existing commercial games by using quantitative analysis (RQ3). This convergent parallel mixed-methods design [45] was used to identify important criteria that could affect the gaming experience of middle-aged and older adults and to determine how existing games can be evaluated based on these criteria.

Methods

Participants

We adopted a human-centered approach by using a living laboratory that promotes active involvement of users and applies their experience and feedback to the research [46]. We collaborated with 2 institutions that organize programs for middle-aged and older adults. One institution is the 50 plus campus that provides educational programs targeting adults in their 50s and 60s (middle-aged adult group) and the other institution is a senior welfare center targeting adults aged 65 or older (older adult group). We designed living laboratory activities where middle-aged and older adults played various genres of mobile games with young adult partners. During these activities, participants not only played games but also discussed their playing experiences and provided recommendations to improve game design. A total of 40 middle-aged and older adults participated, and Table 1 shows the demographic information of the participants.

Table 1. Demographic information of the participants (N=40).

Characteristics	Values
Age (years), mean (SD)	66.75 (9.31)
Age range (years), n (%)	
50-59	11 (28)
60-69	14 (35)
70-79	12 (30)
80-85	3 (8)
Location, n (%)	
50 plus campus (middle-aged adult group)	20 (50)
Senior welfare center (older adult group)	20 (50)
Gender, n (%)	
Female	30 (75)
Male	10 (25)
Job status, n (%)	
Employed	12 (30)
Retired/homemaker	28 (70)
Education, n (%)	
Elementary/middle school	13 (33)
High school	7 (18)
Undergraduate school	12 (30)
Graduate school	8 (20)
Monthly income[a], n (%)	
Low: Less than ₩2 million	18 (45)
Middle: ₩2-8 million	16 (40)
High: More than ₩8 million	5 (13)
No response	1 (3)

[a]US $1=₩1200.

The recruitment process was conducted differently in both the places. Web-based recruitment was conducted at the 50 plus campus for a month, and 20 participants were recruited. Our orientation session explained the workshop activities and received informed consent forms for participation. Two participants did not attend the orientation, and one decided not to participate after the orientation. Therefore, we recruited 3 more participants. In the senior welfare center, offline recruitment was conducted for 1 month. Two social workers familiar with the members of the center participated in the recruitment and selected 20 mentally and physically healthy participants. We also held an orientation session and received informed consent forms. The recruitment results showed a high percentage of female participants. It can be assumed that this was because female members attending these 2 institutes tended to prefer new and coactivity programs over male members.

Ten participants were grouped in 1 class, and there were 4 classes in the workshops (2 middle-aged groups and 2 older adult groups). Over 8 sessions, the number of absent days per participant was 0-4 times (average attendance rate, 91.6% [293/320]). However, no participant dropped out during the workshops in both places. Ethical approval was obtained from the institutional review board of the authors' institute. Table 2 shows the prior gaming experience of the participants. There were 13 older adult participants who reported as currently playing games because they had participated in an earlier program that included serious games for cognitive training at the senior welfare center. The age of the first gaming experience was widespread among the middle-aged adult group, and most of the older adult participants started playing games at an older age.

Table 2. Prior gaming experience of the participants.

Characteristic	Values, n (%)
Game experience (n=40)	
No gaming experience	12 (30)
Currently gaming	15 (38)
Have experience but no longer play	13 (33)
Age of first playing games (years)[a], (n=28)	
10-19	3 (11)
20-29	3 (11)
30-39	5 (18)
40-49	2 (7)
50-59	6 (21)
More than 60	9 (32)

[a]Only for participants with gaming experience.

In addition, 11 undergraduate students (9 females and 2 males, mean age 21.45 years) majoring in the Department of the Senior Industry at Kangnam University participated in the living laboratory activities as supporting partners. In each class, 1 young partner was paired with 1 older participant. Young partners participated in all 4 classes and supported a total of 4 older participants per person. The other young partners assisted in the workshop activities. We also held an orientation session for young adult partners to receive their informed consent for participation and provide instructions for the role of the supporting partners. They helped older participants play games, played games together, and received feedback from them through questions.

Games Played in the Workshops

Four graduate students studying game design and human-computer interaction developed the first list of games, which are popular and enjoyed across all age groups. They referred to the annual report of the popular game list and personal experience of playing with older adults. Each student added 20-60 game titles in various genres and platforms and scored the difficulty level of gameplay from 1 to 4 according to the time required to learn them (1: 10 minutes, 2: 1-2 hours, 3: a week, 4: more than a week). A total of 146 games were included in the list, and 3 authors finally selected 9 games to play during the workshops. Table 3 shows the games played in the workshops.

Table 3. Games played in the workshops.

Title	Genre	Game description
Game played by both groups (age range, 50-85 years)		
Homescapes	Puzzle and story	A match-3 puzzle where the player solves puzzles presented with stories and decorates a house [47].
Fruit Ninja	Action	A fruit-slicing action game where the player swipes and slices fruits with a blade [48].
Long Journey of Life	Adventure and story	A story-based game where the player controls a small boat and sails through life stages from birth to death [49].
Games played by the middle-aged group only (age range, 50-64 years)		
Brawl Stars	Action	A multiplayer web-based battle arena game where players attack other players with the same team members [50].
Sally's Law	Action and story	A platformer action game with a story and puzzle [51].
Good Pizza, Great Pizza	Management	The player becomes the owner of a pizza restaurant and makes pizzas according to orders [52].
Games played by the older adult group only (age range, 65-85 years)		
Word Tower-World Trip	Puzzle	A word puzzle game where the player swipes and connects syllables and finds hidden words [53].
Go-stop Plus	Web board	A web-based version of a Korean traditional card game that uses a deck of 48 cards [54].
Lonely One: Hole-in-one	Shooting and sports	A golf-based game where the player adjusts the parabola and releases to hit the ball [55].

The following guidelines were used to choose the games. First, the game should be accessible in both mobile phones and tablet personal computers; 72 games were selected in this stage. Second, the game should be in Korean language; 64 games were further selected in this stage. Third, to ensure that the games were not too complicated for older adults, the difficulty level should be 1 or 2. In this stage, 54 games were selected from the remaining ones. There were 21 puzzle games, 6 action games, 6 web board games, 4 adventure games, 10 management/simulation games, 6 shooting games, and 1 rhythm game. Before we finalized the list, we had discussions with a social worker and a program manager in each collaborating institution regarding participant groups' digital skills [56] in using mobile apps. In addition, we had a pilot session to test the games with 13 participants from the 50 plus campus. Finally, we selected 1-3 games from each genre in order to have an even distribution from a variety of genres such as puzzle, action, web board, adventure, and management games. Considering the digital skills of each participant group, we selected 3 common games for both groups and 3 different games for each group. "Go-stop Plus" and "Lonely One: Hole-in-one" were selected for the older adults' group because they are based on card games and sports, which are familiar to older adults in real life. Participants also played other games not in this list according to their preferences during the living laboratory activities and provided their feedbacks.

Workshop Procedure

We organized intergenerational play workshops twice a week for 4 weeks (8 sessions), and each session lasted 1-2 hours. Each session consisted of a short lecture (10-30 min), coplaying (20-40 min), survey and pair discussion (10-20 min), and group discussion (10-20 min). Figure 1 shows the workshop process.

The lectures dealt with how to play specific games and useful information to enhance the game literacy of those who are not familiar with digital games. After a short lecture, participants and young adult partners were paired, and they played games together. Each pair was provided a 10-inch tablet personal computer, but they could use their personal smartphone if they wanted. Young adult partners not only played games together but also taught participants how to play and get acquainted with the games. After playing a game, the young partner asked, "what were you satisfied with or unsatisfied with during the game play?", to which the participants provided their responses. In addition, participants filled out the survey form by checking the satisfaction level for each element of the game. They discussed their gameplay experience and shared ideas of design recommendation during group discussions. Participants were compensated with ₩10,000 (approximately US $8) per session for participating in the workshops.

Figure 1. Photographs from the intergenerational play workshop.

For over one month of the living laboratory activities, older adult participants could visit the play room at the senior welfare center and use the tablet personal computers, which had the recommended games installed. There was no playroom at the 50 plus campus, but middle-aged adult participants could borrow the tablet personal computers and play games at home. Both groups could also install and play games on their personal mobile phones. Researchers opened a web-based community and asked participants to upload diaries with their experience after playing games at home. This task of writing a diary was only assigned to the middle-aged adult group because the older adult group was not familiar with a web-based community.

Data Collection and Analysis

Qualitative Analysis (RQ1 and RQ2)

For qualitative data entries, conversations during coplaying and discussion sessions were recorded and transcribed, including gaming diaries from middle-aged participants. The purpose of the qualitative analysis was to develop a practical mobile design guide for middle-aged and older adults in a holistic manner. The “customer journey map” suggested by service design studies [57] was utilized to include as many game elements as possible, offering contact points between users and gameplay in a coordinated series of processes from the start of the gameplay to the end.

First, we performed inductive thematic analysis on conversation transcripts and gameplay diaries to identify the positive and negative experiences of gameplay. We followed the 6 steps of thematic analysis proposed by Braun and Clarke [58] in this process. Thematic analysis is a useful method for summarizing the key features of a large data set and for gaining a rich understanding of people’s experiences [59]. To familiarize with the data, the first author read all the transcripts and diaries multiple times (step 1). Then, the first author generated initial codes (step 2) and developed these codes into themes of positive and negative experiences (step 3). Three authors reviewed and defined themes through discussion to reach a consensus (steps 4 and 5). All steps were repeated until a final version of the game design guides was acquired (step 6).

Second, we grouped themes of positive and negative experiences under the categories of elements related to gameplay. In order to define these categories, we first identified the basic elements (eg, audiovisuals, interface) by referring to the fundamental components presented by Ermi and Mäyrä [60]. Then, we added additional elements such as device, platforms, and sales, referring to the contents of a design document presented by Fullerton [61]. Finally, 10 elements were included: (1) cognitive and physical elements that are mainly addressed in existing user experience studies, such as audiovisual, interface, motor skill, and touch interaction; (2) psychological and socioemotional elements related to game content such as game rules, story and character, and social aspects; and (3) consumption contextual elements that are related to the game, such as devices, software installation, advertising, and in-game payments. We included consumption contextual elements in the guide because they also form an integrated user experience with core game elements.

Third, we generated design guides from each theme of experience and recommendation from the participants’ feedback. Three authors and one game design expert reviewed the themes and created proper guides. The final guides were compared with previous works to identify the guides that matched the previous works and the newly discovered guides.

Quantitative Analysis (RQ3)

For quantitative data, we requested survey responses regarding the satisfaction level with each element of the games played at the end of each session. The data were collected because the scores of the commercial games on each element can be a useful indicator for game designers. First, participants rated the satisfaction level of each game's elements on a 5-point scale for the game they played that day (1=very unsatisfied; 2=unsatisfied; 3=average; 4=satisfied; and 5=very satisfied). Table 4 shows the questionnaire regarding the satisfaction level with each game element. Five face emoticons were used for the 5-point scale to facilitate the answering process for participants ().

Second, on the last day of the workshop, participants rated the overall satisfaction level of the game experience for each game on a 5-point scale. This allowed us to examine the satisfaction level when participants first encountered each game and the overall satisfaction after playing the game for a while.

Table 4. Survey questionnaires for measuring the satisfaction level.

Game element	Question
Cognitive and physical elements	
Font size	Was the text legible with a reasonable font size?
Button size	Was the size and location of the button appropriate?
Sound	Were you satisfied with the sound of the game?
Information amount	Was there an appropriate amount of on-screen information?
Button interaction	Did you have any difficulty finding which button to press on each page?
Speed	Was the game played at a proper speed to understand the content?
Agility	Was it difficult when you needed a quick reaction? (eg, selecting moving targets, avoiding obstacles quickly)
Psychological and socioemotional elements	
Objective	Was the goal to achieve in the game clear and appropriate?
Resource/item	Was the method of using in-game resources and items clear and appropriate?
Word comprehension	Did you understand the meaning of the words and sentences in the game?
Mood/character/story	Did you like the atmosphere, character, and story of the game?
Consumption contextual elements	
Setting to start	How was the setting process to start the game? (eg, loading the account, setting a nickname)
Advertising	Were the in-game ads easy to deal with?
Payment	Did you understand the in-game purchase system?

Results

Positive and Negative Experience With Gameplay (RQ1)

We identified 38 themes of positive experiences and 58 themes of negative experiences from dialogs, game diaries, and observations. The most frequently mentioned negative experience was difficulty in understanding the game rules. The second was frustration caused by repeated failures in action games because of lack of agility. The third was difficulty in distinguishing between important objects of similar color or design. Interestingly, the pleasure of playing together was found to be the most positive experience, followed by cognitive training. The third was sensational pleasure derived from special audiovisual effects such as all objects bursting at once. Based on the identified themes of positive and negative experiences, we proposed 45 mobile game design guides for the middle-aged and older adult population. For detailed lists of themes and guides, see the supplementary material (Multimedia Appendix 1).

Mobile Game Design Guides for the Middle-Aged and Older Adult Population (RQ2)

Design Guides Related to Cognitive and Physical Elements

Table 5 shows the design guidelines related to cognitive and physical elements. We presented not only the proposed guides but also previous works that are consistent with our findings.

Table 5. Mobile game design guides on cognitive and physical elements.

Design guide (DG)		Previous works
Audiovisual		
DG1	Design important object/character as distinguishable from others.	[20,22,36,62]
DG2	Provide voice dubbing when presenting stories or speech bubble.	[24]
DG3	Provide options to choose the size of the font and objects.	[20,22,24,25,36]
DG4	Avoid sounds that are too sharp or repetitive.	—[a]
Interface		
DG5	Present functions step-by-step rather than presenting excessive information in one screen.	[20,23]
DG6	Visually express functions of buttons for illiterate players.	[20,23]
DG7	Avoid multiple button controls at the same time unless they are essential to gameplay.	[20,62]
DG8	Provide user manual that explains the control and function of buttons.	[20]
DG9	Place important buttons in easy to find and touch locations.	[26]
DG10	Provide a mini-map.	—
DG11	Highlight touch area rather than suggesting to touch anywhere.	—
DG12	Provide a tutorial in case of repeated incorrect touch interactions.	—
DG13	Automatically confirm the termination of sliding actions after a specific time period.	—
DG14	Limit the area of operating touchpad on the screen.	—
Motor skills		
DG15	Provide practice session for beginners.	—
DG16	Provide hints for control timing when the player fails repeatedly.	—
DG17	Increase the process speed and difficulty incrementally.	[25]
DG18	Provide speed adjustment function.	[25]

[a]Not available.

Several previous studies have reported design recommendations that consider age-related changes such as font size, interface, and motor skills in user interface design, including digital games [20,22-26,36,62]. We found similar experiences, especially in audiovisuals and interfaces. Negative experiences related to declining vision were most frequently identified. In addition, participants had negative experiences in complex interface designs that were cognitively burdensome.

Touch interaction in games is unique for older adult participants who lack gameplay experience and who are not familiar with sliding or swapping interactions. We developed design guides 11-14 based on the failure cases of touch interaction. Most failure cases were observed in older adult participants. We found that participants slid objects in the wrong direction or forgot to pull a finger away from the screen after the action. In addition, some participants mistakenly performed a tap action when a slide was required. Several participants failed to double-click because they tapped once or too slowly. Confusion among participants was aggravated when complex touch-interaction methods were required. The older adult group played "Lonely One: Hole-in-one," in which a parabola appears when the player touches anywhere on the screen. The player should adjust the position of the parabola and release it, causing the object to move along the parabola. This interaction method is similar to the popular game "Angry Birds." Participants often failed to draw a correct parabola because they touched the wrong starting point (too close to the edge of the screen). Moreover, they tried to draw a line with their finger along the parabolic dotted line. In addition, we found that participants felt uncomfortable when the speed of the moving character was too fast or too slow. We developed 4 guides regarding motor skills (design guides 15-18).

Design Guides Related to Psychological and Socioemotional Elements

Table 6 shows the design guidelines related to psychological and socioemotional elements. Similar to that reported in previous research [17,25,36,43], our participants easily adapted to a game that had clear goals and simple rules, but they became frustrated with a game that was difficult and insufficiently explained. They said they forgot the rules quickly over time, although they had learned how to play it. Therefore, we developed design guide 19 that offers sufficient explanation of the game rules [25] and design guide 20 that provides options for selecting the difficulty levels [25,36,43]. Furthermore, new design guides were developed from participants' feedback, such as providing hints (design guide 21), tips during loading time (design guide 22), and notification messages to prevent resource abuse (design guide 23). In addition, many participants disliked the time limitation in a game; they felt nervous and stressed with the time limits. Therefore, we proposed design guide 25, "eliminate time limits or provide alternatives for level-passing," which was also proposed in previous studies [20,36]. Some middle-aged participants were satisfied with restrictions on the play opportunities, such as limiting the number of hearts because it prevented them from becoming too addicted to games. We developed a design guide for the time management option, which was also proposed by De Schutter and Vanden Abeele [41].

In terms of story and character, participants expressed dissatisfaction with stories that were too simple or cliché. They were also dissatisfied with stories that lack sympathy or promoted negative emotions such as depression and gloom. Participants also criticized stereotypical expressions regarding the lives of older adults. Instead, participants liked game stories that appealed to and were empathetic to players, stimulated their curiosity, and provided useful knowledge. One participant felt satisfied when another character in the game praised the player. Similar guides and recommendations were reported from previous works [10,36,43]; however, a new guide was proposed regarding a variety of interactive storytelling options such that the player can reflect their own preference (design guide 30). We found many positive effects of gameplay from participants' feedback and identified those experiences to themes of positive experience under the category of affective aspects: feeling of fun and flow, cognitive skill training, sense of achievement, strategic thinking, concentration, learning something new, killing time, connecting stories to real life experiences and finding meaning, and reminiscing.

Similar to previous studies, we found that many older adult participants perceived gameplay as cognitive training activities [35]. For example, they mentioned "It was good because I feel like [I am] training my brain," "It is likely to prevent dementia," and "I feel like I'm getting smarter because I keep using my brain." Moreover, participants said that the perception of beneficial effects is an important motivation to play. De Carvalho et al [36] also highlighted that clarifying the benefits of the game for older adults is important. We added design guide 31, which informs the affective aspect and beneficial effects of the game such as cognitive training, strategic thinking, learning, connecting real life, and reminiscing. We found that the results were consistent with previous studies on social dimensions [24,41,43]; participants said they had more fun playing with others than when playing alone. They were especially satisfied that they were able to accomplish a difficult mission with supporters on their side. They also felt good when they worked together with an unknown player whom they met in the game to complete the mission. However, some middle-aged participants found it harder to play games with their children because they had different preferences and gameplay skills.

Table 6. Mobile game design guides on psychological and socioemotional elements.

Design guide (DG)		Previous works
Game rules		
DG19	Provide guidance, tutorials, and practice sessions.	[25]
DG20	Provide options for selecting difficulty levels.	[25,36,43]
DG21	Provide hints when the player fails repeatedly or is taking too much time.	—[a]
DG22	Provide game rules and tips during loading time (repetitive tutorials).	—
DG23	Provide a notification message if players abuse resources early in the game.	—
DG24	Provide an appropriate challenge rather than a simple or easy rule.	[36,42]
DG25	Eliminate time limits or provide alternatives for level-passing.	[20,36]
DG26	Provide time management options.	[41]
Story and character		
DG27	Provide familiar languages and concepts to the player (eg, based on culture and age).	[20,26]
DG28	Provide players with complimentary messages or motivational feedback.	[10,36,43]
DG29	Do not indicate life of older adults in static, passive, negative, and depressed tones.	[63]
DG30	Provide options for players to choose stories and reflect diversity.	—
Affective aspect and perceived benefit		
DG31	Inform players of the affective aspect and beneficial effects of the game (eg, cognitive training, strategic thinking, learning, connecting real life, reminiscing)	[36]
Social aspect		
DG32	Consider using multiplayer mode or coplaying context in single play mode.	[24,41,43]
DG33	Remove chat features in competitive games or only allow consensual chat between players.	—
DG34	Restrict the use of abusive language.	—

[a]Not available.

One middle-aged participant was rather satisfied with the lack of chat function when playing “Brawl Stars” because she never heard swear words or slang from other players. Alternatively, it was less burdensome if the chat function was set to be available after establishing some acquaintances. However, participants were uncomfortable when the method for connecting to social media or inviting friends was complicated. Some participants were concerned about other people; they were worried about bothering others while turning the sound on or exhibiting excessive action during gameplay. One participant said that she did not want to connect to social media and show others that she played games because there were many formal relationships on social media.

Design Guides Related to Consumption Contextual Elements

Table 7 shows the design guidelines related to consumption contextual elements, such as devices, software installation, and advertising/payment. Most guides under this category are newly written. We found that older adults preferred using tablet personal computers rather than mobile phones because of their deteriorating eyesight resulting from aging. However, using tablet personal computers was not comfortable when participants unintentionally touched the screen. Furthermore, because of the apparent lack of moisture on their fingers, touch recognition did not work well for some of the participants. We provided additional equipment such as touch pens and controllers; touch pens worked well with games where small objects were to be touched. Similarly, middle-aged adults preferred using a specially developed controller while playing “Brawl Stars.” Therefore, encouraging the use of appropriate devices and additional equipment for middle-aged and older adults may be useful. Installing game software on devices is a challenge for some older adult participants who lack experience in using smartphones. They had difficulty typing and searching for apps. However, they became accustomed to it through repeated practice. Participants said they felt confused with different games having similar titles. Many older adult participants could not remember their IDs and password. Without help, it was difficult for them to register for or log into a game. This issue has been reflected in design guide 38, but special care is needed to prevent misuse of personal information. In contrast, middle-aged adults did not have much trouble with their installation. However, waiting for large files to download was boring for some participants and they did not want to pay for data usage to download games.

Participants felt uncomfortable and bored when there were too many ads or when the ads were too long. In addition, they felt uncomfortable when they are suddenly required to watch an ad while playing because it interrupts their game flow. They were upset that advertisements with inappropriate content such as adult content suddenly appear sometimes without any notification. Participants sometimes became confused whether the pop-up interaction was an advertisement or the game they were playing. They were also confused because they did not know how to turn the ads off, especially when the clickable “close” button feature was too small or when there was a countdown before the appearance of the close button. Participants hoped to be provided with items, hints, and other benefits of gameplay when watching advertisements. Some older adult participants were very generous regarding advertisements. They said that it was okay because they could take a short break and were provided new information during the advertisement.

Regarding payment, older adult participants had no intention of purchasing games or items. They did not connect their credit card to the app store because they were afraid of being mischarged. However, some middle-aged adults paid to buy items to gain an advantage in the game. Both groups were offended when phrases inducing such charges appeared at a difficult level.

Table 7. Mobile game design guides on consumption contextual elements.

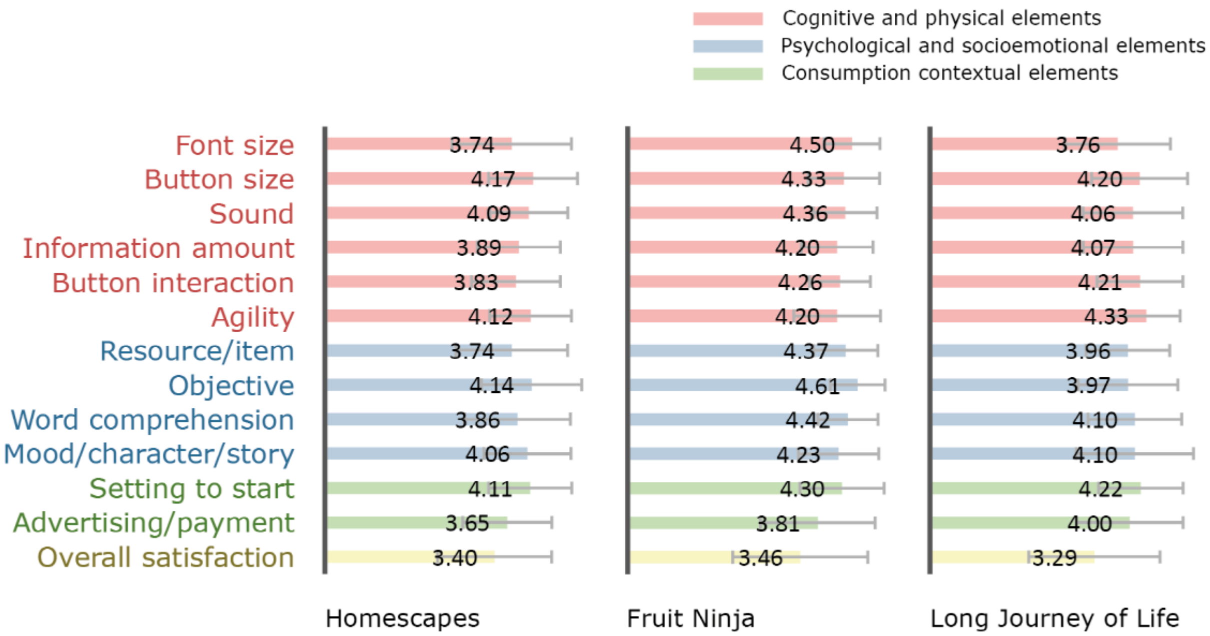
Design guide (DG)	Description
Devices	
DG35	Able to play both in a tablet personal computer and mobile phone.
DG36	Provide supportive equipment (eg, touch pen, controller).
Installation and setting to start	
DG37	Inform expected install time and file sizes.
DG38	Load account information automatically with privacy precautions.
DG39	Use easy and unique game titles, which do not overlap with other games.
Advertisement and payment	
DG40	Provide items and hints after viewing ads.
DG41	Avoid excessive or long ads, which interrupt gaming.
DG42	Make it easy to turn off the ads.
DG43	Present suitable ads for ages and preference (avoid inappropriate content).
DG44	Let the player choose the timing for advertisements (avoid sudden ads that interrupt the game flow).
DG45	Noninteractive ads are preferred over interactive ads.

Satisfaction Level in Each Element of the Games (RQ3)

Figures 2-4 show participants' satisfaction level for each game's elements and overall satisfaction on the 5-point scale. The goal of this descriptive statistics is to provide game designers with a visual profile of how players are satisfied with game elements in the game case. The satisfaction level for each game's elements was measured in each session; therefore, they were treated as first impressions after the first 30 minutes of gameplay. The overall satisfaction level in each game was measured on the last day of the workshops. This score is a satisfaction evaluation by comparing it to other games after several days of gameplay. Overall, we can observe from the figures that the satisfaction level with the individual elements performed on each session tended to be higher than the overall satisfaction evaluated on the last day of the workshop. For additional information on Figures 2-4 such as the mean and standard deviation (SD) of the satisfaction levels in each game, see the supplementary material (Multimedia Appendix 2).

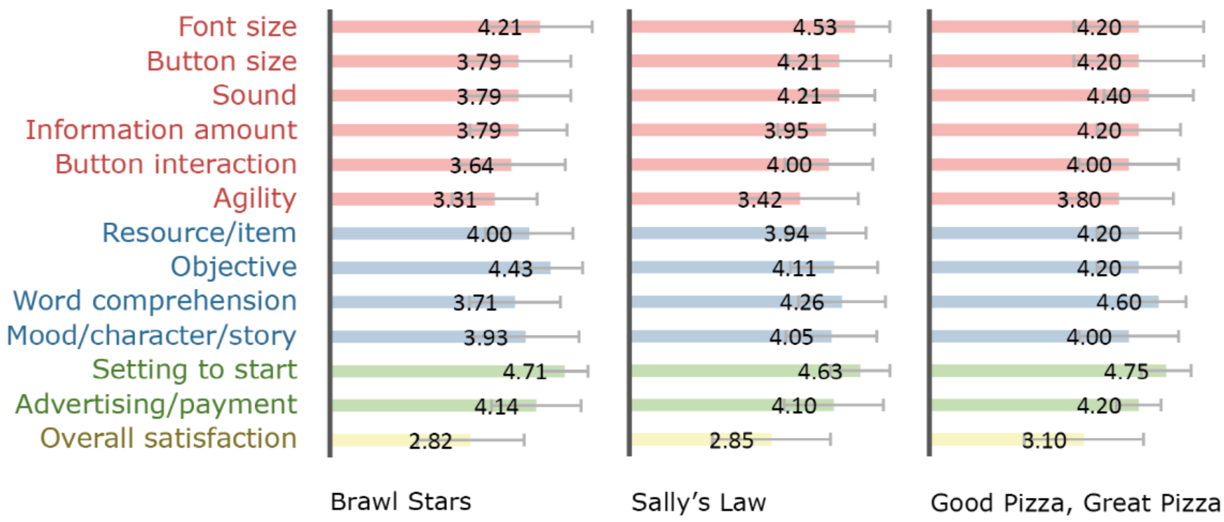
Regarding games played by both groups (Figure 2), the mean scores of "Fruit Ninja" in all game elements were over 4.0, except for advertising/payment, but the overall satisfaction was only 3.46 (SD 1.35). It is assumed that even though this game was intuitive and easy to use, there were wide differences in the overall preference between the participants. The scores of "font size" were a bit lower on "Homescapes" and "Long Journey of Life" than those of other elements; we found that many participants mentioned that the text is too small in those games.

Figure 2. Satisfaction levels in games played by both groups.



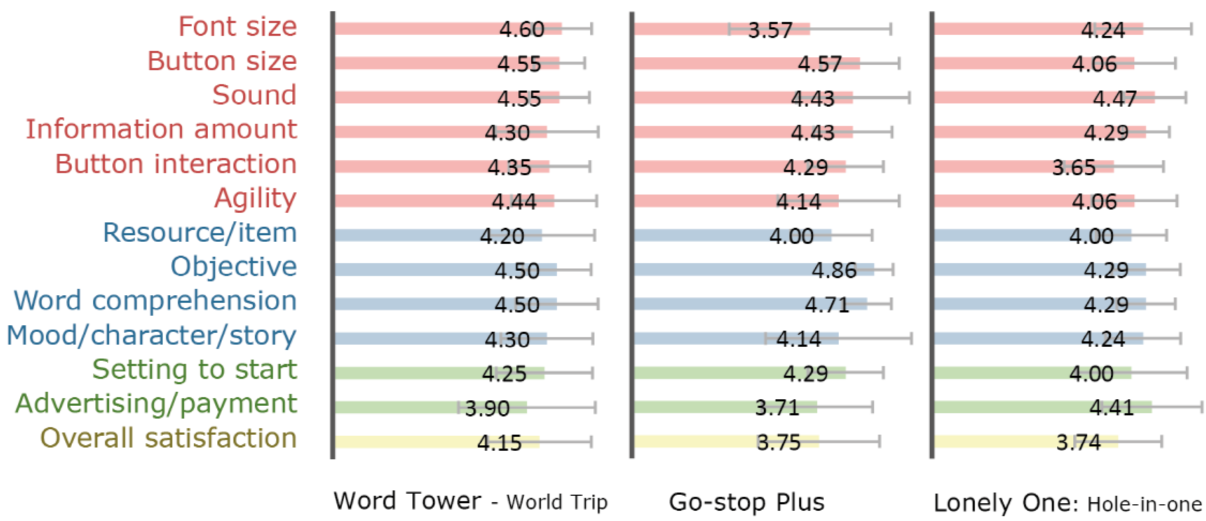
Regarding games played by the middle-aged group only (Figure 3), action games such as "Brawl Stars" and "Sally's law" were found to be less satisfactory with agility than other elements, and the overall satisfaction was low. The 2 games seem to provide participants with too much difficulty in terms of agility. In addition, the button interaction of "Brawl Stars" was also low (mean 3.64 [SD 1.08]), which seems to have been difficult for participants to control buttons simultaneously using both hands.

Figure 3. Satisfaction levels in games played by the middle-aged group.



Regarding games played by the older adult group only (Figure 4), the mean scores of "Word Tower-World Trip" were estimated to be over 4.0 in almost all elements, with a high overall satisfaction level (mean 4.15 [SD 1.04]). Scores of "objective" and "word comprehension" in "Go-stop Plus" were high because it was a digital version of the traditional Korean card game popular with older adults. However, font size was less satisfactory than other elements in this game. Regarding "button interaction," "Lonely One: Hole-in-one" scored only 3.65 (SD 1.00), apparently because adjusting the position of the parabola to determine the direction of the ball's flight was unfamiliar to older adults.

Figure 4. Satisfaction levels in games played by the older adult group.



Timeline of the Study

This project was funded from April 2019 to December 2021 and approved by the institutional review board on June 7, 2019. The data were collected at the workshops involving 40 middle-aged and older adults and 11 young partners from July 1, 2019 to August 28, 2019.

Discussion

Overview of This Study

The purpose of this study was to derive mobile game design guidelines for the middle-aged and older adult population by analyzing user experience through a qualitative thematic analysis. We proposed 45 guides under the framework of 10 game elements. In addition, we reported descriptive statistics showing the satisfaction profiles for 9 game cases. Based on these results, we had the following discussion.

Embrace the Middle-Aged and Older Population: Reaffirmed Guides and Newly Discovered Guides

Previous studies reported that aging issues such as decreased eyesight, memory, and motor skills could lead to difficulties in playing digital games targeted at the younger generation [15,16,18]. Game design guides for older adults should focus on age-related changes and game elements [22-25]. This study reaffirmed these previous issues and noted that these guidelines have not been applied to many commercial mobile games. We rediscovered repeatedly reported issues such as the size of the text and objects, cognitive burden at the interface, and agility problems. These confirmed guidelines should be essential considerations for game designers who want to create games that cater to the middle-aged and older adult population. In previous studies, it was rare to observe all contact points from the beginning to the end of the mobile gaming experience. By including the context of game consumption, this study was able to introduce new guidelines. These are the parts related to the unique touch interaction used in the game and the parts related to the use of the app store and advertising/payment features embedded in the gaming experience.

Prior Experience of Computer and Digital Games

The use of new technologies is more likely to be a response to historical changes in the older generation than age-related declines [64]. We also found that the prior experience in using digital devices and games affects the participants' overall gameplay. This is because many digital games require basic digital skills such as installation, typing, double-clicking, and cancellation. While middle-aged participants who are already experienced in the use of computers and smartphones had no problems with installation, typing, and various touch-based interactions, older adult participants needed more help to familiarize themselves with basic operations. In addition, many games use the design elements of earlier successful games. For example, gamers are familiar with a heart-shaped icon that implies remaining opportunities. Moreover, digital games of the same genre often use similar interaction methods. Therefore, compared to the younger generation who are unconsciously trained on these norms through various games since childhood, older adults, who lack gameplay experience, need time to adapt to a new game. Because of these factors, participants recognized gameplay as a learning activity rather than a leisure activity. Charlier et al [65] stressed that older adults' experiences may include more traditional learning and less game-based learning. Our participants wanted to read specific manuals or learn from others before they started playing. In addition, participants mentioned that they had to recognize the benefits of a game and how it helped them start the gameplay. Therefore, when encouraging older adults to play games, it will be necessary to persuade them to show them how these games are beneficial to them [34,36,37]. In addition, it is recommended to provide game literacy education programs, which introduce the benefits of games, recommend proper games for them, and explain how to play games.

Game Aesthetics and Socioemotional Aspects for the Middle-Aged and Older Adult Population

According to the mechanics, dynamics, and aesthetics framework proposed by Hunicke et al [66], aesthetics refers to the emotions that the player experiences while playing the game. The game designer wants to arouse certain emotional responses in the player, such as sensation, fantasy, narrative, challenge, fellowship, discovery, expression, and submission. De Schutter [67] extended the list to include older adults' gaming experience and presented 6 items as "Geronto-Aesthetics," such as cultivation, compensation, connectedness, contribution, nostalgia, and contemporaneity. Through our empirical research, we discovered themes consistent with those in previous studies: challenge (sense of achievement, strategic thinking, and concentration), sensation (feeling of fun and flow), discovery (learning something new), submission (killing time), cultivation (connecting stories to real life experiences and finding meaning), and nostalgia (reminiscing). Moreover, we identified one more item, that is, cognitive skill training. Gameplay is a meaningful activity for older adults because they are interested in cognitive training and in mitigating dementia [35].

Regarding the social aspect of the gaming experience, participants enjoyed playing games together [38,39] and they positively evaluated the interactions with the younger generation. In this study, playing together does not necessarily mean solely playing multiplayer games. Various social interactions could also occur during games that offer only single-player options, for example, turn-taking, score competition, helping with difficult parts, and vicarious play. Young adult partners acted as supporters and coplayers during the workshop; without this help, it would have been even more difficult for the older population to use the game alone. In addition, participants preferred cultural content that they were familiar with, and many women were reluctant to engage with violent content [16].

Limitations and Future Works

There are several limitations associated with this study and a number of issues that remain unexplored. First, a limited number of games were played because of the short workshop duration. Users' personal preference and styles are important factors, but we did not reflect them when we chose the games. Depending on the diverse interests of middle-aged and older adults, we need to discover and expose them to more games suiting their preferences.

Second, extra caution is required to generalize the results because we could not recruit a more diverse participant pool. All the participants were Koreans, and the results could vary depending on participants from different cultural backgrounds and technology dissemination. In addition, we could not recruit participants who are prescribed or recommended gaming but who have no affinity to games. Moreover, as more female participants were recruited, the study does not include plentiful experiences of male participants. However, since mobile games are especially popular with middle-aged and older women [3], their experiences might be meaningful for mobile game designers.

Third, special attention is required when interpreting the results because the presence of young adults in the workshop setting could impact the gaming experience of participants. Our participants were able to obtain immediate help during the workshop from young partners; however, in many cases, playing alone would likely cause more difficulties.

Fourth, the scope of the quantitative analysis reports descriptive statistics at the exploration stage. The purpose of the quantitative analysis was to provide a useful indicator to game designers by showing the visual profiles of each case. Therefore, these tentative results require further refinement in strictly controlled experiments.

Lastly, with regard to the prior experience with computers and digital games, we found that there were differences between middle-aged adults and older adults in terms of gaming experience, but further research is needed to clarify these differences. In addition, when the digital native generation in their 20s and 30s is older, this guide is likely to have different criteria. Technologies are gradually evolving to be available intuitively without special learning. Therefore, this guide will require continuous updates through user studies from different cohorts.

Conclusion

This study includes considerations when creating games for the middle-aged and older population. Game design guides were produced based on the feedback of middle-aged and older adults during the game workshops, where they played mobile games of various genres. The implications of this study are suggesting a design guideline focusing on 10 categories of game elements. Consequently, game designers can holistically understand the game experience of older adults and easily find the relevant information. In addition, our empirical research was able to reaffirm the proposals from previous works that sought to improve the usability of a user interface. Furthermore, we also provided new guides for game design, such as touch interaction, game rules, stories, and advertising/payments. We found that most elements that participants were uncomfortable with during gameplay could be applied to all generations, regardless of age. The young generation also experiences difficulties with using touchscreen technologies, but they are better at adopting technological changes [68]. Thus, improving older adults' gaming experience may improve the experience for other generation groups as well. Rather than having special game design rules for the older population, we found that a universal design that can be easily accessed by gaming novices irrespective of their generation, is more important. In addition, design guides targeting specific capabilities and limitations are not effective because the older population is diverse in its ranges and combinations of limitations [69]. We hope this guide will be useful for various activities related to gameplay for the middle-aged and older adult population, such as for developing games targeting this population, modifying existing games while considering accessibility, and developing game-related educational programs for this population.

Acknowledgments

The authors would like to thank the Anyang senior welfare center and Seoul 50 plus central campus. The authors also appreciate all participants of the Intergenerational Play Workshop held by Korea Advanced Institute of Science and Technology in 2019. This research was supported by the Ministry of Culture, Sports and Tourism and the Korea Creative Content Agency through the Culture Technology Research and Development Program (Project R2019020010).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mobile game design guides for the middle-aged and older adult population.

[PDF File (Adobe PDF File), 93 KB - games_v9i2e24449_app1.pdf]

Multimedia Appendix 2

Descriptive statistics of participants' satisfaction levels in each game case.

[PDF File (Adobe PDF File), 215 KB - games_v9i2e24449_app2.pdf]

References

1. Digital market outlook: mobile games worldwide. Statista. URL: https://www.statista.com/outlook/211/100/mobile-games/worldwide#market-revenue [accessed 2020-09-04]
2. Mobile fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-09-04]
3. Nelson-Kakulla B. Gaming attitudes and habits of adults aged 50-Plus. AARP. URL: https://www.aarp.org/research/topics/technology/info-2019/2020-gaming-trends-older-americans.html [accessed 2020-09-04]
4. 2020 Game user survey. Korea Creative Content Agency. URL: http://www.kocca.kr/cop/bbs/view/B0000147/1842858.do?menuNo=201825 [accessed 2020-09-04]
5. Basak C, Boot WR, Voss MW, Kramer AF. Can training in a real-time strategy video game attenuate cognitive decline in older adults? Psychol Aging 2008 Dec;23(4):765-777 [FREE Full text] [doi: 10.1037/a0013494] [Medline: 19140648]

6. Green CS, Bavelier D. Exercising your brain: a review of human brain plasticity and training-induced learning. Psychol Aging 2008 Dec;23(4):692-701 [FREE Full text] [doi: 10.1037/a0014345] [Medline: 19140641]
7. Zelinski EM, Reyes R. Cognitive benefits of computer games for older adults. Gerontechnology 2009;8(4):220-235 [FREE Full text] [doi: 10.4017/gt.2009.08.04.004.00] [Medline: 25126043]
8. Bleakley CM, Charles D, Porter-Armstrong A, McNeill MDJ, McDonough SM, McCormack B. Gaming for health: a systematic review of the physical and cognitive effects of interactive computer games in older adults. J Appl Gerontol 2015 Apr;34(3):NP166-NP189. [doi: 10.1177/0733464812470747] [Medline: 24652863]
9. Hall AK, Chavarria E, Maneeratana V, Chaney BH, Bernhardt JM. Health Benefits of Digital Videogames for Older Adults: A Systematic Review of the Literature. Games Health J 2012 Dec;1(6):402-410. [doi: 10.1089/g4h.2012.0046] [Medline: 26192056]
10. Marston HR, Smith ST. Interactive Videogame Technologies to Support Independence in the Elderly: A Narrative Review. Games Health J 2012 Apr;1(2):139-152. [doi: 10.1089/g4h.2011.0008] [Medline: 26193189]
11. Miller KJ, Adair BS, Pearce AJ, Said CM, Ozanne E, Morris MM. Effectiveness and feasibility of virtual reality and gaming system use at home by older adults for enabling physical activity to improve health-related domains: a systematic review. Age Ageing 2014 Mar;43(2):188-195. [doi: 10.1093/ageing/aft194] [Medline: 24351549]
12. Allaire JC, McLaughlin AC, Trujillo A, Whitlock LA, LaPorte L, Gandy M. Successful aging through digital games: Socioemotional differences between older adult gamers and Non-gamers. Computers in Human Behavior 2013 Jul;29(4):1302-1306. [doi: 10.1016/j.chb.2013.01.014]
13. Kahlbaugh PE, Sperandio AJ, Carlson AL, Hauselt J. Effects of Playing Wii on Well-Being in the Elderly: Physical Activity, Loneliness, and Mood. Activities, Adaptation & Aging 2011 Oct;35(4):331-344. [doi: 10.1080/01924788.2011.625218]
14. Osmanovic S, Pecchioni L. Beyond Entertainment: Motivations and Outcomes of Video Game Playing by Older Adults and Their Younger Family Members. Games and Culture 2015 Sep 02;11(1-2):130-149. [doi: 10.1177/1555412015602819]
15. Brown JA. Let's play: understanding the role and meaning of digital games in the lives of older adults. 2012 Presented at: FDG '12: Proceedings of the International Conference on the Foundations of Digital Games; May 29-June 1; Raleigh, NC, USA p. 273-275. [doi: 10.1145/2282338.2282396]
16. Nap HH, de Kort YW, IJsselsteijn WA. Senior gamers: preferences, motivations and needs. Gerontechnology 2009;8(4):247-262. [doi: 10.4017/gt.2009.08.04.003.00]
17. Kaufman D, Sauvé L, Renaud L, Sixsmith A, Mortenson B. Older Adults' Digital Gameplay: Patterns, Benefits, and Challenges. Simulation & Gaming 2016 Apr 21;47(4):465-489. [doi: 10.1177/1046878116645736]
18. Loos E. Designing meaningful intergenerational digital games. In: Utrecht University Repository. 2014 Presented at: International Conference on Communication, Media, Technology and Design; April 24-26, 2014; Istanbul, Turkey URL: https://dspace.library.uu.nl/handle/1874/302902
19. Czaja SJ, Lee CC. The impact of aging on access to technology. Univ Access Inf Soc 2006 Dec 8;5(4):341-349. [doi: 10.1007/s10209-006-0060-x]
20. Johnson J, Finn K. Designing User Interfaces for an Aging Population: Towards Universal Design. Cambridge, MA, USA: Morgan Kaufmann; 2017.
21. Zaphiris P, Kurniawan S, Ghiawadwala M. A systematic approach to the development of research-based web design guidelines for older people. Univ Access Inf Soc 2006 Nov 16;6(1):59-75. [doi: 10.1007/s10209-006-0054-8]
22. Gamberini L, Alcaniz M, Barresi G, Fabregat M, Ibanez F, Prontu L. Cognition, technology and games for the elderly: An introduction to ELDERGAMES Project. PsychNology Journal. 2006. URL: https://www.researchgate.net/publication/200552925_Cognition_technology_and_games_for_the_elderly_An_introduction_to_ELDERGAMES_Project [accessed 2021-04-29]
23. Fua KC, Gupta S, Pautler D, Farber I. Designing serious games for elders. FDG 2013:291-297. [doi: 10.13140/RG.2.1.5166.9848]
24. Ijsselsteijn W, Nap H, de Kort Y, Poels K. Digital game design for elderly users. 2007 Presented at: Future Play '07: Proceedings of the 2007 conference on Future Play; November 15-17, 2007; Toronto, Canada p. 17-22. [doi: 10.1145/1328202.1328206]
25. Gerling K, Schulte F, Smeddinck J, Masuch M. Game design for older adultsffects of age-related changes on structural elements of digital games. 2012 Sep Presented at: ICEC 2012: International Conference on Entertainment Computing; September 26-29, 2012; Bremen, Germany p. 235-242. [doi: 10.1007/978-3-642-33542-6_20]
26. De Barros AC, Leitão R, Ribeiro J. Design and Evaluation of a Mobile User Interface for Older Adults: Navigation, Interaction and Visual Design Recommendations. Procedia Computer Science 2014;27:369-378. [doi: 10.1016/j.procs.2014.02.041]
27. De Schutter B. Never Too Old to Play: The Appeal of Digital Games to an Older Audience. Games and Culture 2010 May 07;6(2):155-170. [doi: 10.1177/1555412010364978]
28. Salmon JP, Dolan SM, Drake RS, Wilson GC, Klein RM, Eskes GA. A survey of video game preferences in adults: Building better games for older adults. Entertainment Computing 2017 Jun;21:45-64. [doi: 10.1016/j.entcom.2017.04.006]

29. Boot WR, Champion M, Blakely DP, Wright T, Souders DJ, Charness N. Video games as a means to reduce age-related cognitive decline: attitudes, compliance, and effectiveness. Front Psychol 2013;4:31. [doi: 10.3389/fpsyg.2013.00031] [Medline: 23378841]
30. Pearce C. The Truth About Baby Boomer Gamers: A Study of Over-Forty Computer Game Players. Games and Culture 2008 Feb 13;3(2):142-174. [doi: 10.1177/1555412008314132]
31. Chesham A, Wyss P, Müri RM, Mosimann UP, Nef T. What Older People Like to Play: Genre Preferences and Acceptance of Casual Games. JMIR Serious Games 2017 Apr 18;5(2):e8 [FREE Full text] [doi: 10.2196/games.7025] [Medline: 28420601]
32. Blocker KA, Wright TJ, Boot WR. Gaming Preferences of Aging Generations. Gerontechnology 2014;12(3):174-184 [FREE Full text] [doi: 10.4017/gt.2014.12.3.008.00] [Medline: 29033699]
33. Sherry J, Lucas K, Greenberg B, 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 URL: https://www.researchgate.net/profile/John-Sherry-2/publication/259583577_Video_game_uses_and_gratifications_as_predictors_of_use_and_game_preference/links/54dc196b0cf28d3de65e9fed/Video-game-uses-and-gratifications-as-predictors-of-use-and-game-preference.pdf
34. De Schutter B, Brown J. Digital Games as a Source of Enjoyment in Later Life. Games and Culture 2015 Jul 12;11(1-2):28-52. [doi: 10.1177/1555412015594273]
35. De Schutter B, Malliet S. The older player of digital games: A classification based on perceived need satisfaction. Communications 2014 Mar 4;39(1):67-88. [doi: 10.1515/commun-2014-0005]
36. De Carvalho R, Ishitani L. Motivational factors for mobile serious games for elderly users. 2012 Presented at: Proceedings of XI SB Games; November 2-4, 2012; Brazil URL: http://sbgames.org/sbgames2012/proceedings/papers/cultura/C_F3.pdf
37. Cota TT, Ishitani L, Vieira N. Mobile game design for the elderly: A study with focus on the motivation to play. Computers in Human Behavior 2015 Oct;51:96-105. [doi: 10.1016/j.chb.2015.04.026]
38. Gajadhar BJ, Nap HH, de Kort YW, IJsselsteijn WA. Out of sight, out of mind: co-player effects on seniors' player experience. 2010 Presented at: Fun and Games '10: Proceedings of the 3rd International Conference on Fun and Games; September 15-17, 2010; Leuven, Belgium p. 74-83. [doi: 10.1145/1823818.1823826]
39. Souders DJ, Boot WR, Charness N, Moxley JH. Older Adult Video Game Preferences in Practice: Investigating the Effects of Competing or Cooperating. Games and Culture 2016;11(1-2):170-120 [FREE Full text] [doi: 10.1177/1555412015603538] [Medline: 29033698]
40. 2019 essential facts about the computer and video game industry. Entertainment Software Association. URL: https://www.theesa.com/wp-content/uploads/2019/05/ESA_Essential_facts_2019_final.pdf [accessed 2020-09-04]
41. De Schutter B, Vanden Abeele V. Designing meaningful play within the psycho-social context of older adults. 2010 Presented at: Fun and Games '10: Proceedings of the 3rd International Conference on Fun and Games; September 2010; Leuven, Belgium p. 84-93. [doi: 10.1145/1823818.1823827]
42. McLaughlin A, Gandy M, Allaire J, Whitlock L. Putting fun into video games for older adults. Ergonomics in Design 2012 Apr 20;20(2):13-22 [FREE Full text] [doi: 10.1177/1064804611435654]
43. Marston HR. Design Recommendations for Digital Game Design within an Ageing Society. Educational Gerontology 2013 Feb;39(2):103-118. [doi: 10.1080/03601277.2012.689936]
44. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: Defining "gamification". 2011 Presented at: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; September 28-30, 2011; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]
45. Creswell JW. Research design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA, USA: SAGE Publications, Inc; 2014.
46. Bergvall-Kåreborn B, Ståhlbröst A. Living Lab: an open and citizen-centric approach for innovation. International Journal of Innovation and Regional Development 2009;1(4):356-370. [doi: 10.1504/ijird.2009.022727]
47. Homescapes. Playrix. URL: https://www.playrix.com/mobile-games/ [accessed 2019-08-28]
48. Fruit ninja. Halfbrick. URL: https://www.halfbrick.com/ [accessed 2019-08-28]
49. Long journey of life. Pinx Studio. URL: https://www.pinx.studio/ [accessed 2019-08-28]
50. Brawl stars. Supercell. URL: https://supercell.com/en/ [accessed 2019-08-28]
51. Sally's law. Nanali studio. 2018. URL: http://www.nanali.net/official.php/home/eng [accessed 2019-08-28]
52. Good pizza, great pizza. TapBlaze. URL: https://www.tapblaze.com/ [accessed 2019-08-28]
53. Word tower - world trip. Monster Planet. URL: https://play.google.com/store/apps/details?id=com.MonsterPlanet.WordTower2 [accessed 2019-08-28]
54. Go-stop plus. Buzzpowder. URL: https://play.google.com/store/apps/details?id=com.encom.gostop_play&hl=en_CA&gl=US [accessed 2019-08-28]
55. Lonely one: hole-in-one. LoadComplete. URL: https://loadcomplete.com/ [accessed 2019-08-28]
56. Digital skills critical for jobs and social inclusion. UNESCO. URL: https://en.unesco.org/news/digital-skills-critical-jobs-and-social-inclusion [accessed 2020-12-31]

57. Richardson A. Using customer journey maps to improve customer experience. Harvard Business Review. URL: http://www.iimagineservicedesign.com/wp-content/uploads/2015/07/Experience-Maps-Using-Customer-Journey-Maps-to-Improve-Customer-Experience.pdf [accessed 2020-09-04]
58. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
59. Nowell LS, Norris JM, White DE, Moules NJ. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. International Journal of Qualitative Methods 2017 Oct 02;16(1):160940691773384 [FREE Full text] [doi: 10.1177/1609406917733847]
60. Ermi L, Mäyrä F. Fundamental components of the gameplay experience: Analysing immersion. In: Changing Views: Worlds in Play. 2005 Presented at: DiGRA 2005; June 16-20, 2005; Vancouver, Canada URL: https://ir.lib.sfu.ca/item/260
61. Fullerton T. Game Design Workshop: A Playcentric Approach to Creating Innovative Games 4th edition. Boca Raton, FL, USA: CRC Press; 2018.
62. De Almeida RXE, Ferreira SBL, Soares HP. Recommendations for the Development of Web Interfaces on Tablets/iPads with Emphasis on Elderly Users. Procedia Computer Science 2015;67:140-149. [doi: 10.1016/j.procs.2015.09.258]
63. De Schutter B, Vanden Abeele V. Towards a Gerontoludic Manifesto. Anthropology & Aging 2015 Nov 19;36(2):112-120. [doi: 10.5195/AA.2015.104]
64. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
65. Charlier N, Ott M, Remmele B, Whitton N. Not just for children: game-based learning for older adults. 2012 Presented at: The 6th European Conference on Games Based Learning; October 4-5, 2012; Cork, Ireland p. 102-108 URL: https://www.researchgate.net/profile/Bernd-Remmele/publication/256591353_Not_just_for_children_Game-based_learning_for_older_adults/links/59106071aca272f38d59827a/Not-just-for-children-Game-based-learning-for-older-adults.pdf
66. Hunicke R, LeBlanc M, Zubek R. MDA: A formal approach to game design and game research. AAAI. URL: https://www.aaai.org/Papers/Workshops/2004/WS-04-04/WS04-04-001.pdf [accessed 2020-09-04]
67. De Schutter B. Gerontoludic design: extending the MDA framework to facilitate meaningful play for older adults. International Journal of Gaming and Computer-Mediated Simulations (IJGCMS) 2017;9(1):45-60. [doi: 10.4018/ijgcms.2017010103]
68. Page T. Touchscreen mobile devices and older adults: a usability study. International Journal of Human Factors and Ergonomics 2014;3(1):65-85. [doi: 10.1504/IJHFE.2014.062550]
69. Kascak LR, Rébola CB, Sanford JA. Integrating universal design (UD) principles and mobile design guidelines to improve design of mobile health applications for older adults. 2014 Presented at: 2014 IEEE International Conference on Healthcare Informatics; September 15-17; Verona, Italy p. 343-348. [doi: 10.1109/ICHI.2014.54]

Abbreviations

RQ: research question

Edited by N Zary; submitted 21.09.20; peer-reviewed by C Victor, H Söbke; comments to author 08.12.20; revised version received 01.01.21; accepted 12.02.21; published 20.05.21.

Please cite as:
Lee S, Oh H, Shi CK, Doh YY
Mobile Game Design Guide to Improve Gaming Experience for the Middle-Aged and Older Adult Population: User-Centered Design Approach
JMIR Serious Games 2021;9(2):e24449
URL: https://games.jmir.org/2021/2/e24449
doi:10.2196/24449
PMID:34014175

Original Paper

A Virtual Reality–Based App to Educate Health Care Professionals and Medical Students About Inflammatory Arthritis: Feasibility Study

Philipp Klemm[1*], MD; Arnd Kleyer[2,3*], MD; Koray Tascilar[2,3], MD; Louis Schuster[2,3], MSc; Timo Meinderink[2,3], MSc; Florian Steiger[2,3]; Uwe Lange[1], MD; Ulf Müller-Ladner[1], MD; Johannes Knitza[2,3], MD; Philipp Sewerin[4], MD; Johanna Mucke[4], MD; Alexander Pfeil[5], MD; Georg Schett[2,3], MD; Fabian Hartmann[2,3], BSc; Axel J Hueber[6], MD; David Simon[2,3], MD

[1]Department of Rheumatology, Immunology, Osteology and Physical Medicine, Justus-Liebig University Gießen, Campus Kerckhoff, Bad Nauheim, Germany

[2]Department of Internal Medicine 3, Rheumatology and Immunology, Friedrich-Alexander University Erlangen-Nuremberg and Universitätsklinikum Erlangen, Erlangen, Germany

[3]Deutsches Zentrum Immuntherapie, Friedrich-Alexander University Erlangen-Nuremberg and Universitätsklinikum Erlangen, Erlangen, Germany

[4]Department and Hiller Research Unit for Rheumatology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[5]Department of Internal Medicine 3, Jena University Hospital, Friedrich Schiller University, Jena, Germany

[6]Sektion Rheumatologie, Sozialstiftung Bamberg, Bamberg, Germany

[*]these authors contributed equally

Corresponding Author:
David Simon, MD
Department of Internal Medicine 3, Rheumatology and Immunology
Friedrich-Alexander University Erlangen-Nuremberg and Universitätsklinikum Erlangen
Ulmenweg 18
Erlangen, 91054
Germany
Phone: 49 091318543253
Fax: 49 091318539091
Email: david.simon@uk-erlangen.de

Abstract

Background: Inflammatory arthritides (IA) such as rheumatoid arthritis or psoriatic arthritis are disorders that can be difficult to comprehend for health professionals and students in terms of the heterogeneity of clinical symptoms and pathologies. New didactic approaches using innovative technologies such as virtual reality (VR) apps could be helpful to demonstrate disease manifestations as well as joint pathologies in a more comprehensive manner. However, the potential of using a VR education concept in IA has not yet been evaluated.

Objective: We evaluated the feasibility of a VR app to educate health care professionals and medical students about IA.

Methods: We developed a VR app using data from IA patients as well as 2D and 3D-visualized pathological joints from X-ray and computed tomography–generated images. This VR app (Rheumality) allows the user to interact with representative arthritic joint and bone pathologies of patients with IA. In a consensus meeting, an online questionnaire was designed to collect basic demographic data (age, sex); profession of the participants; and their feedback on the general impression, knowledge gain, and potential areas of application of the VR app. The VR app was subsequently tested and evaluated by health care professionals (physicians, researchers, and other professionals) and medical students at predefined events (two annual rheumatology conferences and academic teaching seminars at two sites in Germany). To explore associations between categorical variables, the χ^2 or Fisher test was used as appropriate. Two-sided *P* values ≤.05 were regarded as significant.

Results: A total of 125 individuals participated in this study. Among them, 56% of the participants identified as female, 43% identified as male, and 1% identified as nonbinary; 59% of the participants were 18-30 years of age, 18% were 31-40 years old, 10% were 41-50 years old, 8% were 51-60 years old, and 5% were 61-70 years old. The participants (N=125) rated the VR app as excellent, with a mean rating of 9.0 (SD 1.2) out of 10, and many participants would recommend use of the app, with a mean recommendation score of 3.2 (SD 1.1) out of 4. A large majority (120/125, 96.0%) stated that the presentation of pathological bone formation improves understanding of the disease. We did not find any association between participant characteristics and evaluation of the VR experience or recommendation scores.

Conclusions: The data show that IA-targeting innovative teaching approaches based on VR technology are feasible.

(JMIR Serious Games 2021;9(2):e23835) doi:10.2196/23835

KEYWORDS

feasibility; virtual reality; inflammatory arthritis; psoriatic arthritis; rheumatoid arthritis

Introduction

Inflammatory arthritides (IA) such as rheumatoid arthritis (RA) or psoriatic arthritis (PsA) are complex diseases affecting millions of people worldwide, causing functional impairment and a reduced quality of life [1,2]. Owing to the chronic inflammation, in addition to joint symptoms, some patients with RA or PsA can also experience localized or generalized bone disease. Local bone changes (erosions, bone proliferation) result in functional impairment, while systemic bone loss and worsening biomechanical properties increase the risk of developing osteoporosis and pathologic fractures [3-5].

Imaging is one of the key tools for an early diagnosis of IA, as well as for monitoring and understanding the bone pathology. Various imaging techniques can help to detect pathologic structural changes, and to assess bone density and quality. X-ray, as a 2D imaging instrument, is well-established in routine clinical practice, and is frequently used for assessing catabolic (erosion) and anabolic (bone proliferation) changes or joint space narrowing. Using high-resolution peripheral quantitative computed tomography (HR-pQCT) can further help to characterize joint disease in IA with an in-depth 3D analysis of erosive and osteoproliferative changes, precise quantification of trabecular and cortical bone densities, and assessment of biomechanical properties [6-9].

Recent data show that there is still room for improvement in the understanding and recognition of pathologies, interpretation of image findings, as well as in the appropriate usage of proper imaging techniques [10-12]. Accordingly, there seems to be some concern about musculoskeletal imaging, particularly among health professionals in training [10-12]. Moreover, physicians, medical students, and health professionals explicitly state a need for more in-depth training in imaging and its interpretation regarding musculoskeletal disorders [10,13-15].

To improve the teaching of musculoskeletal imaging, conventional teaching methods such as textbooks, online services, and simulators already exist, which impart knowledge through a systematic and standardized approach. One of these teaching tools is the imaging course of the European League Against Rheumatism (EULAR). Virtual reality (VR) technology can be used to complement these established traditional methods. VR refers to a realistic and immersive simulation of a 3D environment created with interactive software and hardware that is experienced and controlled by movements of the body. In contrast to traditional teaching tools, VR technology has the advantage of visualizing images of various imaging tools such as computed tomography in three spatial axes. The VR technique allows users to see and interact with 3D objects. This so-called “immersive” experience enables the consumer to touch, rearrange, scale, and walk through the objects (to see the inside of the image objects). Previously, VR apps were investigated in several patients with different diseases [16]. In the field of rheumatology, two studies have focused on different patient groups and evaluated how VR apps can lead to improved health care. Venuturupalli et al [17] showed that VR apps could be a viable solution for the treatment of pain and anxiety in rheumatic patients. In addition, Botella et al [18] demonstrated that usage of VR can have a long-term benefit with significantly reduced pain and depression in patients with fibromyalgia. However, all of these studies were patient-focused applications of VR.

Recently, University Hospital Erlangen and the Bamberg Hospital (a teaching hospital) started developing a VR app (Rheumality) that contains 2D and 3D data of the pathologic joints from real patients with RA and PsA at different stages of the disease [19]. These imaging data were combined with the corresponding medical history and clinical data. This app uses immersive sensation that could improve the understanding of chronic inflammatory changes in the joints and bones of patients with IA. Users can touch, scale, or even immerse themselves in the typical bone and joint pathologies of RA or PsA, such as erosions, new bone formations, or bone loss. The advanced visualization and interactivity of 3D bone images paired with the corresponding medical history offer potential to improve training and education in rheumatology. Ideally, the virtual world will contribute to enhancing end-user knowledge and help to acquire confidence in the management of IA.

However, the feasibility of a VR app in teaching about IA has not yet been evaluated, and therefore its potential benefits have not yet been assessed. In this study, we evaluated the feasibility of the Rheumality VR app that was developed specifically to educate health care professionals and medical students about IA. We investigated the general impression, knowledge gain, and possible applications using data from physicians, researchers, and other health care professionals working in the field of rheumatology as well as medical students.

Methods

Development of a VR App for Educational Purposes in IA

In 2016, University Hospital Erlangen (Erlangen, Germany) and Bamberg Hospital (Bamberg, Germany) started to develop an educational VR app (called Rheumality) based on real anonymized patient cases with corresponding clinical and imaging data to provide health care professionals and medical students with a better understanding of IA [19]. The imaging data focused particularly on peripheral X-rays and HR-pQCT images of the arthritic joints and bone structures. The version (2.1) of Rheumality used for this study included three case examples (two RA cases and one PsA case), and was developed for the above-mentioned targeted population (Figure 1). All patients agreed to the use of their anonymized data and signed an informed consent form, which was approved by the Ethics Committee of Friedrich-Alexander-University (324_16B).

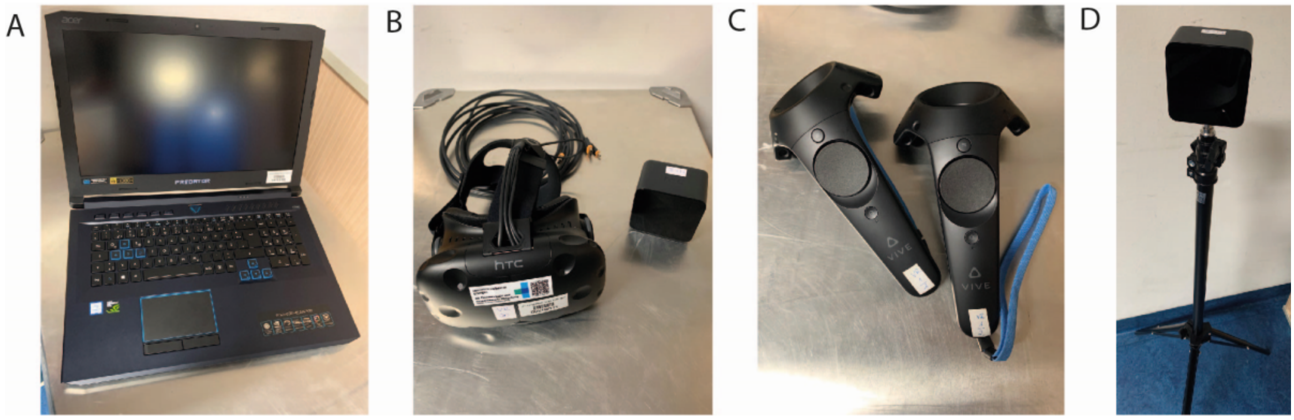
For each case, brief clinical information about disease history (eg, initial diagnosis, symptoms, treatment, response to treatment, disease perception, patient reported outcomes) is given to the user (Multimedia Appendix 1). The user has to participate actively, and has to indicate the typical arthritic joint and bone pathologies on the basis of 2D X-ray images of the hands (Figure 1A) and 3D HR-pQCT images of the finger joints (Figure 1C). Subsequently, the participant is asked to estimate how much the patient is restricted in different areas of life, taking into account the pathologies depicted in the images (questions based on the Modified Stanford Health Assessment Questionnaire [20]) (Figure 1B). Each case (tutorial) is designed to last 10 minutes.

Figure 1. Learning tasks in the VR application for teaching of IA. Depiction of 2D hand X-ray of a patient with arthritis on which the user has interactively identified joint pathologies (green circles) (A). In another task, the user is asked to assess the disease burden of the same patient after evaluating the X-ray images (B), 3D CT image of the same pathologies of the X-ray image, the user can rotate and immerse with these images and is asked to define specific pathologies such as erosions (red circle) (C). (A, C) show images of a patient with long-standing (inadequately treated) RA, (D) CT images of a patient with early RA.

Technical Equipment for the VR Experience

A notebook Acer Predator Helios 500 (Intel i7-8750H, NVIDIA GeForce GTX 1070; acquisition cost: US $2417) and HTC VIVE VR headset (2x OLED 1080×1200, 90 Hz, FOV 110°, 2x Lighthouse Tracking, Motion Controller; acquisition cost: US $628) were used in this study. The setup time was 5 minutes to clear the required space in a room (safety measure), connect all devices, switch them on, and set up and run the VR app. Pictures of the equipment are provided in Figure 2.

Figure 2. Applied technical virtual reality (VR) equipment. For the VR experience, a Notebook (Acer Predator Helios 500) (A), VR set (VR Headset HTC VIVE) (B), motion controller (C), and Lighthouse Tracking (D) are necessary.



Questionnaire Development

To evaluate possible applications and benefits for improved disease understanding from using the VR app, a questionnaire was designed in a consensus meeting (web-based, involving authors PK, AK, DS, AH, PS, AF) after initial individual research of the current literature and a definition of the issues to be addressed [21,22]. Results of the respective consultations were then assessed by the group: each single item and wording was presented to all experts, followed by a vote about the inclusion. For a question to be accepted for the final questionnaire, a majority vote of 75% was required. This method was chosen in analogy to the EULAR recommendation task force [23].

Thirteen questions were included in the final questionnaire covering how the VR app and the VR itself were perceived. All questions of the questionnaire are depicted in Multimedia Appendix 2. In addition, basic demographic data (gender, age category, and profession) were collected.

Evaluation of the VR App as a Teaching Tool

The VR app (Rheumality version 2.1) was provided to a professional audience at predefined events according to a standardized scheme and subsequently evaluated using the developed questionnaire. The professional community consisted of health professionals working in the field of rheumatology (physicians, researchers, and other health professionals) and medical students. Three different methods for exploitation and evaluation were selected to reach the broadest possible target population: (i) the annual German Society of Rheumatology conference 2019 in Dresden, Germany; (ii) the German Imaging Course in Rheumatology (Deutscher Bildgebungskurs) 2019 in Duesseldorf, Germany; and (iii) academic teaching seminars (from October 2019 to March 2020) for medical students at two different German universities (Campus Kerckhoff of Justus-Liebig-University Gießen and Friedrich-Alexander-University Erlangen). During events (i) and (ii), the possibility of participation was advertised to the public through flyers, posters, and a trade fair booth. Regarding (iii), the students were directly approached in corresponding teaching seminars about their interest in participating. Inclusion criteria were aged above 18 years and belonging to one of the following professional backgrounds: medical student, resident, specialist (hospitalist), specialist (outpatient practitioner), researcher, other health care professional (eg, nurse, study nurse). Pregnant participants as well as participants suffering from nausea, vomiting, dementia, motion sickness, stroke, seizure, and epilepsy were not allowed to participate. To qualify for the evaluation, all individuals had to complete the full VR tutorial (10 minutes). Directly after the VR experience, the participants were asked to answer the questionnaire once providing informed consent. Prior to each VR session, the participants were informed by a trained member of the research team (LS, TM, FS, PK, DS, AK) about the background of the app (presentation of real patient cases in combination with the respective images of the associated pathologic joints), basic control mechanism, and safety measures (eg, what to do when feeling uncomfortable). Each VR session was accompanied by a qualified staff member (LS, TM, FS, PK, DS, AK), who also evaluated subsequent side effects. The app was evaluated by participants directly after the VR session using the developed questionnaire based on a web-based survey (SurveyMonkey Inc).

Statistical Analysis

Nominal and ordinal variables are described by counts and percentages, and interval scaled variables are described by mean and SD. To explore associations between categorical variables, the χ^2 or Fisher test was used as appropriate. Two-sided *P* values ≤.05 were regarded as statistically significant. Data were analyzed using SSPS for Windows (SPSS 26, IBM Corporation).

Results

Study Participants

A total of 133 individuals were tested for eligibility, 8 of whom had to be excluded (history of nausea, epilepsy, and motion sickness; presence of pregnancy), resulting in a final dataset from 125 individuals included for analysis. Among these 125 participants, 50 were physicians, including 27 specialists in rheumatology (16 hospitalists, 11 outpatient practitioners), 23 residents, 5 researchers, and 4 other health care professionals (Table 1). The remaining participants were medical students (n=66). Seventy of the 125 participants (56.0%) identified as female, 54 (43.2%) identified as male, and 1 (0.8%) identified as nonbinary. Among the 125 participants, 74 (59.2%) were 18-30 years of age, 22 (17.6%) were 31-40 years old, 13 (10.4%) were 41-50 years old, 10 (8.0%) were 51-60 years old, and 6 (4.8%) were 61-70 years old. Overall, 16% (20/125) of the participants already had experience with VR technology.

Table 1. Professional background of participants (N=125).

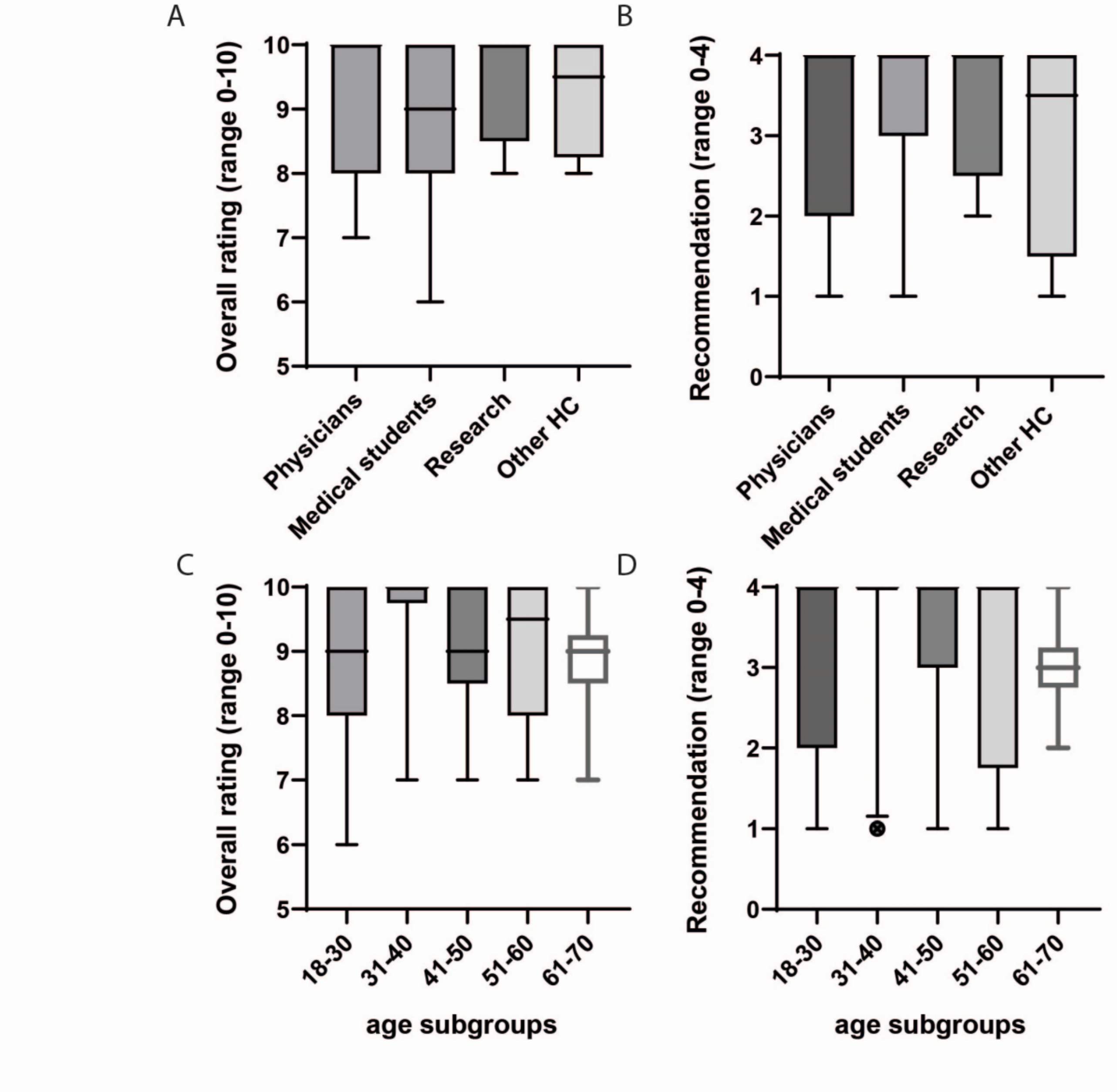
Profession	Participants, n (%)
Medical student	66 (52.8)
Resident	23 (18.4)
Specialist (hospitalists)	16 (12.8)
Specialist (outpatient practitioner)	11 (8.8)
Researcher	5 (4.0)
Health care professional	4 (3.2)

General Feedback on the Imaging VR App

The participants generally rated the app as excellent. The overall rating (best rating: 10, worst rating: 0) of the VR app was excellent (mean 9.0, SD 1.2), and participants could imagine recommending the app to others (mean 3.2, SD 1.1), based on a scale from 0 (not at all) to 4 (definitely) (Figure 3). Using the χ^2 test, neither age nor profession was significantly associated with a given evaluation of the VR app (χ^2=21.7, *P*=.15 and χ^2=12.7, *P*=.89, respectively) or a special recommendations to friends and family (χ^2=14.6, *P*=.89 and χ^2=13.8, *P*=.54, respectively).

The majority (102/125, 81.6%) of participants indicated that they would use the app themselves if it was available for download in an app store. When asked to evaluate the intuitiveness of the VR app, 109 of the 125 participants (87.2%) considered the control to be intuitive, whereas only 7 out of 125 (5.6%) found the VR app to be confusing. All 7 participants rating the VR to be confusing were under 41 years of age (6 were 18-30 years old and 1 was 31-40 years old). However, there was no statistically significant association with age (χ^2=2.647, *P*=.62). The vast majority (109/125, 87.2%) also rated the amount of time they spent on the virtual patient case to be optimal.

Figure 3. Feedback on the virtual reality (VR) app Rheumality. Overall rating (range 0-10) on the VR app divided into four different professional subgroups (A) and age subgroups (C). Recommendations of the VR app in the four different professional subgroups (B) and age subgroups (D) on a scale of 0 (not at all) to 4 (definitely). In the boxplots, crossbars represent medians, whiskers represent the 5th-95th percentiles (points below the whiskers are drawn as individual points); the box always extends from the 25th to 75th percentiles (hinges of the plot). HC: health care professionals.



Knowledge Dissemination by the VR App

When asked whether the presentation of the pathologic bone formation improved disease understanding, 120 of the 125 participants (96.0%) answered "yes." In addition, 124 of the 125 participants (99.2%) stated that the way the disease was represented by the VR app was helpful for gaining a better understanding of the disease. Similarly, 123 of the 125 participants (98.4%) stated an interest in obtaining additional tutorials including new case studies or for other rheumatic diseases.

Feedback on Potential Areas of Application

When asked about further prospects of application, 123 of the 125 participants (98.4%) thought that the developed VR concept based on real medical histories and imaging data could have a positive influence on teaching IA and rheumatology. In accordance with this, 118 of 125 participants (94.4%) were of the opinion that a VR app could be successfully used for teaching about IA and rheumatology. In addition, 51 of the 125 participants (40.8%) saw potential for the app in research, 95 of 125 (76.0%) saw potential for patient education, and 98 of 125 (78.4%) saw potential for training of health professionals.

Harms

One study participant reported nausea and had to stop the VR experience after 6 minutes.

Discussion

Principal Findings

This study demonstrated that a broad professional audience considers an educational VR app (called Rheumality) based on real patient data to have potential for knowledge transfer about IA. The combination of medical history and image data for the same individual convinced the large majority of the participants of the educational benefit of the app. Interestingly, the acceptance of this teaching approach was evident across all age groups and was independent of medical education. Thus, physicians, researchers, and health care professionals working in the field of rheumatology, as well as medical students, would accept and recommend this teaching format as a good knowledge transfer tool. Besides the high acceptance, we also found that the VR app seems to be a safe educational tool, since only one participant had to stop the VR experience due to nausea. With regard to further fields of application, a considerable majority of the participants stated that VR technology has potential for broader areas of implementation. A substantial percentage of the participants could imagine that, in addition to classical teaching and training aspects, patient education could also be a promising field of application.

Traditional teaching approaches based on textbooks or online texts are successful in transmitting knowledge, but the intuitive depiction of 3D imaging reality is limited due to the 2D characteristic of these teaching tools. Therefore, new teaching methods focusing on imaging education can be used as an additional knowledge transfer tool.

Our study shows that many individuals currently consider that the developed VR concept, which utilizes the latest hardware and software possibilities, could have a positive influence on teaching.

There are also various technological developments in rheumatic and musculoskeletal teaching and patient care (such as mobile apps or activity trackers) that take this aspect into account [24,25]. However, to our knowledge, there is no app currently available that transports 3D imaging of real patients in combination with actual patient data on a technological VR basis. Given the recognized importance of IA imaging for diagnosis, monitoring, and prognosis, such a VR app can be valuable in the future to reliably train key players of the health care system for clinical routine practice and training to improve the care of IA patients [14,26,27].

VR technology (goggles and VR-ready computers) is becoming more and more affordable, immersive, and flexible, leading to the expectation that its use in teaching purposes may increase in the coming years [28]. Our study shows that VR apps are already technically at such an advanced level that teaching about IA with VR technology seems feasible.

Limitations

The major limitation of our study is that the effect of learning with this VR app compared with the standard (eg, textbooks) was not evaluated, which is needed to ultimately demonstrate its practical usefulness. For this purpose, further studies will have to be performed in the future. In addition, a specific, advanced VR app was used in this study, which relied on the combination of patient data and real imaging data. Therefore, the results of our survey cannot be transferred to all VR apps. Further, confirmatory studies with larger numbers of participants and more differentiated target audiences, especially patients, are necessary to verify and validate our study results. In addition, the current VR equipment is, compared to traditional teaching methods, relatively expensive for an implementation in clinical practice, with costs of several hundred dollars for a powerful computer and the necessary goggles. Only when acquisition costs continue to decrease, a comprehensive introduction into the clinical health care environment will be possible. It would be useful to perform future studies with additional health care professionals from several countries to transfer our results to the international community.

Comparison With Prior Work

VR apps have been investigated in several patients with different diseases [16]. However, VR apps have rarely been applied in the field of rheumatology to date, and this prior work has thus far been largely patient-focused [17,18]. To our knowledge, there has been no study specifically evaluating the teaching potential of VR in rheumatology or IA until now. The "digital divide" is frequently mentioned when assessing new technologies for a broad audience, meaning that older individuals seem to experience more difficulty in adopting new technologies than younger individuals [29-31]. Interestingly, the tested VR app in this study was rated positively independent of age and profession. Furthermore, only a small minority of participants found the VR experience to be confusing, all of whom were younger than 41 years. A conclusive judgment on this matter would be difficult at this point; however, it could be speculated that this minority of younger users had different expectations in advance due to their previous experience and therefore considered the VR concept to be confusing.

Conclusions

Our study showed that novel teaching approaches based on VR technology are feasible for teaching about IA. The use of VR apps such as Rheumality enables disease-specific knowledge visualization and may add a new educational pillar.

Acknowledgments

Rheumality is a registered trademark of Lilly Deutschland GmbH. Rheumality was developed in cooperation with the University Hospital Erlangen, Germany. This study was supported by the Deutsche Forschungsgemeinschaft (DFG-FOR2886 PANDORA and the CRC1181). Additional funding was received by the Bundesministerium für Bildung und Forschung (BMBF; project METARTHROS, MASCARA), H2020 GA 810316- 4D-Nanoscope ERC Synergy Project, IMI-funded project RTCure, Emerging Fields Initiative MIRACLE of the Friedrich-Alexander-Universität Erlangen-Nürnberg, and Else Kröner-Memorial Scholarship

(DS, no. 2019_EKMS.27). The present work was performed in (partial) fulfillment of the requirements for obtaining the degree “Dr. rer. biol. hum.” for DS at Friedrich-Alexander-University Erlangen-Nürnberg. Further support was provided by the Deutsche Rheumastiftung (to DS and PK) and NWT-med of Justus-Liebig-University Gießen (to PK).

Conflicts of Interest

PK, AK, FH, DS are advisors to Eli Lilly and Company, Germany.

Multimedia Appendix 1
Video of the developed virtual reality app for knowledge transfer on rheumatic and musculoskeletal diseases.
[MOV File , 51867 KB - games_v9i2e23835_app1.mov]

Multimedia Appendix 2
Final developed questionnaire.
[DOCX File , 17 KB - games_v9i2e23835_app2.docx]

References

1. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med 2011 Dec 08;365(23):2205-2219. [doi: 10.1056/NEJMra1004965] [Medline: 22150039]
2. Ritchlin CT, Colbert RA, Gladman DD. Psoriatic arthritis. N Engl J Med 2017 Mar 09;376(10):957-970. [doi: 10.1056/NEJMra1505557] [Medline: 28273019]
3. Ogdie A, Harter L, Shin D, Baker J, Takeshita J, Choi HK, et al. The risk of fracture among patients with psoriatic arthritis and psoriasis: a population-based study. Ann Rheum Dis 2017 May 16;76(5):882-885 [FREE Full text] [doi: 10.1136/annrheumdis-2016-210441] [Medline: 28093419]
4. Simon D, Kleyer A, Faustini F, Englbrecht M, Haschka J, Berlin A, et al. Simultaneous quantification of bone erosions and enthesiophytes in the joints of patients with psoriasis or psoriatic arthritis - effects of age and disease duration. Arthritis Res Ther 2018 Aug 31;20(1):203 [FREE Full text] [doi: 10.1186/s13075-018-1691-z] [Medline: 30170626]
5. Stemmler F, Simon D, Liphardt A, Englbrecht M, Rech J, Hueber AJ, et al. Biomechanical properties of bone are impaired in patients with ACPA-positive rheumatoid arthritis and associated with the occurrence of fractures. Ann Rheum Dis 2018 Jul 23;77(7):973-980 [FREE Full text] [doi: 10.1136/annrheumdis-2017-212404] [Medline: 29475856]
6. Simon D, Kleyer A, Bui CD, Hueber A, Bang H, Ramming A, et al. Micro-structural bone changes are associated with broad-spectrum autoimmunity and predict the onset of rheumatoid arthritis. Arthritis Rheumatol 2020 Mar 18:41229. [doi: 10.1002/art.41229] [Medline: 32067367]
7. Simon D, Kleyer A, Stemmler F, Simon C, Berlin A, Hueber AJ, et al. Age- and sex-dependent changes of intra-articular cortical and trabecular bone structure and the effects of rheumatoid arthritis. J Bone Miner Res 2017 Apr 18;32(4):722-730. [doi: 10.1002/jbmr.3025] [Medline: 27787923]
8. Simon D, Kleyer A, Englbrecht M, Stemmler F, Simon C, Berlin A, et al. A comparative analysis of articular bone in large cohort of patients with chronic inflammatory diseases of the joints, the gut and the skin. Bone 2018 Nov;116:87-93. [doi: 10.1016/j.bone.2018.07.017] [Medline: 30048820]
9. Simon D, Tascilar K, Unbehend S, Bayat S, Berlin A, Liphardt A, et al. Bone mass, bone microstructure and biomechanics in patients with hand osteoarthritis. J Bone Miner Res 2020 Sep 02;35(9):1695-1702. [doi: 10.1002/jbmr.4046] [Medline: 32395822]
10. Marino K, Merrick D, Edwards K, Pratten M. Musculoskeletal radiology teaching at a UK medical school: do we need to improve? Anat Sci Educ 2019 May 05;12(3):257-263. [doi: 10.1002/ase.1834] [Medline: 30395696]
11. Atsina K, Parker L, Rao VM, Levin DC. Advanced imaging interpretation by radiologists and nonradiologist physicians: a training issue. AJR Am J Roentgenol 2020 Jan;214(1):W55-W61. [doi: 10.2214/AJR.19.21802] [Medline: 31691611]
12. Krusche M, Sewerin P, Kleyer A, Mucke J, Vossen D, Morf H, Rheumadocs und Arbeitskreis Junge Rheumatologie (AGJR). Specialist training quo vadis? Z Rheumatol 2019 Oct 29;78(8):692-697. [doi: 10.1007/s00393-019-00690-5] [Medline: 31468164]
13. Beyer C, Ramiro S, Sivera F, Mandl P, Machado PM, Ospelt C, EMEUNET Country Liaisons Group. Educational needs and preferences of young European clinicians and physician researchers working in the field of rheumatology. RMD Open 2016 Jul 27;2(2):e000240 [FREE Full text] [doi: 10.1136/rmdopen-2015-000240] [Medline: 27493789]
14. Mandl P, Ciechomska A, Terslev L, Baraliakos X, Conaghan P, D'Agostino MA, et al. Implementation and role of modern musculoskeletal imaging in rheumatological practice in member countries of EULAR. RMD Open 2019 Jun 25;5(2):e000950 [FREE Full text] [doi: 10.1136/rmdopen-2019-000950] [Medline: 31321076]
15. Joshua F, Bailey C, Marabani M, Romas E, White R, Wong P. Perceptions of Doppler ultrasound for rheumatoid arthritis disease activity assessment and education. Int J Rheum Dis 2019 Jan;22(1):55-61. [doi: 10.1111/1756-185X.13034] [Medline: 28205411]

16. Snoswell AJ, Snoswell CL. Immersive virtual reality in health care: systematic review of technology and disease states. JMIR Biomed Eng 2019 Sep 26;4(1):e15025. [doi: 10.2196/15025]
17. Venuturupalli RS, Chu T, Vicari M, Kumar A, Fortune N, Spielberg B. Virtual reality-based biofeedback and guided meditation in rheumatology: a pilot study. ACR Open Rheumatol 2019 Dec 11;1(10):667-675. [doi: 10.1002/acr2.11092] [Medline: 31872189]
18. 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]
19. Kleyer A, Simon D, Hartmann F, Schuster L, Hueber A. "Virtual rheumatology" : A new teaching concept for rheumatology of the future? Z Rheumatol 2019 Mar;78(2):112-115. [doi: 10.1007/s00393-019-0594-y] [Medline: 30694358]
20. Pincus T, Summey JA, Soraci SA, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum 1983 Nov;26(11):1346-1353. [doi: 10.1002/art.1780261107] [Medline: 6639693]
21. Mosadeghi S, Reid MW, Martinez B, Rosen BT, Spiegel BMR. Feasibility of an immersive virtual reality intervention for hospitalized patients: an observational cohort study. JMIR Ment Health 2016 Jun 27;3(2):e28 [FREE Full text] [doi: 10.2196/mental.5801] [Medline: 27349654]
22. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]
23. van der Heijde D, Aletaha D, Carmona L, Edwards CJ, Kvien TK, Kouloumas M, et al. 2014 Update of the EULAR standardised operating procedures for EULAR-endorsed recommendations. Ann Rheum Dis 2015 Jan 26;74(1):8-13 [FREE Full text] [doi: 10.1136/annrheumdis-2014-206350] [Medline: 25261577]
24. Knitza J, Tascilar K, Messner E, Meyer M, Vossen D, Pulla A, et al. German mobile apps in rheumatology: review and analysis using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2019 Aug 05;7(8):e14991 [FREE Full text] [doi: 10.2196/14991] [Medline: 31381501]
25. Gossec L, Guyard F, Leroy D, Lafargue T, Seiler M, Jacquemin C, et al. Detection of flares by decrease in physical activity, collected using wearable activity trackers in rheumatoid arthritis or axial spondyloarthritis: an application of machine learning analyses in rheumatology. Arthritis Care Res (Hoboken) 2019 Oct 27;71(10):1336-1343. [doi: 10.1002/acr.23768] [Medline: 30242992]
26. Kaeley GS, Bakewell C, Deodhar A. The importance of ultrasound in identifying and differentiating patients with early inflammatory arthritis: a narrative review. Arthritis Res Ther 2020 Jan 02;22(1):1 [FREE Full text] [doi: 10.1186/s13075-019-2050-4] [Medline: 31898524]
27. Tins BJ, Butler R. Imaging in rheumatology: reconciling radiology and rheumatology. Insights Imaging 2013 Dec 15;4(6):799-810 [FREE Full text] [doi: 10.1007/s13244-013-0293-1] [Medline: 24127271]
28. VR headsets are falling in price - is this good or bad news for virtual reality? Kei Studios. URL: https://kei-studios.com/vr-headsets-are-dropping-in-price-is-this-good-or-bad-news-for-virtual-reality/ [accessed 2021-04-26]
29. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
30. Czaja SJ, Sharit J. Age differences in attitudes toward computers. J Gerontol B Psychol Sci Soc Sci 1998 Sep 01;53(5):P329-P340. [doi: 10.1093/geronb/53b.5.p329] [Medline: 9750571]
31. Ellis RD, Allaire JC. Modeling computer interest in older adults: the role of age, education, computer knowledge, and computer anxiety. Hum Factors 1999 Sep 02;41(3):345-355. [doi: 10.1518/001872099779610996] [Medline: 10665203]

Abbreviations

EULAR: European League Against Rheumatism
HR-pQCT: high-resolution peripheral quantitative computed tomography
IA: inflammatory arthritis
PsA: psoriatic arthritis
RA: rheumatoid arthritis
VR: virtual reality

Edited by N Zary; submitted 27.08.20; peer-reviewed by S Stones, S Yamaguchi; comments to author 08.10.20; revised version received 23.11.20; accepted 13.03.21; published 11.05.21.

Please cite as:
Klemm P, Kleyer A, Tascilar K, Schuster L, Meinderink T, Steiger F, Lange U, Müller-Ladner U, Knitza J, Sewerin P, Mucke J, Pfeil A, Schett G, Hartmann F, Hueber AJ, Simon D
A Virtual Reality–Based App to Educate Health Care Professionals and Medical Students About Inflammatory Arthritis: Feasibility Study
JMIR Serious Games 2021;9(2):e23835
URL: https://games.jmir.org/2021/2/e23835
doi:10.2196/23835
PMID:33973858

Original Paper

User Experience With Dynamic Difficulty Adjustment Methods for an Affective Exergame: Comparative Laboratory-Based Study

Ali Darzi[1], PhD; Sean M McCrea[2], PhD; Domen Novak[1], PhD

[1]Department of Electrical and Computer Engineering, University of Wyoming, Laramie, WY, United States

[2]Department of Psychology, University of Wyoming, Laramie, WY, United States

Corresponding Author:
Domen Novak, PhD
Department of Electrical and Computer Engineering
University of Wyoming
1000 E University Avenue
Dept 3295
Laramie, WY, 82071
United States
Phone: 1 307 766 2279
Email: dnovak1@uwyo.edu

Abstract

Background: In affective exergames, game difficulty is dynamically adjusted to match the user's physical and psychological state. Such an adjustment is commonly made based on a combination of performance measures (eg, in-game scores) and physiological measurements, which provide insight into the player's psychological state. However, although many prototypes of affective games have been presented and many studies have shown that physiological measurements allow more accurate classification of the player's psychological state than performance measures, few studies have examined whether dynamic difficulty adjustment (DDA) based on physiological measurements (which requires additional sensors) results in a better user experience than performance-based DDA or manual difficulty adjustment.

Objective: This study aims to compare five DDA methods in an affective exergame: manual (player-controlled), random, performance-based, personality-performance–based, and physiology-personality-performance–based (all-data).

Methods: A total of 50 participants (N=50) were divided into five groups, corresponding to the five DDA methods. They played an exergame version of Pong for 18 minutes, starting at a medium difficulty; every 2 minutes, two game difficulty parameters (ball speed and paddle size) were adjusted using the participant's assigned DDA method. The DDA rules for the performance-based, personality-performance–based, and all-data groups were developed based on data from a previous open-loop study. Seven physiological responses were recorded throughout the sessions, and participants self-reported their preferred changes to difficulty every 2 minutes. After playing the game, participants reported their in-game experience using two questionnaires: the Intrinsic Motivation Inventory and the Flow Experience Measure.

Results: Although the all-data method resulted in the most accurate changes to ball speed and paddle size (defined as the percentage match between DDA choice and participants' preference), no significant differences between DDA methods were found on the Intrinsic Motivation Inventory and Flow Experience Measure. When the data from all four automated DDA methods were pooled together, the accuracy of changes in ball speed was significantly correlated with players' enjoyment (r=0.38) and pressure (r=0.43).

Conclusions: Although our study is limited by the use of a between-subjects design and may not generalize to other exergame designs, the results do not currently support the inclusion of physiological measurements in affective exergames, as they did not result in an improved user experience. As the accuracy of difficulty changes is correlated with user experience, the results support the development of more effective DDA methods. However, they show that the inclusion of physiological measurements does not guarantee a better user experience even if it yields promising results in offline cross-validation.

(JMIR Serious Games 2021;9(2):e25771) doi:10.2196/25771

KEYWORDS

affective computing; dynamic difficulty adaptation; exergames; physiological measurements; task performance; personality characteristics; psychophysiology

Introduction

Affective Exergames

Exercise games (commonly shortened to exergames) are used to promote enjoyable, intensive exercise in applications such as weight loss and maintenance [1,2], healthy aging [3], and motor rehabilitation [4,5]. In such exergames, tailoring the game difficulty to the player's abilities and preferences can often improve the user experience, enabling more enjoyable, frequent, and intensive exercise. Such difficulty adjustment can always be performed manually by a player via an interface. However, although manual adjustment is generally accurate, it can cause interruptions in game flow, as users must stop playing the game to adjust difficulty [6]. This has led to the development of dynamic difficulty adjustment (DDA) methods, where the game assesses the player's current state and automatically adjusts the difficulty to bring the player into a more desirable state.

The simplest and most popular DDA methods are based only on performance measures (eg, in-game score), which provide an easily measurable and interpretable indicator of perceived game difficulty. However, although many studies have shown the positive effects of performance-based DDA on user experience [7], performance alone does not necessarily provide insight into the player's psychological state, for example, a frustrated player can get a high score in a game without enjoying it. This has led to the development of affective games, an emerging type of videogame that adapts difficulty based on a combination of the player's performance and their psychological (cognitive and affective) state. This psychological state can be defined as, for example, the level of anxiety [8,9] or via the two-dimensional valence-arousal model [9,10] and is commonly inferred from measurements such as an electrocardiogram (ECG), galvanic skin response, and electroencephalography (EEG) using automated classification algorithms. Once the psychological state has been classified, DDA can be performed using rules such as "if performance is high and anxiety is low, increase difficulty." Thus, incorporating psychological information allows affective games to potentially achieve more effective and personalized DDA than performance-based methods [11].

Affective exergames are most prominent in motor rehabilitation, where physiological measurements have been used to estimate psychological states and adjust the difficulty of exercises for both the upper [5,12-16] and lower limbs [17,18]; they have also been used for general exercise enhancement in several studies [19,20]. However, as affective exergames are more complex than simple performance-based DDA, the question arises: does the additional cost and complexity result in a better user experience?

Do Affective Exergames Improve User Experience?

As mentioned, in affective exergames, psychological states are extracted from measurements such as ECG, EEG, and galvanic skin response using signal processing and machine learning techniques. The ground truth for such machine learning is generally the user's self-reported state or opinion, for example, how anxious they are or how they would like difficulty to be adjusted [8,11,21]. The *accuracy* of DDA is then estimated as the percentage of times that the affective exergame makes the same DDA decision (based on the extracted psychological state) as the user would. This accuracy is never perfect; however, to justify the use of affective games, affective DDA should reach a higher accuracy than performance-based DDA and result in a better user experience than performance-based DDA.

Many studies in both affective exergaming and other fields of affective computing have shown that the addition of physiological measurements allows for a more accurate classification of the player's psychological state compared with using only performance measurements [12,17,22-24]. Furthermore, some studies have shown that affective exergames that perform DDA based on physiological measurements achieve a positive user experience [5,13-15]. However, there is little evidence that the user experience achieved with affective exergames is better than that achieved with a performance-based exergame in the same context. To the best of our knowledge, only one study has examined this difference and has found that an affective exergame was more engaging than an equivalent performance-based exergame [14]. Even outside affective exergaming, the evidence in favor of affective DDA is limited, with a few studies in entertainment games showing better user experience with affective DDA than performance-based DDA [8] and a few studies showing a better user experience with affective DDA than manual (user-guided) DDA [25,26], although the user's decisions are frequently used as the ground truth for training affective DDA.

This lack of evidence showing that affective technologies result in a better user experience than simpler technologies has been acknowledged as a significant issue in affective computing [27,28], as calculating the accuracy of psychological state classification or demonstrating a positive user experience with affective DDA may not be enough to show the superiority of affective technologies. For example, our own *Wizard of Oz* studies in affective games have shown that, although user psychological state classification accuracy is correlated with user experience, the relationship between the two is complex and nonlinear [29,30]. Furthermore, studies from other fields of human-machine interaction indicate that offline classification accuracy does not always strongly correlate with user experience or even with real-time (dynamic) classification accuracy [31,32]. Thus, lessons learned from offline classification studies cannot be directly transferred to physiology-based DDA. The state of the art in affective exergaming can be summarized as follows: there is evidence that affective exergames result in a positive user experience and that physiology-based psychological state classification is more accurate than performance-based psychological state classification. However, there is very limited evidence that physiology-based DDA (used in affective exergames) results in a better user experience than performance-based DDA (used in simpler exergames). Additional evidence in this regard is needed to justify the broader adoption of affective exergames.

Goal of the Study

This study aims to evaluate the user experience in an exergame where DDA is performed using one of five methods: manual, random, performance-based (PE), personality-performance–based (PEPE), and physiology-personality-performance–based (all-data). The last three DDA methods were based on classifiers developed for offline psychological state classification in our previous study [33]. In this study, these classifiers were connected to *if-then* difficulty adjustment rules and used as a basis for DDA. The research questions were as follows:

- Research question 1: do the PEPE and all-data methods result in a better user experience than the PE method? The all-data method can be considered an affective exergame, and the all-data with PE comparison thus represents a direct comparison of affective and performance-based exergames. In our previous study, including personality and physiological measurements increased psychological state classification accuracy compared with using only performance measures [33], but this is not guaranteed to result in a better user experience. The inclusion of personality characteristics was considered separately because tailoring DDA to personality may improve user experience without additional sensors [34];
- Research question 2: do manual and random DDA result in the best and worst user experience, respectively? The user's preference is commonly used as theground truth for training DDA algorithms in affective games [8,11,21], so manual DDA should result in a very positive user experience. At the same time, some studies have shown better user experience with physiology-based DDA than with manual DDA [25,26]. Random DDA should result in a poor experience but is included as a baseline;
- Research question 3: is user experience positively correlated with the accuracy of psychological state classification during gameplay? Our previous Wizard of Oz [29,30] studies and others' studies [28] indicated a correlation between the two, but the nature of this relationship in actual games remains unclear.

Methods

Overview

This paper describes a comparative evaluation of five DDA methods in an affective exergame. Three of these methods were developed based on data recorded in a previous study [33]; thus, we briefly refer to a previous study for better understanding. Both studies were approved by the University of Wyoming Institutional Review Board (protocol no 20190822DN02495).

This section is divided into five subsections. The first subsection presents the exergame used for DDA evaluation, which is identical for both studies. The second subsection describes the different measurement types used as a basis for DDA, which were nearly identical for both studies. The third subsection summarizes the previous study used to train the DDA methods for this study (described in detail in our previous paper [33]). The fourth subsection presents the study protocol. Finally, the fifth subsection presents the outcome variables and data analysis performed to compare the five DDA methods in this study.

Exergame

An exergame version of Pong was reused from our previous research, originally intended for two-player rehabilitation exergaming [35,36], and a single-player version was created for our previous open-loop study [33]. It consists of 2 paddles and a ball on a board. The participant controlled the bottom paddle, while the top paddle was controlled by a computer opponent. If the ball passed either player's paddle, the other player would score a point, and the ball would start moving again from the middle of the screen. The player moved their paddle left and right by tilting the Bimeo (Kinestica) arm tracking device left and right with their dominant hand. The game was played on a 21-inch screen, with the participant seated approximately 60 cm from the screen. A photograph of this study setup is shown in Figure 1.

The game difficulty can be adjusted using two parameters: ball speed and paddle size. Although the game allows the participant's and opponent's paddle sizes to be changed independently, the evaluated DDA methods always changed the 2 paddles simultaneously and identically so that the participant's paddle was never larger or smaller than that of the computer opponent. The DDA methods used in this study were based on multiple measurements, with the decision-making rules trained based on data recorded in a previous study, as described in the following sections.

Figure 1. A participant playing the exergame (on the screen) using the Bimeo device (right hand) while wearing the different physiological sensors on the head and nondominant hand. The eye tracker is visible below the screen.

Measurements Used as the Basis for DDA

A total of 3 data types were recorded and used as a basis for DDA: physiological responses, performance, and personality characteristics. Physiological responses and performance were recorded continuously during gameplay and thus varied as the difficulty of the game changed. Personality characteristics were collected using questionnaires at the start of the session (after the researcher demonstrated the game but before participants played it) and did not change during gameplay but were expected to influence the participant's DDA preferences.

The physiological recording process was identical in this study and our previous open-loop study [33] so that the previous data could be directly used to create the DDA methods; parts of the text in this section and Multimedia Appendix 1 are thus rephrased from the previous paper.

Physiological Signals

This study used 2 g.USBamp signal amplifiers and associated sensors (g.tec Medical Engineering GmbH) to measure six physiological signals: 8-channel EEG, 2-channel electrooculogram, ECG, respiration, galvanic skin response, and skin temperature. A seventh physiological signal, point of gaze, was recorded using a GP3 eye tracker (Gazepoint). The sensors used are shown in Figure 1. Detailed information about the measurements is available in Multimedia Appendix 1 but, in brief, they were performed as follows:

- EEG was recorded from 8 locations based on the 10-20 placement system [37]: AF3, AF4, F1, F2, F5, F6, C1, and C2. Feature extraction methods included the lateral power spectrum density [38] and dispersion entropy [39].
- Electrooculogram was recorded from 2 channels reflecting up-down and left-right eye movements. The extracted features were based on its first derivative. In addition, it was used to denoise the EEG signals.
- ECG was recorded using 4 electrodes on the trunk. Extracted features included heart rate and time- and frequency-domain estimates of heart rate variability [40].
- Respiration was recorded using a thermistor-based sensor in front of the nose and mouth. Extracted features included respiration rate and time-domain estimates of respiratory rate variability.
- The skin temperature was recorded using a sensor attached to the little finger of the nondominant hand. Extracted features included the mean temperature and changes in skin temperature across time.
- Galvanic skin response was recorded by attaching a sensor to the index and middle fingers of the nondominant hand. Features were extracted from both the tonic (low frequency) component and the phasic skin conductance responses [41].
- From the eye tracker data, the extracted features included the size of each pupil and the mean gaze velocity.

Physiological features were normalized by dividing each calculated feature value by the feature value obtained during the baseline period.

Performance

The in-game score was used as the only feature to assess participants' performance. It was defined as the difference between the participant's score and the computer opponent's score.

Personality Characteristics

Participants completed four personality questionnaires: the Learning and Performance Goal Orientation measure [42], the Behavioral Inhibition and Activation Scales [43], the Self-Efficacy Scale [44], and the 10-Item Personality Inventory [45]. Questionnaire details are provided in Multimedia Appendix 1.

Summary of the Previous Study

In our previous study [33], 30 healthy university students (mean 24.2, SD 4.4 years; 11 women) participated in a 1-hour session. Physiological signals were first recorded for a 2-minute baseline period, during which participants were instructed to relax, remain motionless, and look at the computer screen. Then, nine game difficulty configurations consisting of combinations of three possible ball speeds (slow, medium, and fast) and three possible paddle sizes (small, medium, and large) were played in random order for 2 minutes each. After each 2-minute interval, a short questionnaire was filled out to assess the perceived difficulty, enjoyment, and two subjective preferences about game difficulty (desired change to ball speed and paddle size). Participants had seven options for perceived difficulty and enjoyment (1 for very low and 7 for very high) and five options for how they would like to change each difficulty parameter, ranging from a decrease by two levels (−2) to an increase by two levels (+2). The order of game difficulty settings was preset randomly for each participant, and the participant's desired changes to the ball speed and paddle size did not actually affect the game.

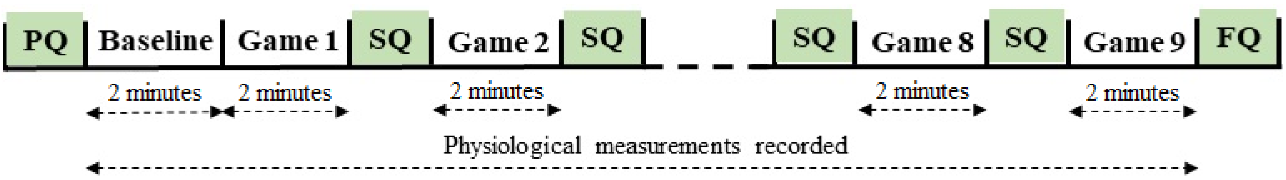
The study protocol resulted in a data set with two outputs (desired changes to ball speed and paddle size) and 64 input features calculated during 2-minute gameplay periods: the physiological, personality, and performance features described in Multimedia Appendix 1. Multiple classifiers were developed to categorize the two subjective difficulty adjustment preferences into three classes: increase (+1 and +2), no change (0), and decrease (−1 and −2). Separate classifiers were developed and evaluated for ball speed and paddle size, and separate classifiers were trained for different combinations of the three recorded data types (Multimedia Appendix 1). These classifiers were separated so that a classifier did not have access to the results of another classifier (eg, the ball speed classifier did not have access to the results of paddle size classification and vice versa); however, they were still trained on data from the same participants and gameplay periods, and all had access to current ball speed and current paddle size (as the current game state would be realistically available to any classifier). In all cases, stepwise forward feature selection [46] with an inclusion threshold of P=.05 was first used to select a subset of features. This reduced subset was then used to train four classifier types, of which multiple linear regression resulted in the highest classification accuracy and was thus chosen for this study.

Study Protocol for Comparison of DDA Methods

This study compared five DDA methods using the same study setup and a protocol similar to the open-loop study. A total of 50 healthy university students (mean 25.1, SD 5.9 years; 13 women; 5 left-handed) participated in the study, with 10 assigned to each DDA method. The 5 participant groups went through the same study protocol, which differed only according to the DDA method they were assigned to. Participants were not told which method they were assigned to and had no way of identifying it.

After signing the consent form, filling out the four personality questionnaires, putting on the physiological sensors, and recording the responses for a 2-minute baseline period (same as in the *Summary of the Previous Study* section), participants started to play the game with a medium difficulty level (speed level 3 in a range of 1-6 and paddle size 2 in a range of 1-4). Every 2 minutes, the game was paused, and participants filled out a short questionnaire to report their perceived difficulty, enjoyment, and the way they would like to change the game parameters (same as in the *Summary of the Previous Study* section). Participants were allowed to ask for longer breaks in case of dizziness or arm fatigue; however, this only happened after a 2-minute interval among all 50 participants, and dizziness and fatigue were not otherwise tracked. Once the participant was ready to continue, the difficulty level was adjusted by changing the ball speed and paddle size using the DDA method to which the participant was assigned. In all DDA methods, the ball speed and paddle size were adjusted independently of each other. All participants played the game for a total of 18 minutes (nine 2-minute intervals, with difficulty adjustment after each interval) and then completed two outcome questionnaires (see the Outcome Variables and Data Analysis section) to complete the study protocol. The protocol is summarized in Figure 2.

Figure 2. Summary of the study protocol. All participants first completed the personality questionnaires, rested for a 2-minute baseline period, then went through nine 2-minute gameplay intervals (game 1-9), with a short questionnaire after each. Physiological measurements were recorded throughout the baseline and gameplay and separated into 2-minute intervals for analysis. At the end, participants filled out the final outcome questionnaires. FQ: final questionnaire; PQ: personality questionnaire; SQ: short questionnaire.



As mentioned earlier, the participant groups differed according to the DDA method; in total, five methods were used:

1. Manual adjustment: the ball speed and paddle size were adjusted based on the participant's preferences expressed in the short questionnaire. For each parameter, they had three options: increase by 1 level, no change, and decrease by 1 level.
2. Random adjustment: for both ball speed and paddle size, one of the three options available to the manual method (increase, no change, or decrease by one level) was chosen entirely randomly.
3. PE method: the ball speed and paddle size were adjusted based on three features: the current ball speed, current paddle size, and the in-game score achieved by participants. To perform the adjustment, two multiple linear regression models were trained based on data from a previous study using the three features as the inputs and the adjustment to the ball speed and paddle size as outputs (the exact regression coefficients are presented in Table A1 in Multimedia Appendix 1). The output of each regression model was converted to discrete classes and used to perform difficulty adjustment: increase by 1 level (output of model above 0.5), no change (output between −0.5 and 0.5), and decrease by 1 level (output below −0.5).
4. PEPE method: the ball speed and paddle size were adjusted based on multiple features selected among current game difficulty, in-game scores, and personality characteristics. The adjustment was performed using two regression-based classifiers, as with the PE method, but trained using both performance and personality data. A mix of personality characteristics from all four questionnaires was included in the classifiers, with only self-efficacy and agreeableness included in both classifiers (exact regression coefficients are presented in Tables A2 and A3 in Multimedia Appendix 1).
5. All-data method: the ball speed and paddle size were adjusted based on multiple features selected among current game difficulty, in-game score, personality characteristics (eg, extroversion), and physiological responses (eg, respiration rate). Adjustments were made using two regression-based classifiers developed based on data from a previous study. For both ball speed and paddle size classifiers, the first two selected features were from skin temperature and respiration, and multiple EEG features and multiple personality characteristics were selected (exact regression coefficients are presented in Tables A4 and A5 in Multimedia Appendix 1).

Among the five DDA methods, only the manual method took the participants' preferences regarding game difficulty into account. Furthermore, limits were put in to ensure that the ball speed and paddle size did not exceed the preset limits. Specifically, if a DDA method's decision would have caused a parameter to exceed a minimum or maximum value (1-6 for ball speed; 1-4 for paddle size), it instead stayed at that extreme value.

Outcome Variables and Data Analysis

The primary outcome of the study was the effect of different DDA methods on user experience. In addition, two secondary analyses were performed. First, the closed-loop classification accuracy was calculated for the 50 participants and correlated with the user experience. Second, the classifiers were retrained on the data from the 50 participants using the same methods as in the previous study.

Effect of DDA Methods on User Experience

The effect of the different DDA methods on user experience was assessed using two self-report questionnaires at the end of the 18-minute gameplay period: the Intrinsic Motivation Inventory (IMI) [47] and the Flow Experience Measure (FEM) [48]. The IMI is an 8-item questionnaire that assesses effort/importance, perceived competence, interest/enjoyment, and pressure/tension with two items per assessed variable. The same version was used in our previous exergaming research [35,36]. The FEM assesses a single, variable flow, using 8 items. This resulted in 5 outcome variables in total, which were compared between the methods using two-tailed two-sample *t* tests. Intrinsic motivation and flow are perhaps the two most commonly evaluated short-term outcomes of serious games and exercise [6,19,23,25,44,47] and were thus considered appropriate for this study.

Accuracy of Speed and Paddle Size Changes

In this study, three DDA methods (PE, PEPE, and all-data) used regression-based classifiers developed based on data from a previous study [33]. Their accuracy was defined as the percentage of agreement between the classifier's opinion and the participant's preference for changes in ball speed or paddle size (measured via the short questionnaire after each 2-minute gameplay interval). This accuracy was expected to correlate with user experience when the classifier was used as a basis for difficulty adjustment. In a previous study, the all-data method resulted in the most accurate classification, whereas PE resulted in the least accurate classification; however, the accuracies in this study may be different for multiple reasons. For example, the range of ball speeds in this study was wider than in the previous study, and new situations, even within the same scenario, may induce ungeneralizable physiological responses, significantly reducing classification accuracy [27]. Simply having new participants may also reduce classification accuracy, although this was expected to have a smaller influence as participants in this and previous studies were drawn from the same broad pool. Thus, the classification accuracy was recalculated for each DDA method using data from this study.

To determine whether the accuracy of changes in ball speed and paddle size was correlated with user experience, Spearman correlation coefficients were calculated between each participant's accuracy of ball speed and paddle size adjustments and the IMI and FEM outcomes. This was done across 40 participants; the manual group was dropped because they all had an accuracy of 100%.

Classifier Retraining and Validation

As mentioned in the previous subsection, we expected that the accuracy of the three classifiers used for the three DDA methods

(PE, PEPE, and all-data) would not be the same in this study as it was in the previous study [33]. Although the data collection in this study was performed with prebuilt classifiers, these classifiers could be retrained offline after the data collection conclusion to determine whether the all-data method is still the most accurate if validated on the new data.

To perform this training and revalidation, three combinations (PE, PEPE, and all-data) of the three data types (physiology, performance, and personality characteristics) from 80 participants (30 from a previous study and 50 from this study) were used as inputs to train classifiers that categorize perceived difficulty, enjoyment, desired change to ball speed, and desired change to paddle size (obtained from the short questionnaire) into three classes: *low/decrease*, *medium/no change*, and *high/increase*. Reference output values for all classifiers were obtained from the short questionnaire, and Table 1 presents the selected reference ranges of the short questionnaire answers for each class and variable. These ranges were chosen to provide the most even possible spread of samples among the three classes for each variable (although an even spread was not always feasible because of biases in the data). The current ball speed and paddle size were added to all input data combinations because they indicate the current game state and are available to any practical model.

Table 1. Defined ranges for retraining the three-class classifiers.

Class	Difficulty	Enjoyment	Speed change	Paddle size change
Low				
Range	1 to 2	1 to 4	−2 to −1	−2 to −1
Sample, n	180	246	62	151
Medium				
Range	3 to 5	5	0	0
Sample, n	364	184	303	368
High				
Range	6 to 7	6 to 7	1 to 2	1 to 2
Sample, n	176	290	355	355

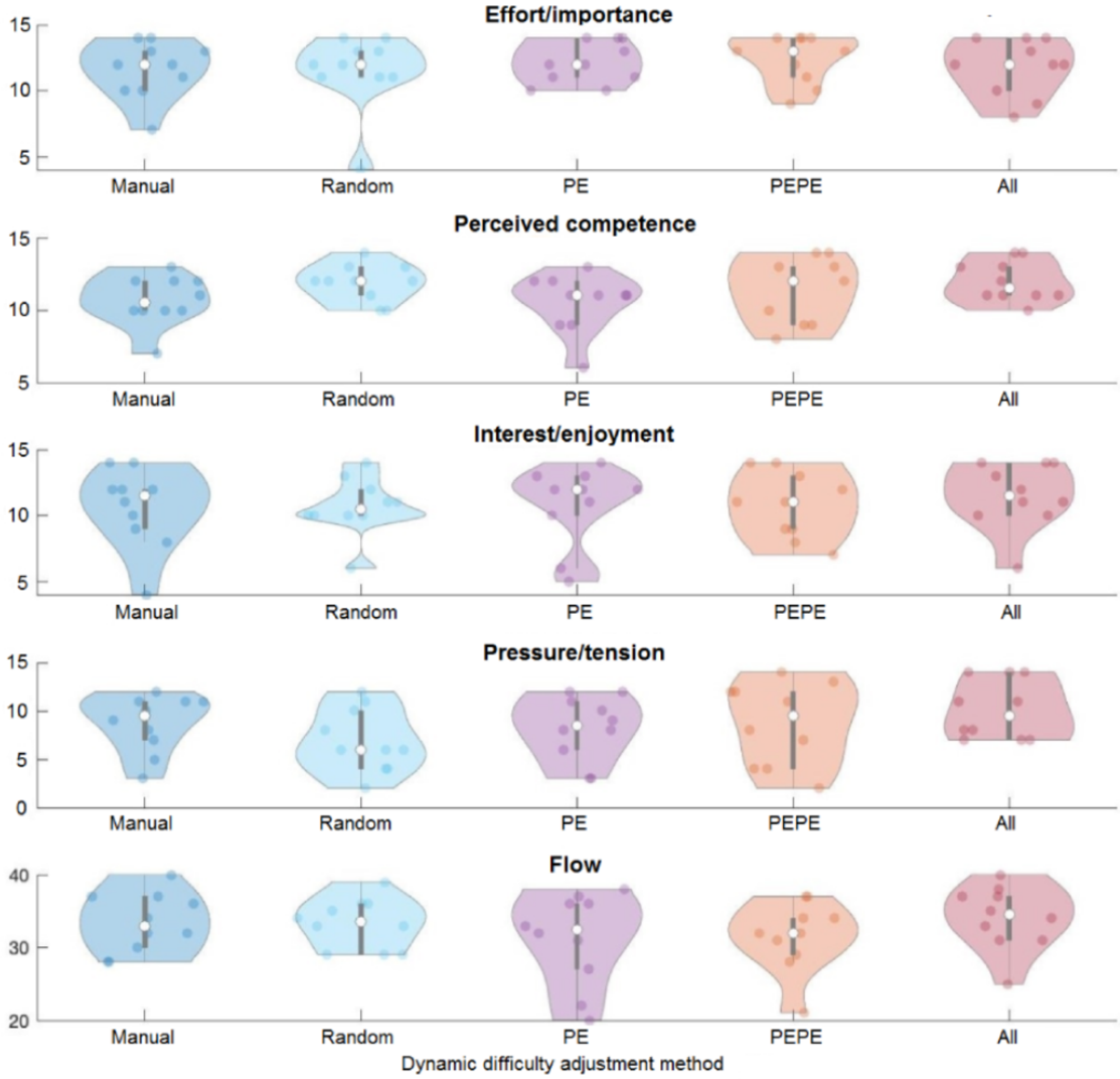
To reduce the number of features before classification, stepwise forward feature selection [46] with an inclusion threshold of 0.05 was used to find the most informative set of features. Then, to classify the input data, four different classifiers were used: a support vector machine with a linear kernel, linear discriminant analysis, ensemble decision tree, and multiple linear regression. The classifiers were validated using 10-fold cross-validation (72 participants' data used to train and 8 participants' data to validate the classifier; the procedure was repeated 10 times with each participant in the validation data set once).

Results

Effects of DDA Methods on User Experience

All extracted features and questionnaire results are available in Multimedia Appendix 2. Figure 3 shows the five outcomes of the IMI and FEM for all five DDA methods, presented as a violin plot of the 10 participants assigned to each method. The outcomes of the IMI and FEM were compared between different DDA methods using two-tailed two-sample *t* tests; however, only one difference was significant: pressure/tension was higher in the all-data method than in the random method (*P*=.04).

Figure 3. Violin plot of postexperiment self-report questionnaire outcomes. all data: based on physiology, personality, and performance. Each participant is indicated as a dot. The first four questionnaire outcomes have a possible range of 2-14, whereas flow has a possible range of 8-40. PE: adjustment based on performance; PEPE: adjustment based on personality and performance.



Classifier Accuracy and Retraining

Table 2 shows the accuracy of each DDA method in the dynamic prediction of the participants' desired changes to the game difficulty parameters. Accuracy was calculated by comparing the classifier's prediction and the participants' preference expressed on the short questionnaire after all 2-minute intervals. The manual method trivially achieves 100% accuracy because it uses the participants' answers as a basis for difficulty adjustment. Among the other four methods, the all-data method was the most accurate for predicting the desired adjustments to ball speed and paddle size.

Figure 4 depicts the ball speed and paddle size for nine 2-minute intervals of each DDA method.

Table 3 presents the accuracies of three-class classification of the four answers to the short questionnaire when retraining and validating the classifiers using three combinations of input data types obtained from 80 participants. Combining all data types (all-data) resulted in the most accurate classifiers.

Table 2. Prediction accuracy of changes to the game difficulty parameters.

Difficulty parameter	Difficulty adjustment method (%)				
	Manual	Random	PE[a]	PEPE[b]	All-data[c]
Ball speed	100	32.5	53.7	43.7	62.5
Paddle size	100	41.2	46.2	33.5	55.0

[a]Based on performance.

[b]Based on personality and performance.

[c]Based on physiology, personality, and performance.

Figure 4. Ball speed (left) and paddle size (right) of five dynamic difficulty adjustment methods for nine 2-minute game intervals across 10 participants. Error bars indicate 95% CI. PE: based on performance; PEPE: based on personality and performance; all data: based on physiology, personality, and performance.

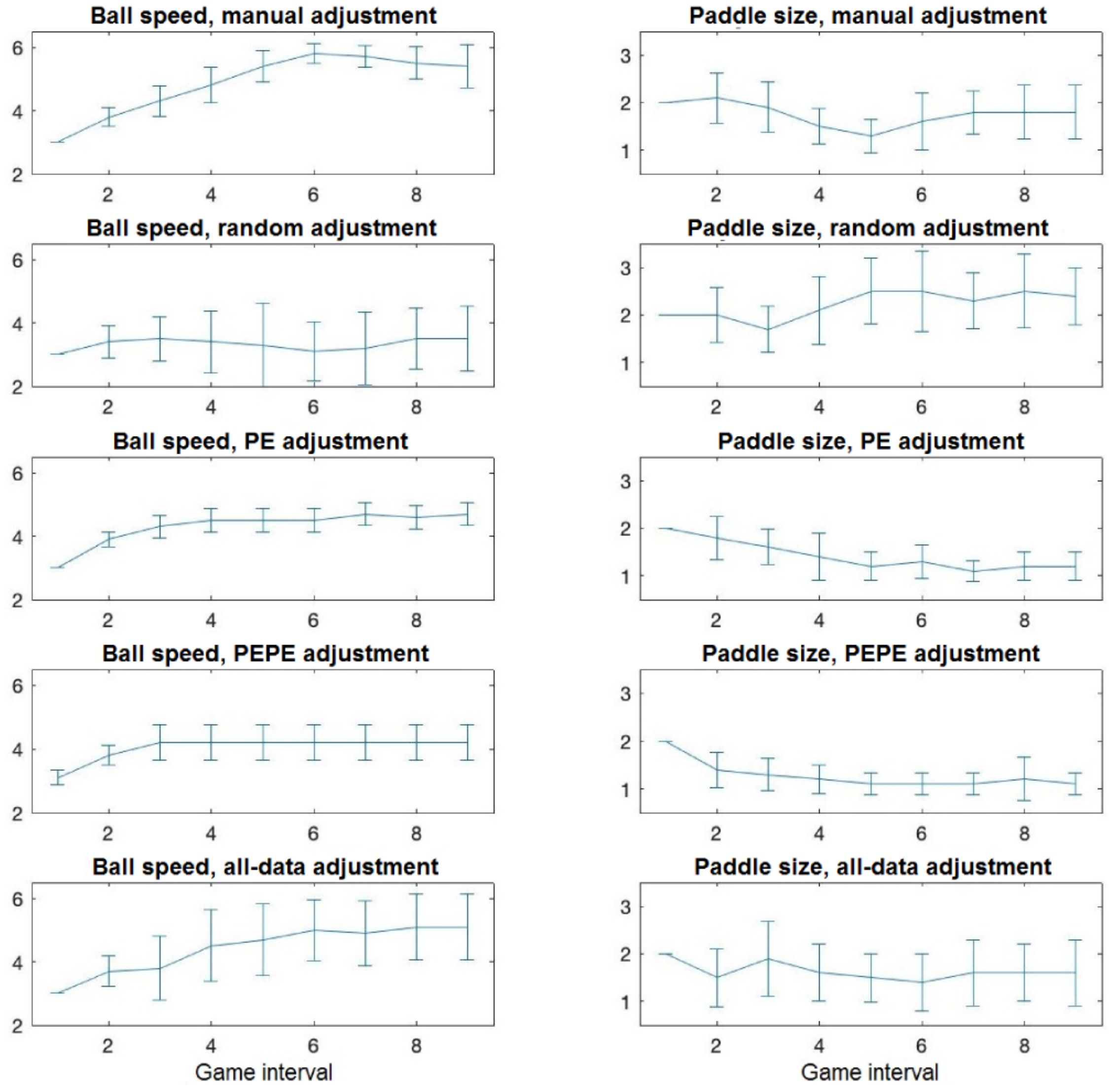


Table 3. Mean three-class classification accuracies for different combinations of input data modalities.

Input data	Outcome variable (%)			
	Difficulty	Enjoyment	Speed change	Paddle size change
PE[a]	62.4 (S[b])	50.2 (L[c])	64.6 (E[d])	55.8 (E)
PEPE[e]	62.5 (S)	50.7 (S)	62.9 (S)	59.9 (R[f])
All-data	64.2 (S)	50.7 (S)	62.6 (E)	61.7 (R)

[a]Based on performance.

[b]S: support vector machine.

[c]L: linear discriminant analysis.

[d]E: ensemble decision tree.

[e]Based on personality and performance.

[f]R: multiple linear regression.

Correlation Between Classifier Accuracy and User Experience

Finally, Table 4 presents the Spearman correlation coefficients of the outcomes of the IMI and FEM questionnaires and the accuracy of ball speed and paddle size adjustments across all DDA methods except manual (where it was always 100%).

Table 4. Spearman correlation coefficients between user experience indicators and accuracy of difficulty adjustments.

User experience aspect	Difficulty parameter			
	Ball speed		Paddle size	
	Coefficient	*P* value	Coefficient	*P* value
Effort/importance	0.27	.09	−0.14	.37
Enjoyment/interest	0.38	.01	−0.10	.51
Competence	0.16	.31	−0.03	.84
Pressure/tension	0.43	.005	−0.04	.81
Flow	0.30	.06	−0.05	.77

Discussion

Practical Implications

No differences were observed between the DDA methods. The exception was higher pressure/tension with the all-data method than with the random method, which was expected as the all-data method resulted in greater difficulty (Figure 4); similar higher pressure/tension as a result of adaptation versus no adaptation has been observed in our previous work [35]. Although we shall discuss the limitations of this study later, we first discuss the practical implications of the results.

As there was no clear benefit to the all-data method, our results do not support the use of physiological measurements for affective exergaming in practical environments. Physiological sensors are expensive, time-consuming, and inconvenient; thus, they should not be used unless they show a clear benefit. This result disagrees with the study of Xu et al [14], which found positive effects of physiology-based DDA, but it is difficult to compare the 2 studies because of significant methodological differences; for example, the Xu study had only three difficulty levels in total. At first, the result also appears to disagree with a previous study by Liu et al [8], which found positive effects of physiology-based DDA in an entertainment game. However, Liu et al [8] also did not find significant differences between physiology-based DDA and performance-based DDA based on outcome measures obtained at the end of the session; they were only able to show a difference because outcome measures were also obtained at multiple time points during gameplay itself, and those midgame outcomes were significant.

The prediction accuracy of changes to game difficulty parameters was higher with the all-data method than all other automated methods (Table 2), so the inclusion of physiological measurements did increase accuracy. For example, the accuracies for ball speed and paddle size were 62.5% and 55.0% with the all-data method compared with 53.7% and 46.2% with the PE method, respectively. The difference in accuracy may have been too low to improve the user experience meaningfully. Our previous Wizard of Oz study, for example, found that users did not reliably perceive a difference between psychological state classification accuracies that differed by less than 10% [30]. Thus, other games where adding physiological measurements increased DDA accuracy by less than 10% (including our own previous work [12] and the work of Chanel et al [9]) also may not find an improvement in user experience as a result.

As seen in Table 4, the accuracy of DDA correlated with user enjoyment/interest (r=0.38; P=.01) when the four nonmanual DDA groups were pooled together. This result agrees with our previous research [29,30] and with studies from other authors [28]. Thus, our study supports the value of more effective, personalized DDA methods; it simply does not support the premise that physiology-based DDA will be more effective than performance-based DDA.

Finally, a follow-up analysis was conducted to identify any clear patterns in when and how the DDA methods made mistakes. A clear pattern was observed: all DDA methods were more accurate at extreme difficulties (eg, very high/low paddle sizes) than at moderate difficulties. This is, in our opinion, unsurprising: when the difficulty is set to an extreme level, the correct choice is very likely to be moving away from the extreme. However, when difficulty is moderate, the correct choice is more dependent on the individual player.

Differences in Classifier Performance

Three DDA methods used classifiers that were trained on data from a previous open-loop study [33]. Studies from other fields of human-machine interaction have shown that classifiers that perform well in an offline, open-loop setting may not perform equally well in an online, real-time setting where they are used to influence machine behavior [31,32]. Thus, although our primary research focus was on user experience, we were also interested in whether the previously trained classifiers would exhibit lower accuracy in this study. Indeed, the accuracies in this study (Table 3) were lower than those in the previous one. For example, in the previous study, the all-data method resulted in a classification accuracy of 84.1% for both ball speed and paddle size preferences [33]; in this study, the accuracies were 62.5% for ball speed and 55% for paddle size. As a validation step, we reran the classification offline (post hoc) to ensure there was no bug with real-time processing; this resulted in the same results as during the data collection sessions.

We believe that the primary reason for the decrease in classification accuracy is the difference in the experienced situations. In a previous study [33], participants were exposed to three different ball speeds and three paddle sizes. In this study, participants were exposed to six different ball speeds and four paddle sizes. This is partially a weakness of our study design; however, in realistic applications, psychological state classification algorithms would inevitably be exposed to situations not seen in the training data. As a follow-up analysis, we retrained and evaluated the same types of classifiers using data from all 50 participants in this study and the 30 participants from the previous one (80 participants in total). As shown in Table 3, this also resulted in lower accuracies than those in the previous study [33], and we again believe that the primary reason for the difference was the broader range of experienced situations.

Nonetheless, as another practical implication, this result suggests that psychological state classification accuracies obtained during offline classifier training on previously recorded data may not transfer to a real-time DDA context. Thus, although a higher DDA accuracy correlates with user experience (Table 4), any exergame studies that only show differences in offline classification accuracy (including our own previous work [33]) should be taken with a grain of salt because these offline results are not guaranteed to have practical benefits when applied to DDA.

Could Other Study Designs Be More Sensitive?

As our study did not show differences in user experience among the DDA methods, we must ask whether other study designs would be more sensitive to these differences. Indeed, an intuitive follow-up study would be to test the same DDA methods in a within-subjects design, with each participant experiencing multiple DDA methods. Xu et al [14] and Liu et al [8] found differences in user experience between physiology-based and performance-based DDA, both of which used a within-subjects design, similar to our previous Wizard of Oz study [30]. In this study, we opted for a between-subjects design so that participants could experience their assigned DDA method for a longer amount of time; using multiple DDA methods would have resulted in much longer sessions. However, we acknowledge that this is not necessarily optimal.

Instead of a within-subjects design, it would also be possible to increase the sample size. The 10 participants assigned to each DDA method in this study constitute a relatively small sample size that may have been insufficient to find differences even if they did exist. We originally envisioned a larger sample size; however, data collection was carried out in early 2020 and then interrupted by the COVID-19 pandemic, forcing us to limit ourselves to the collected 50 participants.

Finally, in addition to a within-subjects design and/or a larger sample size, additional qualitative data could have been collected. For example, players could have been asked open-ended questions about their experiences. This data collection was done in the study by Liu et al [8], where participants reported lower anxiety with physiology-based DDA in open-ended questions but did not report lower anxiety on postgame forced-choice questionnaires. Although no systematic qualitative data collection was performed in this study, we mention an experimenter observation: although participants did not know which DDA method they had been assigned to, some appeared to have initial biases (eg, whether useful information can be extracted from physiology), and some attempted to determine whether their physiology influenced the system by, for example, breathing rapidly. This is in line with previous publications about participant bias and bidirectional relationships in affective computing [27,28] and may warrant further investigation.

Limitations of the Used Exergame

Finally, the results of our study do not necessarily generalize to all affective exergames. We believe that the choice of three-class classification followed by increasing or decreasing or not changing difficulty is not controversial, as it represents a classic approach to affective gaming [21]. However, our study adjusted two difficulty parameters (ball speed and paddle size) simultaneously, and the results may not be generalizable to games where only a single difficulty parameter is changed, such as in the work of Liu et al [8]. Furthermore, our study changed difficulty in relatively small steps, and our results may not generalize to games where a single DDA decision may immediately change the difficulty to a very high or very low level, such as in the work of Xu et al [14].

In addition, our study is based on the classic affective gaming approach of classifying difficulty based on prerecorded data and then adapting it at a constant frequency without user input [21]. Thus, it may not generalize to designs where DDA is adapted with a variable frequency, although such designs are not yet common in affective computing; it also may not generalize to designs where users can provide feedback about DDA decisions, potentially increasing future accuracy. For example, reinforcement learning, which remains largely unexplored in affective exergames, could be used for this purpose—users could give *rewards* for correct DDA decisions, gradually leading to a more positive user experience. Furthermore, the results may not generalize to designs that omit classification entirely and use other schemes, such as fuzzy control [5] or random forest–based regression [16].

Finally, different results might be obtained in the same scenario using different data analysis methods. For example, improved physiological feature extraction and classification may increase the accuracy of physiology-based DDA, resulting in an improved user experience in the same setting; conversely, improved performance feature extraction may also increase the accuracy of performance-based DDA. If we were to reuse the data for further research, we would likely first investigate alternatives to stepwise feature selection, which has potential issues and may not select the optimal subset of features [49]. As a follow-up analysis, we retrained and evaluated the same classifiers with two different feature reduction methods (principal component analysis and the lasso method, recommended by an article critiquing stepwise methods [49]) but did not find an improvement in classification accuracy as a result of these methods.

Conclusions

Five aspects of user experience were compared among five DDA methods of a Pong exergame: manual, random, PE, PEPE, and all-data. The last three methods adjusted the ball speed and paddle size in the game using regression-based classifiers developed in a previous (open-loop) study with 30 participants.

Though the all-data method exhibited the highest user state classification accuracy (approximately 10% higher than the PE method), no significant differences in user experience were observed among DDA methods. Thus, our results do not support the addition of physiological measurements to affective exergames; as such measurements are expensive and time-consuming, they should only be added if they meaningfully improve user experience. The study found that user experience was correlated with user state classification accuracy, thus supporting the development of more effective DDA methods, which simply does not support the notion that additional measurements will automatically improve the user experience.

The study results are limited by a somewhat suboptimal study design: each participant experienced only one DDA method, and a within-subjects design in which participants experience multiple methods may find positive effects of more complex DDA methods on user experience. Nonetheless, few studies have examined the relative effects of physiology-based DDA on user experience, and our study thus adds to the limited body of evidence on the effects of physiological measurements in affective games on user experience.

Acknowledgments

This research was supported by the National Science Foundation under grant 1717705 and by the National Institute of General Medical Sciences of the National Institutes of Health under grant 2P20GM103432. The sponsors had no role in the review and approval of the manuscript for publication.

Authors' Contributions

AD led the data collection and analysis and wrote most of the manuscript. SMM contributed to the literature review, study design, and data analysis. DN supervised the entire study, led the study design, and contributed to data analysis and manuscript writing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details about physiological measurements, questionnaires, and linear regression–based models.
[DOC File , 104 KB - games_v9i2e25771_app1.doc]

Multimedia Appendix 2
Results from individual participants in each game condition.
[XLS File (Microsoft Excel File), 965 KB - games_v9i2e25771_app2.xls]

References

1. Ko J, Jang S, Lee HT, Yun H, Kim YS. Effects of virtual reality and non-virtual reality exercises on the exercise capacity and concentration of users in a Ski exergame: comparative study. JMIR Serious Games 2020 Oct 28;8(4):e16693 [FREE Full text] [doi: 10.2196/16693] [Medline: 33112240]
2. O'Loughlin EK, Barnett TA, McGrath JJ, Consalvo M, Kakinami L. Factors associated with sustained exergaming: longitudinal investigation. JMIR Serious Games 2019 Jul 31;7(2):e13335 [FREE Full text] [doi: 10.2196/13335] [Medline: 31368440]
3. Brauner P, Ziefle M. Serious motion-based exercise games for older adults: evaluation of usability, performance, and pain mitigation. JMIR Serious Games 2020 Apr 01;8(2):e14182 [FREE Full text] [doi: 10.2196/14182] [Medline: 32234703]
4. Laver KE, George S, Thomas S, Deutsch JE, Crotty M. Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 2015 Feb 12:CD008349 [FREE Full text] [doi: 10.1002/14651858.CD008349.pub3] [Medline: 25927099]
5. Rodriguez-Guerrero C, Knaepen K, Fraile-Marinero JC, Perez-Turiel J, Gonzalez-de-Garibay V, Lefeber D. Improving challenge/skill ratio in a multimodal interface by simultaneously adapting game difficulty and haptic assistance through psychophysiological and performance feedback. Front Neurosci 2017;11:242 [FREE Full text] [doi: 10.3389/fnins.2017.00242] [Medline: 28507503]
6. Chen J. Flow in games (and everything else). Commun ACM 2007 Apr 01;50(4):31-34. [doi: 10.1145/1232743.1232769]
7. Silva MP, Silva VD, Chaimowicz L. Dynamic difficulty adjustment on MOBA games. Entertain Comput 2017 Jan;18:103-123. [doi: 10.1016/j.entcom.2016.10.002]
8. Liu C, Agrawal P, Sarkar N, Chen S. Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. Int J Hum-comput Int 2009 Aug 12;25(6):506-529. [doi: 10.1080/10447310902963944]
9. Chanel G, Rebetez C, Bétrancourt M, Pun T. Emotion assessment from physiological signals for adaptation of game difficulty. IEEE Trans Syst Man Cybern, Part A Syst Humans 2011 Nov;41(6):1052-1063. [doi: 10.1109/tsmca.2011.2116000]
10. Rani P, Sarkar N, Smith CA, Kirby LD. Anxiety detecting robotic system – towards implicit human-robot collaboration. Robotica 2004;22(1):85-95. [doi: 10.1017/s0263574703005319]

11. Ng Y, Khong C, Thwaites H. A review of affective design towards video games. Procedia Soc Behav Sci 2012;51:687-691. [doi: 10.1016/j.sbspro.2012.08.225]
12. Novak D, Mihelj M, Ziherl J, Olenšek A, Munih M. Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation. IEEE Trans Neural Syst Rehabil Eng 2011 Aug;19(4):400-410. [doi: 10.1109/TNSRE.2011.2160357] [Medline: 21708507]
13. Shirzad N, Van der Loos HF. Evaluating the user experience of exercising reaching motions with a robot that predicts desired movement difficulty. J Mot Behav 2016 May 06;48(1):31-46. [doi: 10.1080/00222895.2015.1035430] [Medline: 25945816]
14. Xu G, Gao X, Pan L, Chen S, Wang Q, Zhu B, et al. Anxiety detection and training task adaptation in robot-assisted active stroke rehabilitation. Int J Adv Robot Syst 2018 Nov 05;15(6):172988141880643. [doi: 10.1177/1729881418806433]
15. Ozkul F, Barkana DE, Masazade E. Dynamic difficulty level adjustment based on score and physiological signal feedback in the robot-assisted rehabilitation system, RehabRoby. IEEE Robot Autom Lett 2021 Apr;6(2):447-454. [doi: 10.1109/lra.2020.3046353]
16. Novak D, Beyeler B, Omlin X, Riener R. Workload estimation in physical human–robot interaction using physiological measurements. Interact Comput 2015;27(6):616-629. [doi: 10.1093/iwc/iwu021]
17. Koenig A, Novak D, Omlin X, Pulfer M, Perreault E, Zimmerli L, et al. Real-time closed-loop control of cognitive load in neurological patients during robot-assisted gait training. IEEE Trans Neural Syst Rehabil Eng 2011 Aug;19(4):453-464. [doi: 10.1109/TNSRE.2011.2160460] [Medline: 21827971]
18. Knaepen K, Marusic U, Crea S, Guerrero CD, Vitiello N, Pattyn N, et al. Psychophysiological response to cognitive workload during symmetrical, asymmetrical and dual-task walking. Hum Mov Sci 2015 Apr;40:248-263. [doi: 10.1016/j.humov.2015.01.001] [Medline: 25617994]
19. Leiker AM, Miller M, Brewer L, Nelson M, Siow M, Lohse K. The relationship between engagement and neurophysiological measures of attention in motion-controlled video games: a randomized controlled trial. JMIR Serious Games 2016 Apr 21;4(1):e4 [FREE Full text] [doi: 10.2196/games.5460] [Medline: 27103052]
20. Barathi S, Proulx M, O'Neill E, Lutteroth C. Affect recognition using psychophysiological correlates in high intensity VR exergaming. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20: CHI Conference on Human Factors in Computing Systems; April, 2020; Honolulu HI USA p. 1-15. [doi: 10.1145/3313831.3376596]
21. Novak D, Mihelj M, Munih M. A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. Interact Comput 2012 May;24(3):154-172. [doi: 10.1016/j.intcom.2012.04.003]
22. Bian D, Wade J, Swanson A, Weitlauf A, Warren Z, Sarkar N. Design of a physiology-based adaptive virtual reality driving platform for individuals with ASD. ACM Trans Access Comput 2019 Feb 23;12(1):1-24. [doi: 10.1145/3301498]
23. Bontchev B. Adaptation in affective video games: a literature review. Cybern Inf Technol 2016;16(3):3-34. [doi: 10.1515/cait-2016-0032]
24. Jerčić P, Sundstedt V. Practicing emotion-regulation through biofeedback on the decision-making performance in the context of serious games: a systematic review. Entertain Comput 2019 Mar;29:75-86. [doi: 10.1016/j.entcom.2019.01.001]
25. 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;10:223 [FREE Full text] [doi: 10.3389/fnhum.2016.00223] [Medline: 27242486]
26. Nacke L, Kalyn M, Lough C, Mandryk R. Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May, 2011; Vancouver BC Canada p. 103-112. [doi: 10.1145/1978942.1978958]
27. Fairclough SH, Lotte F. Grand challenges in neurotechnology and system neuroergonomics. Front Neuroergono 2020 Nov 30;1:602504. [doi: 10.3389/fnrgo.2020.602504]
28. Fairclough S, Karran A, Gilleade K. Classification accuracy from the perspective of the user: real-time interaction with physiological computing. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15: CHI Conference on Human Factors in Computing Systems; April, 2015; Seoul, Republic of Korea p. 3029-3038. [doi: 10.1145/2702123.2702454]
29. Novak D, Nagle A, Riener R. Linking recognition accuracy and user experience in an affective feedback loop. IEEE Trans Affect Comput 2014;5(2):168-172. [doi: 10.1109/taffc.2014.2326870]
30. McCrea S, Geršak G, Novak D. Absolute and relative user perception of classification accuracy in an affective video game. Interact Comput 2017;29(2):271-286. [doi: 10.1093/iwc/iww026]
31. Chase SM, Schwartz AB, Kass RE. Bias, optimal linear estimation, and the differences between open-loop simulation and closed-loop performance of spiking-based brain-computer interface algorithms. Neural Netw 2009 Nov;22(9):1203-1213 [FREE Full text] [doi: 10.1016/j.neunet.2009.05.005] [Medline: 19502004]
32. Hargrove LJ, Scheme EJ, Englehart KB, Hudgins BS. Multiple binary classifications via linear discriminant analysis for improved controllability of a powered prosthesis. IEEE Trans Neural Syst Rehabil Eng 2010 Feb;18(1):49-57. [doi: 10.1109/TNSRE.2009.2039590] [Medline: 20071277]

33. Darzi A, Wondra T, McCrea S, Novak D. Classification of multiple psychological dimensions in computer game players using physiology, performance, and personality characteristics. Front Neurosci 2019;13:1278 [FREE Full text] [doi: 10.3389/fnins.2019.01278] [Medline: 31849589]
34. Nagle A, Wolf P, Riener R. Towards a system of customized video game mechanics based on player personality: relating the Big Five personality traits with difficulty adaptation in a first-person shooter game. Entertain Comput 2016 Mar;13:10-24. [doi: 10.1016/j.entcom.2016.01.002]
35. Goršič M, M A, Novak D. Comparison of two difficulty adaptation strategies for competitive arm rehabilitation exercises. In: Proceedings of the International Conference on Rehabilitation Robotics (ICORR). 2017 Presented at: International Conference on Rehabilitation Robotics (ICORR); July 17-20, 2017; London, UK p. 640-645. [doi: 10.1109/icorr.2017.8009320]
36. Gorsic M, Clapp JD, Darzi A, Novak D. A brief measure of interpersonal interaction for 2-player serious games: questionnaire validation. JMIR Serious Games 2019 Jul 23;7(3):e12788 [FREE Full text] [doi: 10.2196/12788] [Medline: 31339107]
37. Klem GH, Lüders HO, Jasper HH, Elger C. The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. Electroencephalogr Clin Neurophysiol Suppl 1999;52:3-6. [Medline: 10590970]
38. Fitzgibbon S, DeLosAngeles D, Lewis T, Powers D, Grummett T, Whitham E, et al. Automatic determination of EMG-contaminated components and validation of independent component analysis using EEG during pharmacologic paralysis. Clin Neurophysiol 2016 Mar;127(3):1781-1793. [doi: 10.1016/j.clinph.2015.12.009] [Medline: 26780994]
39. Azami H, Rostaghi M, Abásolo D, Escudero J. Refined composite multiscale dispersion entropy and its application to biomedical signals. IEEE Trans Biomed Eng 2017 Dec;64(12):2872-2879. [doi: 10.1109/tbme.2017.2679136]
40. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. Eur Heart J 1996 Mar 01;17(3):354-381. [doi: 10.1093/oxfordjournals.eurheartj.a014868]
41. Boucsein W. Electrodermal Activity. Switzerland: Springer; 2012:1-618.
42. Kim TT, Lee G. Hospitality employee knowledge-sharing behaviors in the relationship between goal orientations and service innovative behavior. Int J Hosp Manag 2013 Sep;34:324-337. [doi: 10.1016/j.ijhm.2013.04.009]
43. Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. J Pers Soc Psychol 1994;67(2):319-333. [doi: 10.1037/0022-3514.67.2.319]
44. Hsia L, Huang I, Hwang G. Effects of different online peer-feedback approaches on students' performance skills, motivation and self-efficacy in a dance course. Comput Educ 2016 May;96:55-71. [doi: 10.1016/j.compedu.2016.02.004]
45. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. J Res Pers 2003 Dec;37(6):504-528. [doi: 10.1016/S0092-6566(03)00046-1]
46. Quinn GP, Keough MJ. Multiple and complex regression. In: Experimental Design and Data Analysis for Biologists. Cambridge, UK: Cambridge University Press; 2002:111-154.
47. Markland D, Hardy L. On the factorial and construct validity of the Intrinsic Motivation Inventory: conceptual and operational concerns. Res Q Exerc Sport 1997 Mar;68(1):20-32. [doi: 10.1080/02701367.1997.10608863] [Medline: 9094760]
48. Sung H, Hwang G, Yen Y. Development of a contextual decision-making game for improving students' learning performance in a health education course. Comput Educ 2015 Mar;82:179-190. [doi: 10.1016/j.compedu.2014.11.012]
49. Flom P, Cassell D. Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. In: Proceedings of the NESUG Conference. 2007 Presented at: NESUG 2007 Conference; November 11-14, 2007; Baltimore, Maryland URL: https://www.lexjansen.com/nesug/nesug07/sa/sa07.pdf

Abbreviations

DDA: dynamic difficulty adjustment
ECG: electrocardiogram
EEG: electroencephalography
FEM: Flow Experience Measure
IMI: Intrinsic Motivation Inventory
PE: performance based
PEPE: personality-performance based

Edited by N Zary; submitted 14.11.20; peer-reviewed by A Teles, P Brauner; comments to author 08.01.21; revised version received 23.03.21; accepted 16.04.21; published 31.05.21.

Please cite as:
Darzi A, McCrea SM, Novak D
User Experience With Dynamic Difficulty Adjustment Methods for an Affective Exergame: Comparative Laboratory-Based Study
JMIR Serious Games 2021;9(2):e25771
URL: https://games.jmir.org/2021/2/e25771
doi:10.2196/25771
PMID:34057423

Original Paper

Developing Adaptive Serious Games for Children With Specific Learning Difficulties: A Two-phase Usability and Technology Acceptance Study

Oguzcan Yildirim[1], MSc; Elif Surer[1], PhD

[1]Department of Modeling and Simulation, Graduate School of Informatics, Middle East Technical University, Ankara, Turkey

Corresponding Author:
Elif Surer, PhD
Department of Modeling and Simulation
Graduate School of Informatics
Middle East Technical University
Universiteler Mahallesi
Dumlupınar Bulvarı, No:1
Ankara, 06800
Turkey
Phone: 90 5452107888
Email: elifs@metu.edu.tr

Abstract

Background: Specific learning difficulties (SpLD) include several disorders such as dyslexia, dyscalculia, and dysgraphia, and the children with these SpLD receive special education. However, the studies and the educational material so far focus mainly on one specific disorder.

Objective: This study's primary goal is to develop comprehensive training material for different types of SpLD, with five serious games addressing different aspects of the SpLD. The second focus is measuring the impact of adaptive difficulty level adjustment in the children's and their educators' usability and technology acceptance perception. Receiving feedback from the children and their educators, and refining the games according to their suggestions have also been essential in this two-phase study.

Methods: A total of 10 SpLD educators and 23 children with different types of SpLD tested the prototypes of the five serious games (ie, Word game, Memory game, Category game, Space game, and Math game), gave detailed feedback, answered the System Usability Scale and Technology Acceptance Model (TAM) questionnaires, and applied think-aloud protocols during game play.

Results: The games' standard and adaptive versions were analyzed in terms of average playtime and the number of false answers. Detailed analyses of the interviews, with word clouds and player performances, were also provided. The TAM questionnaires' average and mean values and box plots of each data acquisition session for the children and the educators were also reported via System Usability Scale and TAM questionnaires. The TAM results of the educators had an average of 8.41 (SD 0.87) out of 10 in the first interview and an average of 8.71 (SD 0.64) out of 10 in the second interview. The children had an average of 9.07 (SD 0.56) out of 10 in the first interview.

Conclusions: Both the educators and the children with SpLD enjoyed playing the games, gave positive feedback, and suggested new ways for improvement. The results showed that these games provide thorough training material for different types of SpLD with personalized and tailored difficulty systems. The final version of the proposed games will become a part of the special education centers' supplementary curriculum and training materials, making new enhancements and improvements possible in the future.

(JMIR Serious Games 2021;9(2):e25997) doi:10.2196/25997

KEYWORDS

serious games; adaptive games; specific learning difficulty; usability; system usability scale; technology acceptance model; training; development; adaptation; gaming; learning disability; children; education; teacher

Introduction

Special education is a form of education that is customized to the needs of children who have a disability, illness, or difficulty. In the United States, special education is offered in public schools, and this right is secured by the Individuals with Disabilities Education Act (IDEA) [1]. IDEA categorizes disabilities and conditions demanding special education into 13 types, including autism, attention-deficit/hyperactivity disorder, and specific learning difficulty (SpLD). A specific learning impairment (also known as a specific learning disability or SpLD) is a condition in which children learn specific concepts at a slower rate than would be expected for their age or educational level. Learning disability (or learning disorder; LD) is not associated with SpLD, although it is a medical term used for diagnosis and is often referred to as LD. To avoid any misunderstanding, the word SpLD will be used in this study to refer to specific learning difficulties. SpLD is not linked to intelligence, which means that children with SpLD may have average or superior intelligence but learn slower than their peers. SpLD is divided into five groups by the Learning Disabilities Association of America as dyslexia (reading disability), dyscalculia (math disability), dysgraphia (writing disability), oral and written language disorder and specific reading comprehension deficit, and nonverbal learning disability.

Dyslexia is the most well-known form of SpLD, and it causes spelling mistakes, poor reading, and false reading in children. Children with dyscalculia have difficulties with numbers and abstract principles such as counting and basic numerical operations. Misspelling, writing letters in wrong forms, and writing letters of various sizes in a word are all symptoms of dysgraphia in children. Oral and written language disorder and specific reading comprehension deficit affect children's ability to articulate and comprehend sentences. Dyspraxia can show itself in various behaviors, from a lack of coordination between the lips and tongue to problems with handwriting. There is no agreement on the types of SpLD, and the coverage of SpLD changes in every country. Since the proposed games in this study are intended for the Turkish training system and given that these SpLD types are accepted as SpLD in Turkey [2] and are suggested by educators in special education centers with whom the authors collaborated, they are accepted as SpLD types in this paper. In Turkey, 41,600 people have been diagnosed with SpLD, and 82% of dyslexic students drop out of school before attending university. To date, the only verified therapy for SpLD is special education. In addition to the regular school education, children with SpLD attend extra classes taught by special education teachers in groups or individually. This special education can include reading, writing, algebra, developing motor coordination skills, and language education activities.

The impact of games on children's training with SpLD has been discovered recently, and the number of studies on this topic has increased. Investigating the influence of daily video game playing on children with SpLD was the initial research focus in SpLD research [3]. In a study [4], researchers collaborated with 20 children with dyslexia to see if video games, Rayman Raving Rabbids, a promotional Wii video game, influenced their reading abilities. In another study [5], an app, which included six mini-games, was created to diagnose dyslexia in children aged 6 to 7 years. Each of the games in the app tested a different ability such as word forming, syllabic memory, verbal work memory, auditory memory, and word reading. Teachers found the app helpful in identifying children with potential dyslexia and correcting exercises. Nonetheless, the app was language-dependent, and children should have been able to read and write. Another app [6] for screening dyslexia in children aged 7 to 12 years was also developed to solve the language dependency. There are also findings in the SpLD literature for dyscalculia. A serious virtual reality game was created to help children with dyscalculia learn math [7]. Three separate scenarios and game plays were included in the game, each with three different difficulty levels. The participants in the study were 40 students aged 7 to 9 years. The suggested game was played by half of the children at random, while the other half, the control group, was handled with the conventional Domino system. Children who played the proposed game took substantially less time to fit simple mathematical operations to the correct answer when compared with the control group.

In a study performed in Germany [8], a mobile game–based intervention on syllable stress and literacy was developed for German children with dyslexia. The study focused mainly on the design criteria of the game-based intervention, while the quantitative analysis was planned to be performed on children by using the mobile game 20 minutes a day, 5 days per week. Similarly, Sood et al [9] also focused on developing games where the objectives were detecting, monitoring, and managing dyslexia in young children (aged 4-18 years). The study presented the details of the protocols and game design principles, but the games were not tested with the control group yet. Flogie et al [10] developed intelligent game interventions and used them with 51 children with learning difficulties in the mainstream Slovenian education system. The results showed that intelligent game interventions provide personalized education and can be helpful while designing the specialized curriculum.

This study is adapted from author OY's Master's thesis [11], supervised by author ES. This study concentrates on serious games' usability and technology acceptance results in the training of children with SpLD. As previously mentioned, serious games, or games with more than just entertainment value [12,13], have recently become an important research topic among researchers interested in children's education with SpLD. To the best of the authors' knowledge, this is the first systematic research covering all types of SpLD and measuring the impact of using adaptive difficulty via quantitative and qualitative approaches. This research's primary aim is to create a collection of specially developed games for use in special education for children with any form of SpLD. Usability tests were used to investigate the potential of this series of games. Another aim is to demonstrate how the adaptive difficulty system affects children's playing experience. An adaptive difficulty framework was developed to assess its impact on children with SpLD. To that end, five serious games for children with SpLD were designed and developed. Two of the five games were improved with an adaptive difficulty system to increase the children's experience when playing the games. Children with SpLD and their educators evaluated all five games, and their responses to questionnaires—System Usability Scale (SUS) [14] and Technology Acceptance Model (TAM) [15]—and their comments on the games were classified and analyzed. The results demonstrated the games' usability as enhanced training material in special education. Both the children with SpLD and their educators gave highly positive feedback regarding playing the established games, which was also repeated on the responses to the SUS and TAM questionnaires. This study was split into two parts to implement the children's and educators' recommendations into the framework and thoroughly analyze the outcomes.

Methods

Overview

For this study, five different serious games were designed and developed to train children with any type of SpLD. The details of the games and their versions are explained in detail in the following subsections. The games were developed in Unity (version 2018.3.0f2), a cross-platform game engine that allows its users to build games in 2D and 3D environments. For this study, an ASUS Zenpad 8 tablet is used during the data acquisition so that Unity's Android device settings were used during the game development procedure. After the game development process, the games were tested twice by two participant groups, including 10 educators and 23 students with SpLD between the ages of 7 and 11 years. During the first interview with the participant groups, the SUS and TAM questionnaires were filled by the participants besides applying the think-aloud methods during game play and answering open-ended questions after playing the games. After the first interviews, a rule-based adaptive difficulty enhancement was added to two of the five games to measure the impact of adaptiveness in the developed serious games. During the second interview with the children, only open-ended questions and think-aloud methods were used. For the educators, the same interview procedure was used as before. Multimedia Appendix 1 summarizes the prototyping and interview procedures of the study.

Design Criteria and Themes of the Games

At first, with special educators' help, design criteria were defined to develop games compatible with the children's educational background. The first criterion is about the games' target groups, including children with SpLD who know reading and writing. Therefore, children's age was limited to 11 years since designing a decent game for both new readers and secondary school students could be an unfeasible goal. The second criterion is that games should be easy-to-use given the children's struggle with SpLD. Thus, games do not include multitasking such as picking up stars in the sky while solving multiplication problems on the ground.

Given that children with SpLD will play the games, the themes of the games were decided cautiously. Games do not focus on a single type of SpLD but cover various types of SpLD. The first game's theme was the spelling for children with dyslexia or dysgraphia. Simple mathematical operations were the second game's theme for children with dyscalculia. The third game's theme was writing examples of specific categories developed for children with dysgraphia or dyslexia. Two more themes were added for different purposes (ie, a memory game targeted at improving visual memory and a space game requiring time management and constant concentration). When choosing themes, the educators showed what kind of practices they do in their special education and suggested game ideas, so the opinions of educators were also taken into consideration during the game design. The children had only 30 minutes per week, and they had to play all the games and answer the study questions in 2 weeks, so the number of themes was limited due to the time limitations of the children.

Developed Games

During this study, different versions of the five games were developed. The first version of the games were developed as prototypes, but all the themes and rules were already well-defined. Educators examined the games' prototype versions so that the educators' first interview could be considered an initial quality check (Prototype I). Since the users requested no modifications after Prototype III, only two prototypes were developed in the Word game. The Category game and Math game had three prototypes in total for the same reason (ie, no updates were requested by the users). Four different prototypes were developed for the Memory game and Space game.

Word Game: Prototype I

The word game (Figure 1) is focusing on spelling for children with dyslexia or dysgraphia. In the game, a child character asks 10 questions about a single predefined topic, and the answers are limited to 4 to 6 letters in Turkish. The player needs to drag and drop the letters to the appropriate yellow squares, representing a letter in the answer. If the letter is correct, it replaces the yellow square, and if it is a false letter, the letter does not fill the yellow square and returns to its initial position (bottom left corner). After entering all the correct letters, the child avatar confirms "It is correct!" and asks the next question. The game's background image is chosen to keep the players focused on the topic. For instance, "Kitchen" was the topic of the first designed level, and the players were asked to answer questions about cutlery or fruits, such as "What is the name of the sweet fruit which has red, yellow, and green colors?" and "What do you use to eat your soup?" The player gets 10 points for each correct answer.

Figure 1. Screenshot of the Word game from Prototype I.

Word Game: Prototype II

A skip button was added to the screen so that unsolved questions could be passed. The number of correct answers, the number of false letters, the number of skipped questions, and the total time were recorded in the database.

Word Game: Prototype III

Given that the users did not request any new updates after Prototype II, this game was not modified.

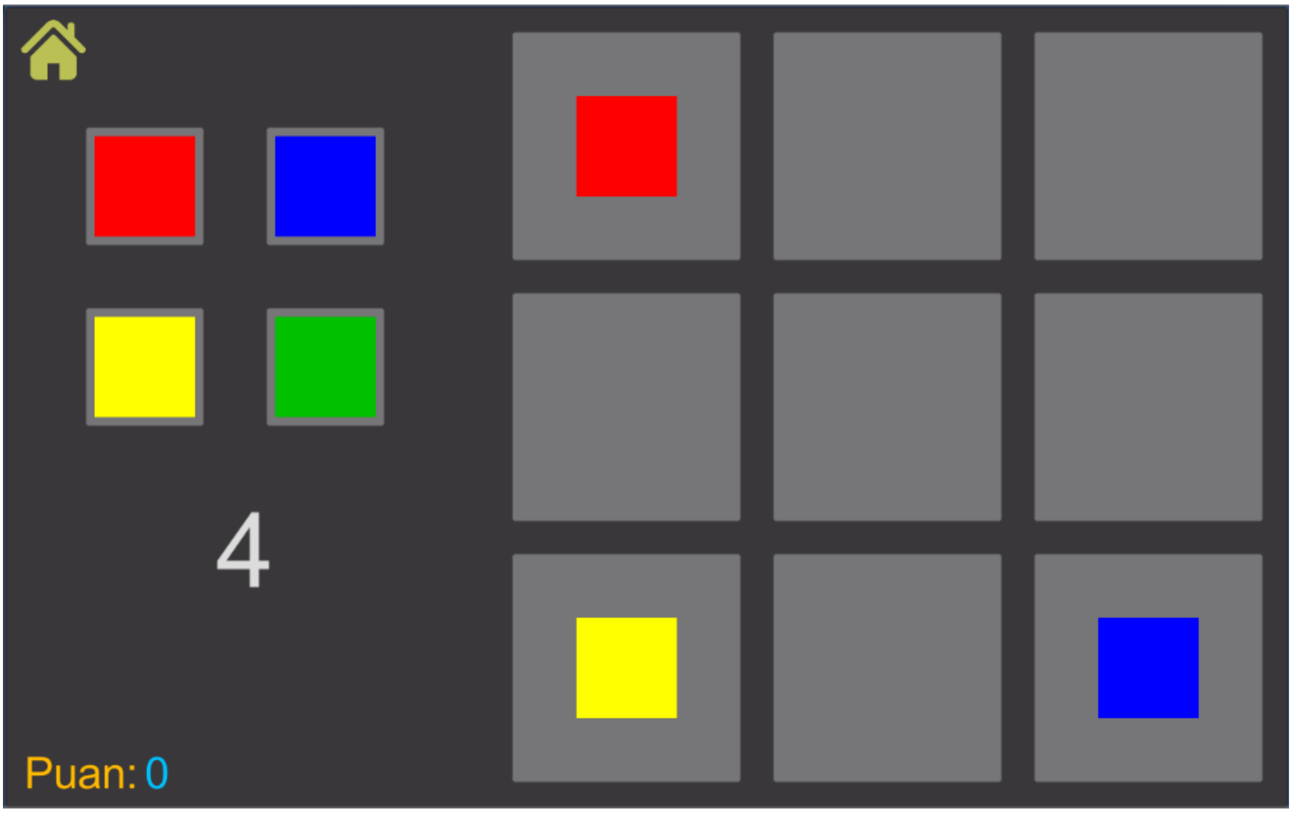
Word Game: Prototype IV

Given that the users did not request any new updates after Prototype II, this game was not modified.

Memory Game: Prototype I

The Memory game (Figure 2) was created to improve the visual memory and attention of children with SpLD. In the game, there are two panels; on the left side of the screen, there are four colors (or four arrows each faced in another direction) and a counter. Nine boxes are aligned as a three-by-three grid on the right side of the screen. When the game starts, initial colors appear on three of the nine boxes randomly, and the counter starts to go down. After the counter hits zero, from five, colors on the right side disappear. The player's goal is to refill the three boxes with the correct colors as quickly as possible. The player first needs to tap the correct color on the left panel and tap the right panel's accurate grid. To erase the grids, the player can tap on a filled grid. Whenever the player fills all three grids correctly, "CORRECT!" feedback shows up at the counter's position. After 3 seconds, a new randomly generated memory question is asked, where each correct answer is 10 points.

Figure 2. Screenshot of the Memory game from Prototype I.



Memory Game: Prototype II

A hint button was added to the screen to show the colors or arrows once again to the players. Additionally, a timer was added to display the countdown. Both the difficulty and timer can be set just before starting the game by the educators or parents. The difficulty setting affects the number of grids to be remembered during the game. The number of correct answers, number of chosen false grids, number of grids filled with wrong color or direction, difficulty level, and session time were recorded in the Memory game's database.

Memory Game: Prototype III

After analyzing the first two prototypes, two games—the Memory game and Space game—were updated and enhanced with the adaptive difficulty level. The games' design or mechanics were not modified. The purpose of the second interview with the students was to determine their behavior and success in games with adaptive difficulty levels.

Memory Game: Prototype IV

To represent the alternative versions of this game, a version with pictures was developed (Multimedia Appendix 2). In this version, sky objects were used, such as sun, moon, cloud, and lightning. Moreover, the appearing time can be set as 3, 4, or 5 seconds before the game starts, increasing the difficulty. Thus, this game has four versions (colors, arrows, letters, and pictures), five difficulty options—one grid, two grids, three grids, four grids, and adaptive (two grids with variable appearing time)—with three different appearing time options (3, 4, and 5 seconds), and an adjustable timer.

Category Game: Prototype I

The Category game was developed to improve the thinking and writing skills of children with dysgraphia or dyslexia—not handwriting but correctly ordering letters in a word. During the game, a couple of categories are given, and the players are expected to find words related to that category. For example, in Figure 3, the category is "Things that we close," and predefined words are "Window," "Door," and "Box." In each category, there are 15 to 20 predefined answers. If the player writes down one of them, it appears on the screen with a random color and position, and the player gains 10 points. If the answer is wrong or misspelled, it does not appear on the screen. The player uses the keyboard of the tablet to write the words. If the player gets stuck at some point, the skip button on the top right corner can be used, and a new category starts. There were 10 different categories in the first version of the game.

Figure 3. Screenshot of the Category game from Prototype I.

Category Game: Prototype II

The category game was separated into two parts. In the first part (Multimedia Appendix 3), the player must choose five correct words about a given category, between 20 available options. After the fifth correct answer, the category changes. When five categories are completed, the second part, which is the same as the previous Category games, starts. The player gives examples of the same categories in the first part. Moreover, a skip button is added to the screen. Session time and the number of correct and false answers were saved in the database. For the second part (Multimedia Appendix 4), the number of correct answers, number of typos, number of false words, and session time were recorded.

Category Game: Prototype III

In this prototype (Multimedia Appendix 5), the input system of the open-ended questions was modified. The player does not use the keyboard of the tablet keyboard anymore but uses an on-screen keyboard aligned in alphabetical order.

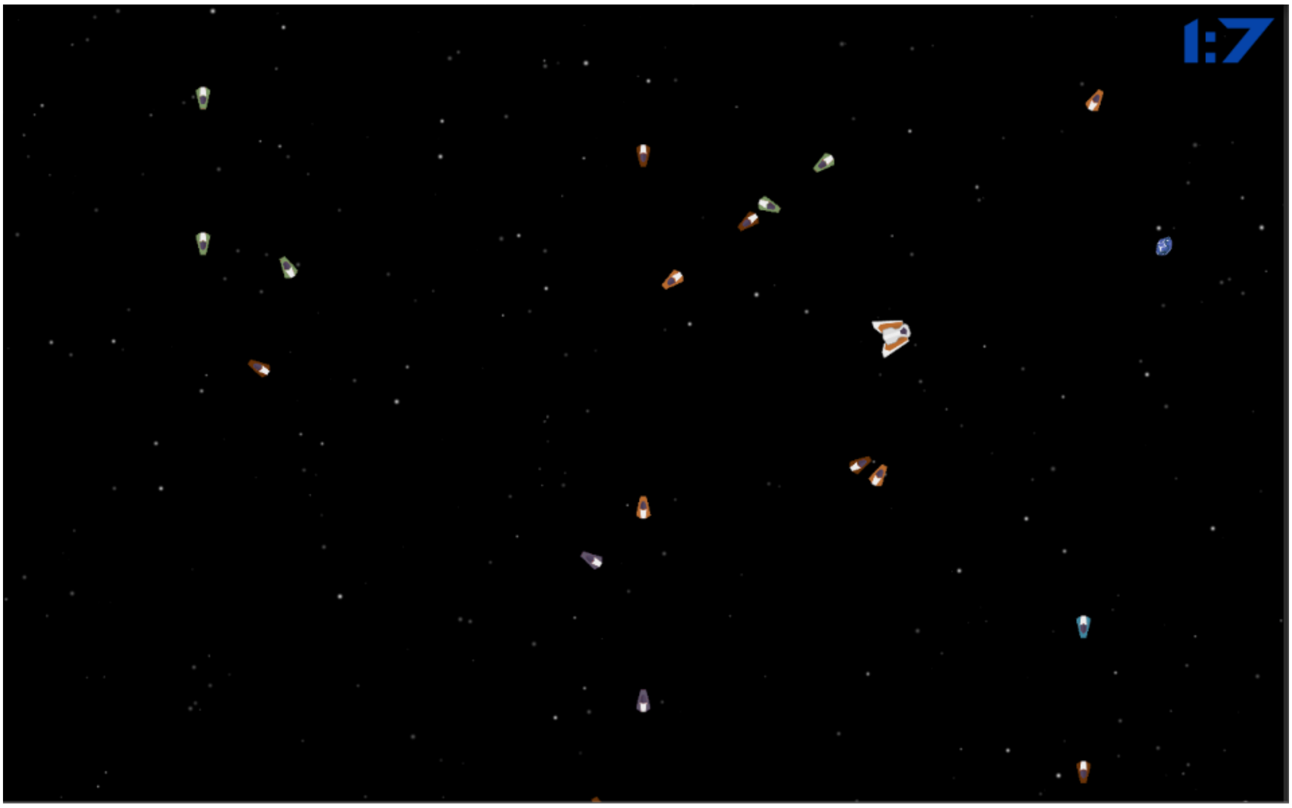
Category Game: Prototype IV

Given that the users did not request any new updates after Prototype III, this game was not modified.

Space Game: Prototype I

In the Space game (Figure 4), the primary objective is to test the time management and concentration levels of children with a theme that is not causally related to their courses. The player's primary goal is collecting gems while trying not to hit tiny spaceships that are programmed to go in a random direction. The initially available time slot is 90 seconds, and the timer counts down. Whenever the player taps into the screen, a spaceship rotates into that point and moves toward the point until it reaches the tapped position. A gem appears every 5 seconds at a random location. If the player cannot collect the gem in 10 seconds, the gem disappears. Every 0.4 seconds, an enemy spaceship spawns off-screen and flies straight through the screen in a random direction. The player gains 2, 4, or 6 seconds randomly after collecting a gem. Whenever a player collides with another spaceship, the player loses 3, 5, or 7 seconds randomly. The game is over when the time is up.

Figure 4. Screenshot of the Space game from Prototype I.



Space Game: Prototype II

No modification was applied in the Space game except forming a database for game statistics. The number of collected gems, the number of crashes with the enemy, and session time were adjusted to be kept in the database.

Space Game: Prototype III

After analyzing the first two prototypes, two games—Memory game and Space game— were updated and enhanced with the adaptive difficulty level. The games' design or mechanics were not modified. The purpose of the second interview with the students was to determine their behavior and success in games with adaptive difficulty levels.

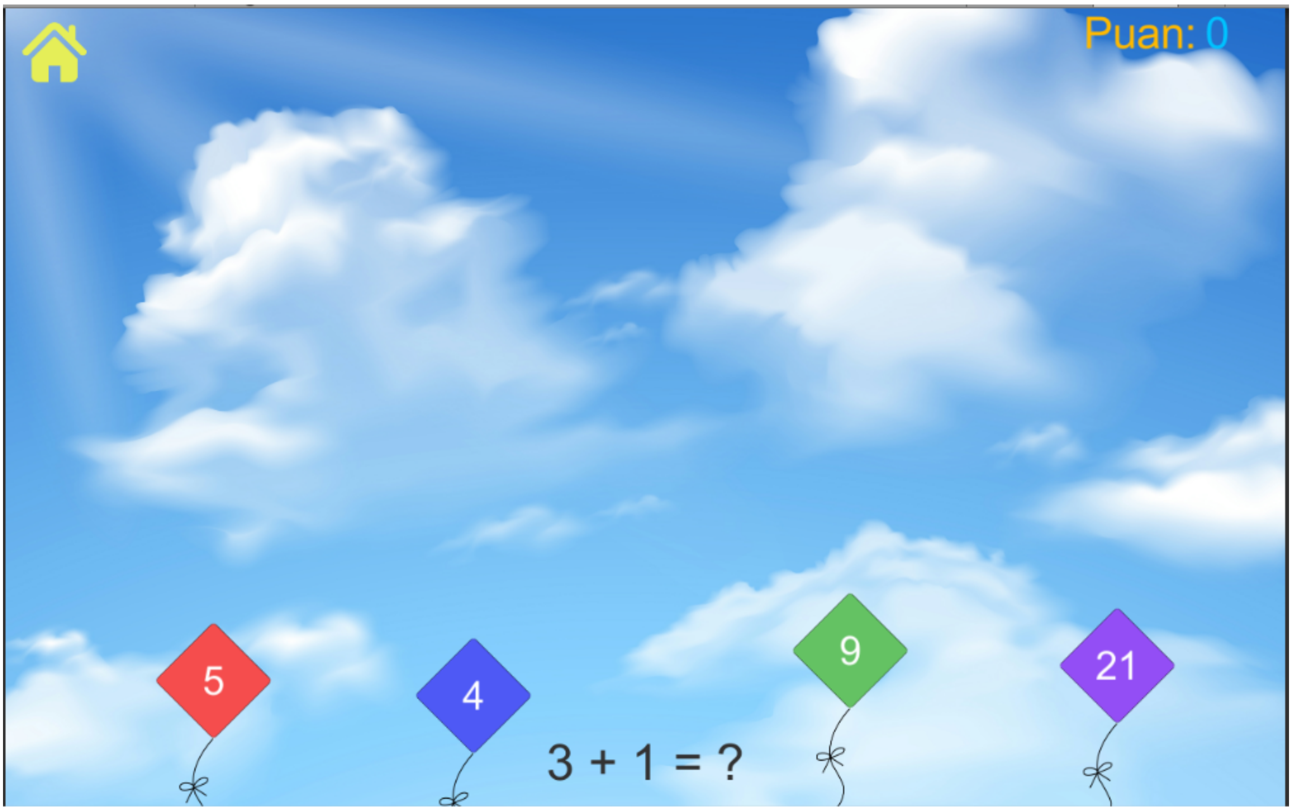
Space Game: Prototype IV

All the objects in the scene, except the user interface, were enlarged. The control mechanism was changed with a joystick, which was added to the screen's bottom right. Additionally, the default enemy spawn frequency was decreased. The waiting time between two enemies' spawn was changed to 0.7 seconds. Moreover, visual feedback on collecting gems and crashing into enemies was added. Players could also see how much time they gained (in blue) or lost (in red) during the game.

Math Game: Prototype I

The Math game (Figure 5) is a simple addition and subtraction game. In the first version, the math question appeared below the screen, and four kites floated up in the air. Questions are created randomly with a set of selected numbers between 0 to 15. The player's primary goal is tapping on the kite containing the correct answer before exiting from the screen. There was no penalty for choosing wrong answers to see the tapping strategy of the children. They could tap all four to see the next question and still get total points.

Figure 5. Screenshot of the Math game from Prototype I.



Math Game: Prototype II

A timer was added as a countdown, and both the timer and difficulty could be set before the game starts. The difficulty level modified the kites' speed, and a maximum limit of numbers for the math operations could also be arranged. The number of correct and false solutions, unanswered questions, session time, and difficulty level were kept in the database.

Math Game: Prototype III

As an option, the multiplication operation was added to the game, and the scoring system was modified so that the player does not gain total points after their first guess. After each guess, the scoring system changes (ie, in the second guess, 8 points; in the third guess, 4 points; and a the last guess, only 1 point is given to the player).

Math Game: Prototype IV

Given that the users did not request any new updates after Prototype III, this game was not modified.

Adaptive Difficulty Mechanism

The adaptive difficulty is designed to check consecutive successful and unsuccessful moves by the player and adjust the game difficulty according to the player's game play level. A simple adaptive difficulty system was developed for the Memory game and Space game. The number of games was decreased to two since these games would be played by the students twice, and the students had time constraints. Besides, these two games had repeatability, and they were liked by children the most. This adaptive difficulty was added to these games to keep the players in the *golden path* or *optimal game play corridor*, which prevents the players from getting bored or frustrated by the difficulty of the game, as it is introduced in the Flow study of Czikszentmihalyi [16] and enhanced by Thomas and Young [17] (Multimedia Appendix 6). For the adaptive difficulty, a rule-based method was selected due to the time constraints of the students. They could only play both games twice, once each for the standard and adaptive versions. Since repetitive testing is no longer an option, methods that use training could not be used. The rule-based difficulty method was chosen to modify the difficulty for a couple of actions instead of every action. These two alternatives would give the same results for large sample sizes in theory; however, the rule-based method was chosen to ensure that every student gets the same results for their successful or unsuccessful set of actions. In addition, since there will be only one game play session per student, the adaptive difficulty should affect the game while the game is continuing, unlike the study of Hendrix et al [18] where the players' previous game sessions changed the game difficulty.

The system was designed to change only one variable to get meaningful results from the adaptive difficulty levels. For the Memory game, the correct answer's display time was the selected variable, while for the Space game, enemy spawn frequency was selected. These variables affect games directly, but they cannot be noticed by the player explicitly, unlike the number of letters in the Memory game or the type of enemies in the Space game, so that the students can give their honest opinion about how they feel about the game difficulty without any a priori clue. The flow of adaptive difficulty change is shown in Multimedia Appendix 7. Since the students could play these games only for a limited time, a single change in the variables was adjusted by around 10%. The difference in these variables could affect the outcome quickly after the game gets easier or more difficult.

Memory Game With Adaptive Difficulty System

For all the students with dyslexia or dysgraphia, a letter version of the Memory game was developed. The Memory game's letter version (Multimedia Appendix 8) contains only four consonants: *p*, *q*, *b*, and *d*. These are the most commonly confused letters for children with dyslexia when reading and writing. Although the Turkish alphabet does not contain the letter *q*, children are familiar with this letter from the keyboard. The adaptive system modifies the first appearing time of the letters on the screen, and the timer counts down seemingly from five. However, it counts faster or slower according to the success of the players in the game. The rules of the adaptive difficulty are simple, as they are explained in Figure 6 in pseudocode. If the player can give two correct answers (C_{CCA}) in a row, the appearing time (t_{LA}) decreases by half a second. If the player provides more than one false solution, false grid, or incorrect letter in a row (C_{CFA} is used for all), the appearing time increases by half a second (A_g is used for *given answer*). Changes in the emerging time are applied to the next question.

Figure 6. Adaptive difficulty memory.

```
procedure ADAPTIVEDIFFICULTYMEMORY (Ag, timer)
  CCCA ← 0
  CCFA ← 0
  tLA ← 5
  while timer != 0 do
    if Ag == true then
      CCCA ← CCCA + 1
      CCFA ← 0
      if CCCA == 2 then
        tLA ← tLA - 0.5
        CCCA ← 0
      end if
    else
      CCFA ← CCFA + 1
      CCCA ← 0
      if CCFA == 2 then
        tLA ← tLA + 0.5
        CCCA ← 0
      end if
    end if
  end while
end procedure
```

The first interview results show that students gave an average of 4.45 correct answers and 2.35 false answers for two grid difficulties in 60 seconds, which led to the selection of two consecutive correct or false answers as the adaptive difficulty changes condition. According to the results, the expected average difficulty change was more than two for the increment and more than one for the decrement during 100 seconds of playtime. Due to these initial outcomes, the amount of the change was set as 0.5 seconds so that neither students can realize the difference directly nor face substantial difficulty changes. The number of correct and false answers before a change in appearing time and the number of changes in appearing time were added to the database.

Space Game With Adaptive Difficulty System

Adaptive difficulty in the Space game only affects the enemy spaceships' spawn frequency (Figure 7). If the player is successful, the frequency increases, and vice versa. The default waiting time (T_{ES}) between the spawn of two enemy spaceships is 0.4 seconds. The waiting time increases by 0.03 seconds when the number of crashes (C_{CE}) is three more than the number of collected gems (C_{CG}).

Waiting time decreases by 0.03 seconds when the number of collected gems is two more than the number of crashes. Both counters reset when the spawning frequency changes (E_{CE} is used for the crashing enemy event, and E_{CG} is used for the collect gem event). According to the first interview results, students collected gems an average of 0.997 times. They crashed the enemies an average of 1.317 times per 10 seconds, and their average playtime was 94.06 seconds. Due to these results, the number of events (crashing enemies or collecting gems) is stored as one variable. Instead of the consecutive events, to change the difficulty, the total number of crashes or collected gems is expected to be more than the other. To decrease the enemy spawn period, the number of collisions is expected to be three more than collected gems instead of two since the students were more likely to crash into an enemy rather than collect a gem. The change amount of the enemy spawn period was selected as 0.03 seconds since the Space game is a more dynamic game that requires more attention than the Memory game. The change was kept below 10% to avoid a profound difference.

Figure 7. Adaptive difficulty space.

```
procedure ADAPTIVEDIFFICULTYSPACE (E_CE, E_CG
timer)
    C_CE ← 0
    C_CG ← 0
    T_ES ← 0.4
    while timer != 0 do
      if E_CE == true then
        C_CE ← C_CE + 1
        if C_CE == (C_CG + 3) then
          T_ES ← T_ES + 0.03
          C_CE ← 0
          C_CG ← 0
        end if
      else if E_CG == true then
        C_CG ← C_CG + 1
        if C_CG == (C_CE + 2) then
          T_ES ← T_ES - 0.03
          C_CE ← 0
          C_CG ← 0
        end if
      end if
    end while
end procedure
```

Participants

In this study, 10 educators participated in the study—7 of them were special education teachers, 2 of them were occupational therapists, and 1 was a psychologist. The mean age of educators was 33 (SD 14.37) years, and 7 of the educators were female. Children participants were selected among the Albatros Special Education Center (Albatros Özel Eğitim Merkezi in Turkish), which provides education for children with SpLD. The final student group includes 23 students in the age group of 7 to 11 years. Before the data acquisition started, permission of their parents was taken both verbally and written. Additionally, children's consent was taken verbally before the acquisitions, and the voluntary nature of the study was emphasized. Only 3 children could not complete the study since they had changed their school before the end of the study. The mean age of the 23 children participants was 8.6 (SD 1.13) years. Most of the children were male (18/23, 80%). A total of 20 children were diagnosed with the SpLD by a doctor. The rest of the participant children had taken the Albatros Special Education Center's test. Given that the special education center's psychologists evaluated them as having SpLD, they were also included in the study.

Permissions and Ethics Approval

This study was performed by two participant groups—students and their educators. This study was not designed as a clinical trial but as a two-phase usability study for the Master's thesis of OY, supervised by ES. The Ethical Approval of Research was approved by the Middle East Technical University Human Subjects Ethics Committee (ID of the permission: 28620816/398). The participants or their parents, if they were children, were informed both with verbal and written communication about the purpose of the study, possible benefits of the study to children with SpLD, and the voluntary nature of participation.

Data Analysis

For both participant groups, the box plot of SUS and TAM questionnaires were plotted using SPSS Statistics 25 software (IBM Corp), while the individual plots were plotted with the Minitab software's trial version (Minitab, LLC). For each interview and questionnaire, the average and SD of the positive statements for each game type were calculated. Student's game statistics, including the average play time, the number of correct answers, and difficulty-based changes, were also presented in detail. In addition, the Wilcoxon signed rank tests for both SUS and TAM questionnaires were calculated. This test was applied to compare the standard and adaptive versions of the games since data were not normally distributed.

Results

Overview

In this section, data acquired from the interviews are analyzed. Throughout this study, two interviews with both participant groups (students and educators) were done. Answers to the following questions were examined:

- Are proposed games suitable for training purposes?
- Are children with SpLD comfortable with different types of the proposed games?
- Can an adaptive difficulty system provide a better game experience for children with SpLD?

The Questionnaire Results of the Students

There are 10 questions in the SUS, where half of the questions consist of positive statements. The TAM questionnaire includes 19 questions about games to understand participants' acceptance level of the proposed games (Table S1 in Multimedia Appendix 9). The playtime of the games was recorded, and the average playing times are shown in Table S2 in Multimedia Appendix 9.

The Comments and Suggestions of the Students

The comments of children (Multimedia Appendix 10) were examined in three sections: comments on current games, future game ideas, and general feedback. For the first part, children were asked to comment on existing game ideas, the difficulty level of games, visuals, and all related things to the games. Most children said that the Math game was too simple in terms of visualization and the question types. They all said that there should be other operations (multiplication and division), and there should be more themes such as planes and cars besides kite themes. Multiplication was added to the final version of the game. Most of the students complained about the number of enemies and their small sizes in the Space game, which were fixed in the final prototype. Most of the students wanted to play the Space game level by level rather than in an endless fashion. In addition, they said that they wanted to see "boss creatures" at the end of each level. Moreover, most children suggested alternative control mechanisms for the Space game, such as buttons, touch joystick, or dragging. After these suggestions, a touch joystick was added in the last version. There were not many common comments on the Word game. Finally, students were asked if they had any other comments in general. Three of them suggested that music or audial feedback in the game will make it more fun. Most of the children proposed an in-game achievement system. Finally, one of the children suggested a ranking system so that players can understand if they won something or not. Multimedia Appendix 11 shows the most common words or phrases in the comments of students. The most common five words were *colorful*, *ship*, *car*, *control*, and *enemy*.

Game Statistics of the Students

In this section, statistics of the games are shown. Tables S3 and S4 in Multimedia Appendix 9 show the data of students in the Memory game from the first interview. *Number of false grids* refers to when the player fills a grid that is supposed to be empty. *Number of false color/arrow markings* refers to cases when the player marks a grid with a wrong color or arrow option. The difficulty of the Memory game changes with the number of grids since they should be memorized to be filled. Students played this game three times: the first color version with low difficulty to get used to the game, then again the color version with greater difficulty, and finally the arrow version. The level of difficulty was set according to the age of the students.

Tables S5-S7 in Multimedia Appendix 9 show the students' game statistics from the first interview of the Space game, Math game, and Word game, respectively.

Statistics of both parts of the Category game are shown in Tables S8 and S9 in Multimedia Appendix 9. For part two, an irrelevant answer to the given category is counted as *False Answer*, while misspelled words are counted as *Typo*. Table S10 in Multimedia Appendix 9 shows the students' data in the Memory game from the second interview, where they played each version of the game once. The Space game statistics from the second interview with the students are shown in Table S11 in Multimedia Appendix 9.

Table 1 shows how often the adaptive difficulty system adjusts the difficulty while the students play the games.

Table 1. Adaptive difficulty changes during game: mean and SD results of in-game adaptive difficulty changes for both games in the second session with students.

Game	Difficulty up, mean (SD)	Difficulty down, mean (SD)
Memory game	2.58 (1.41)	1.25 (0.97)
Space game	1.52 (1.20)	1.56 (0.98)

Wilcoxon signed rank test results can be seen in Table 2. This test was applied to compare the games' standard and adaptive versions since data were not normally distributed. Column *Z* refers to the *z* score, which is closer to 0 when the groups are evenly distributed. A P value of .05 is approximately equal to a *z* score of 1.96. The asymptotic significance column shows the significance of the difference. A value of less than 0.05 is considered significant. Figures 8 and 9 show the box plots of playing time for the Space game and Memory game, respectively.

Table 2. Wilcoxon signed rank test: results of test between data of two versions of both games from the second session with students.

Game	*Z* score	Asymptotic significance (two-tailed)
Memory game: correct answers	–0.567	0.571
Memory game: false grid	–0.138	0.890
Memory game: false letter	–0.536	0.592
Space game: crashed gem	–2.173	0.030
Space game: crashed enemy	–0.924	0.355
Space game: playtime	–1.884	0.060

Figure 8. Box plot of playing time for the Space game from the second session with students.

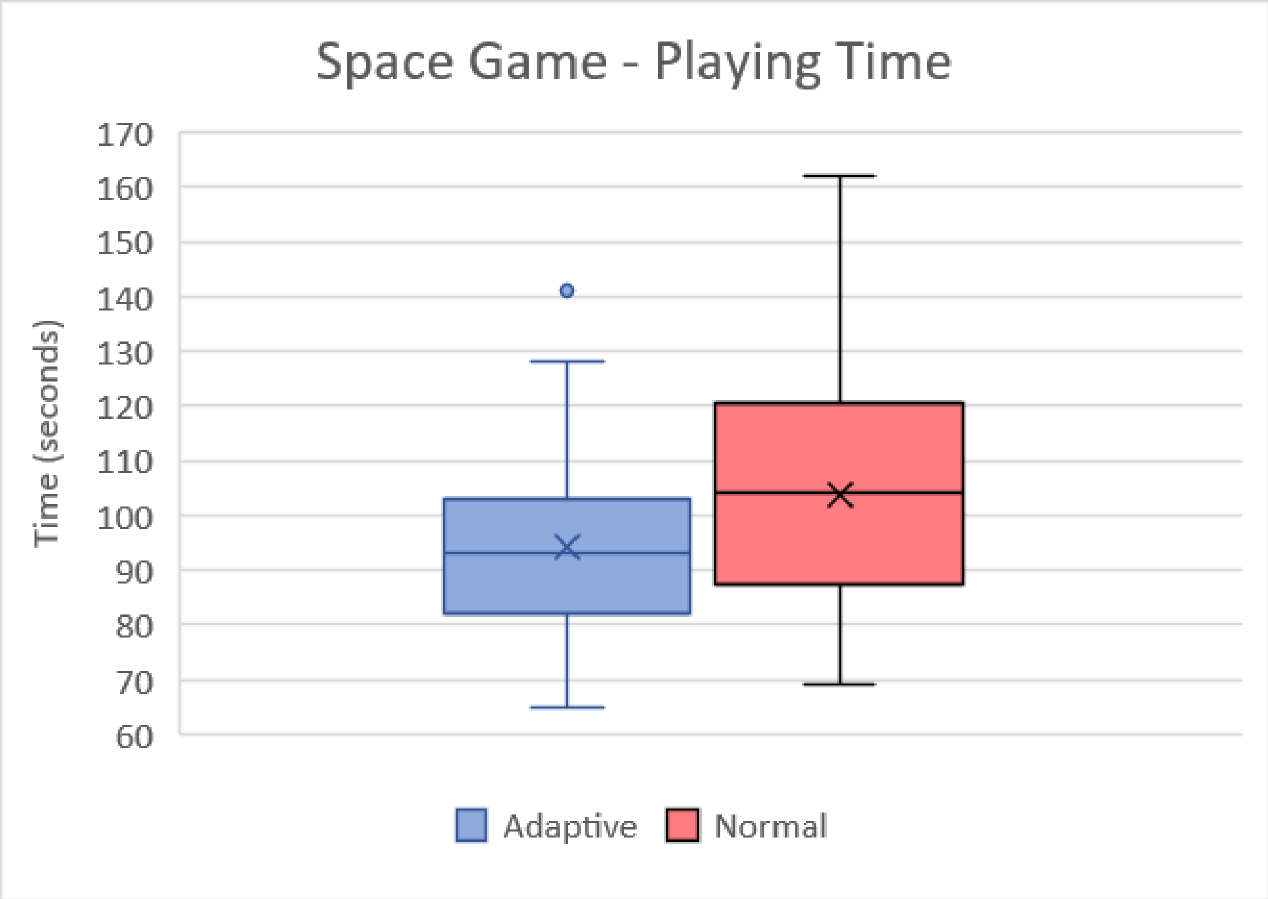
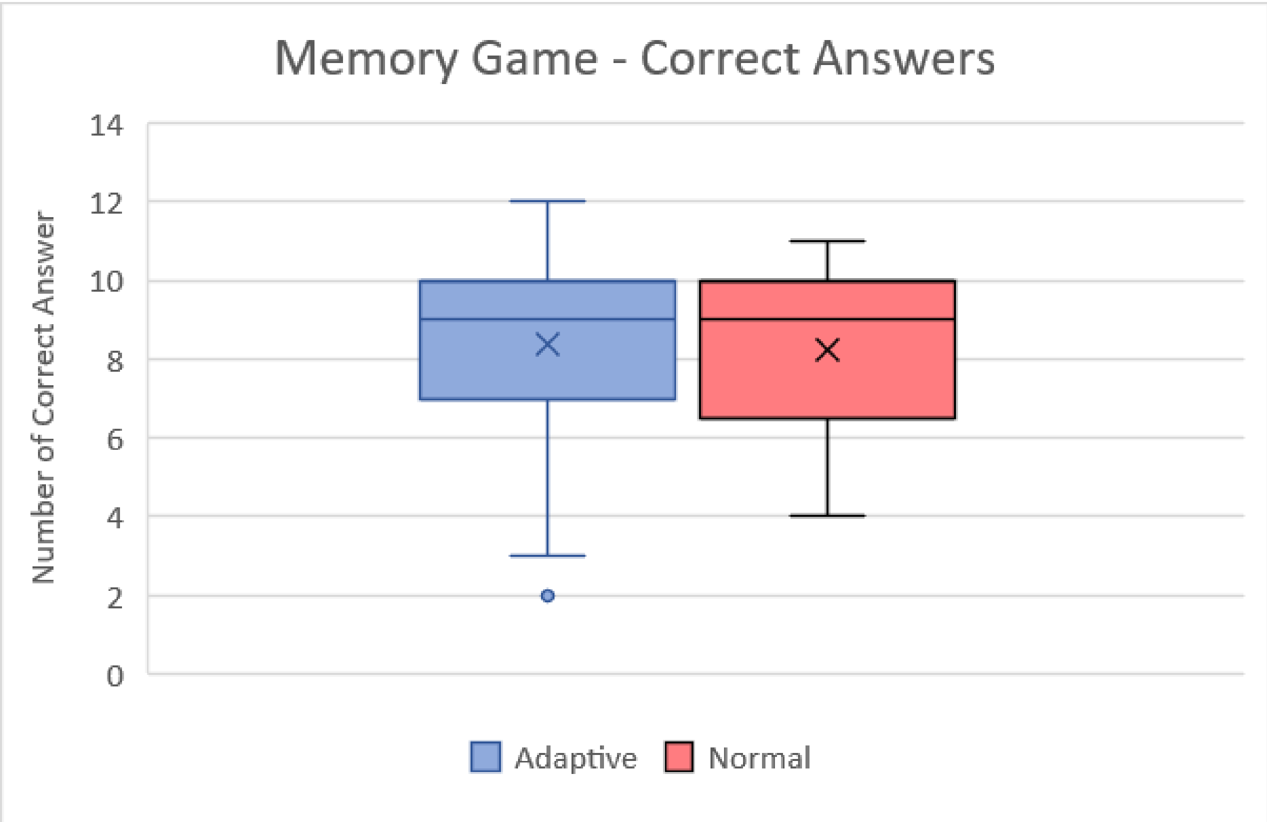


Figure 9. Box plot of correct answers for the Memory game from the second session with students.



The Questionnaire Results of the Educators

The TAM questionnaire results of educators from both interviews are displayed in Table S12 in Multimedia Appendix 9. Table S13 in Multimedia Appendix 9 shows the Wilcoxon signed rank test results of educators for both questionnaires.

Figures 10 and 11 show the box plots of the false answers of the Memory game and the difference between the correct and false answers, respectively. Figures 12 and 13 show the box plots of the SUS and TAM questionnaires, while Multimedia Appendix 12 and 13 display the individual plots of the SUS and TAM questionnaires.

Figure 10. Box plot of false answers for the Memory game from the second session with students.

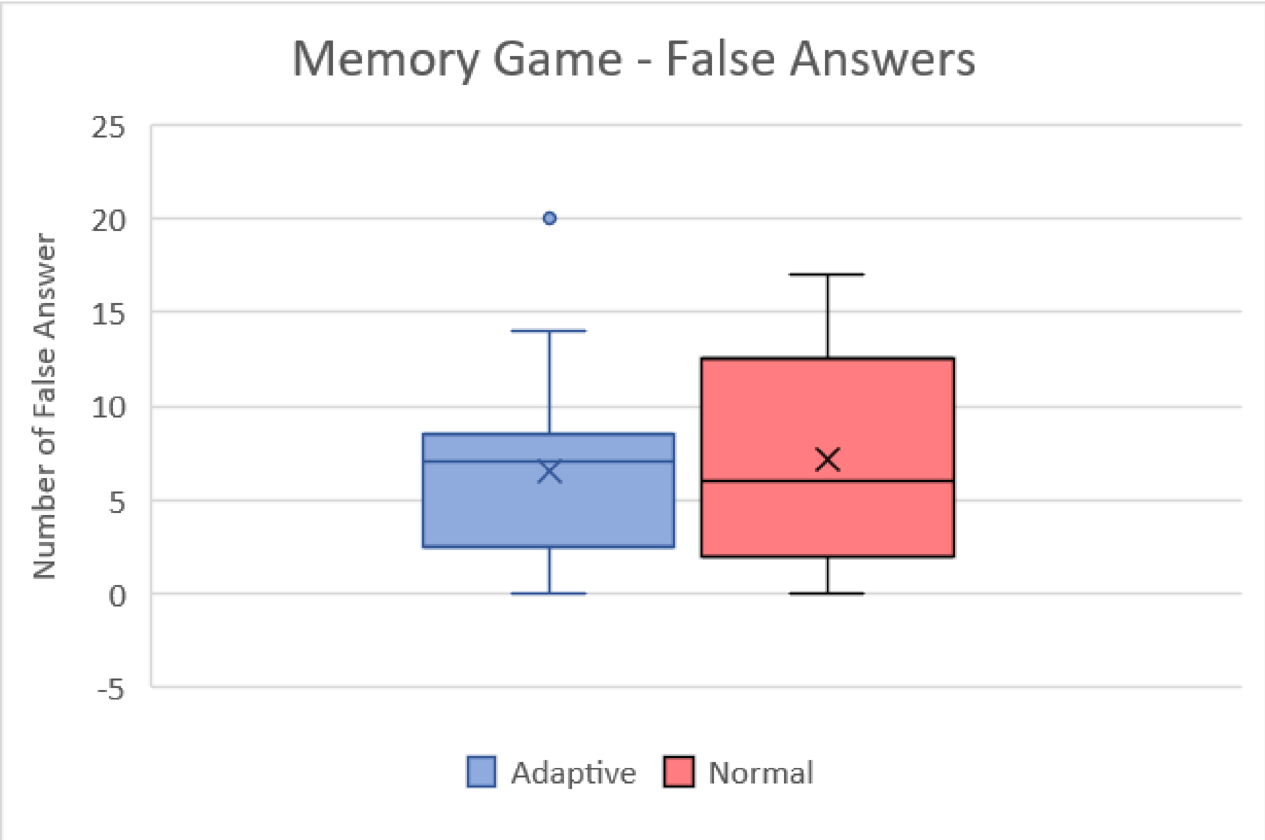


Figure 11. Box plot of the difference between the correct and false answers for the Memory game from the second session with students.

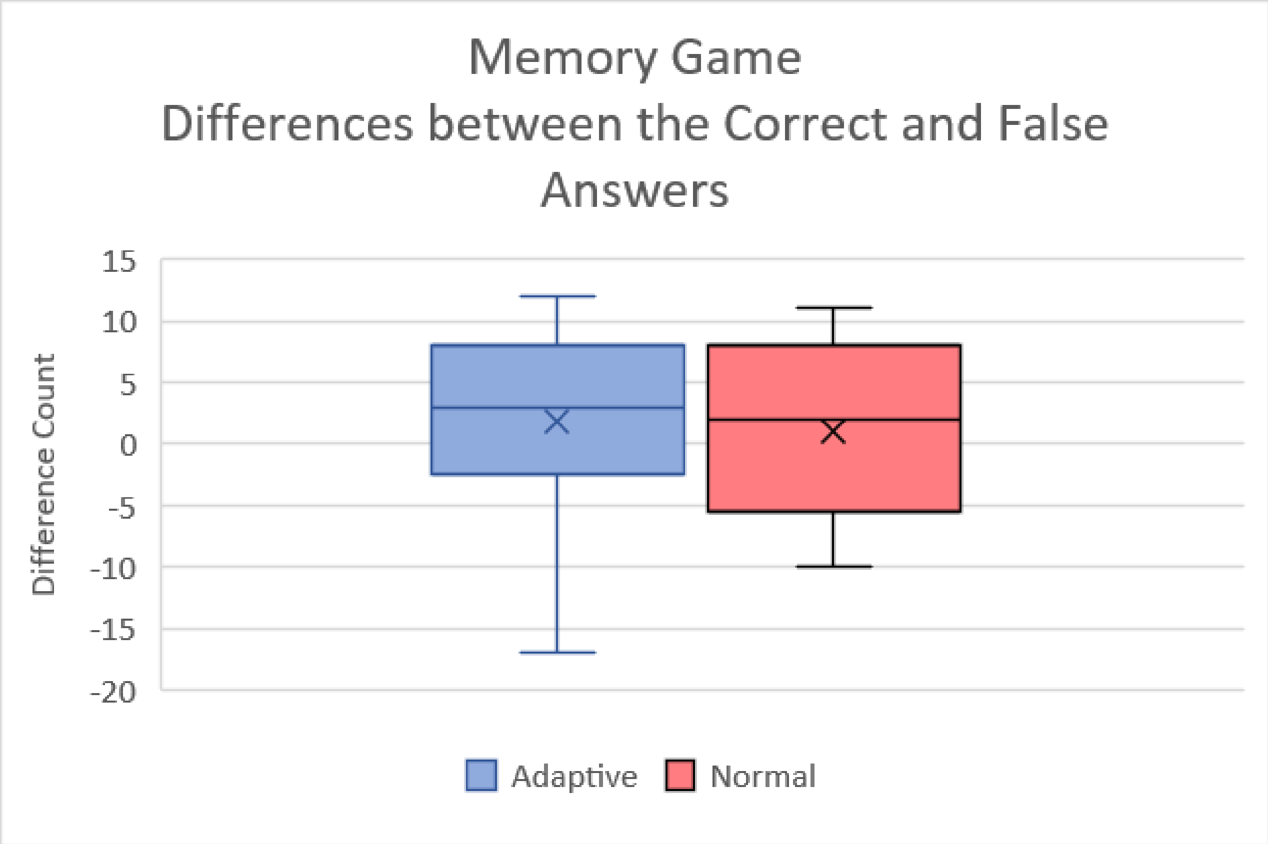


Figure 12. Box plot of SUS scores of the educators from the first (October) and second (March) sessions. SUS: System Usability Scale.

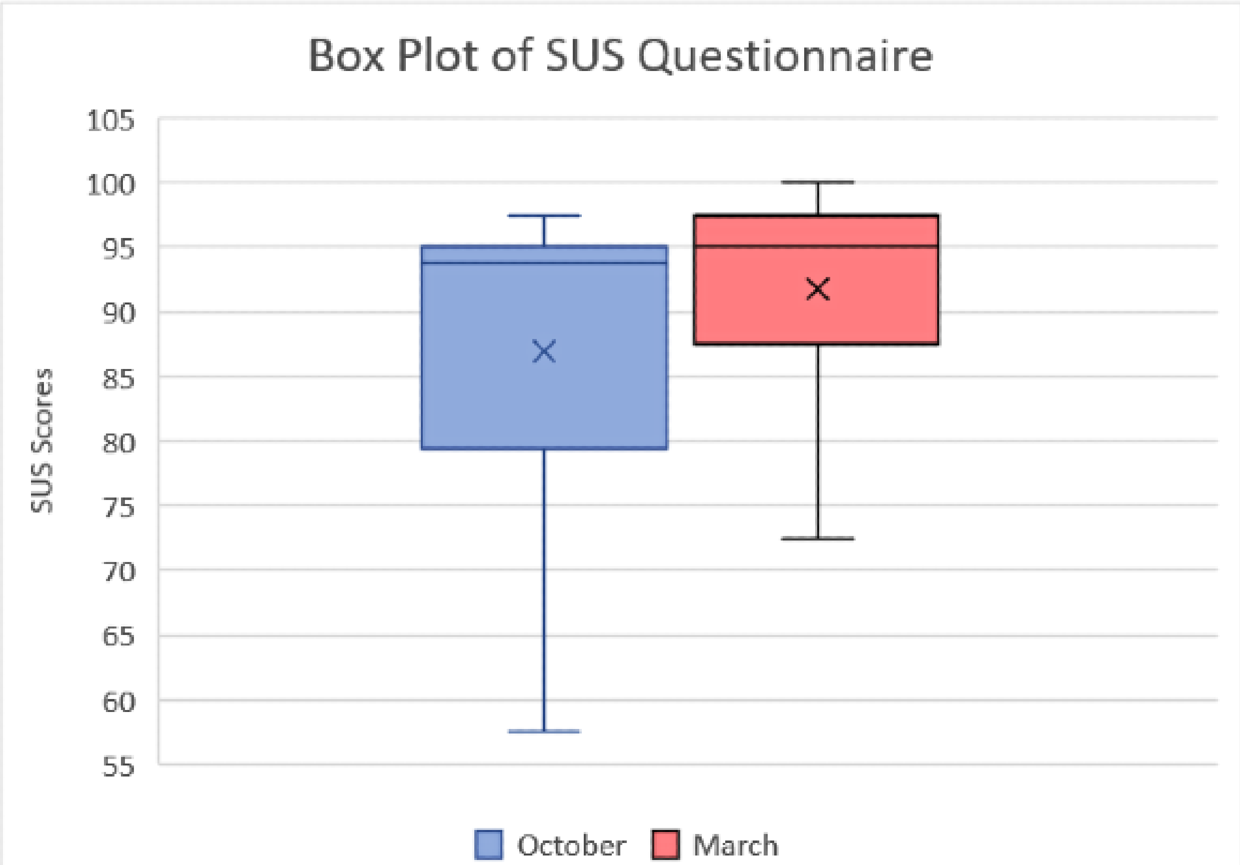
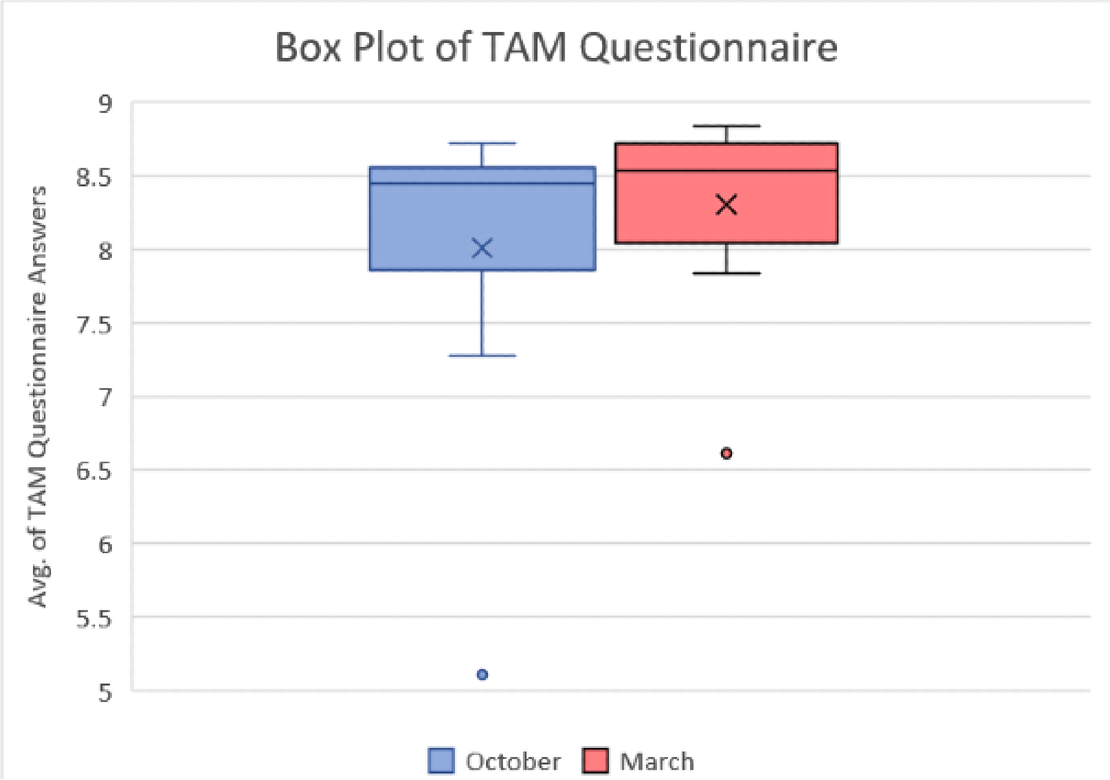


Figure 13. Box plot of TAM questionnaire answers of educators from the first (October) and second (March) sessions. TAM: Technology Acceptance Model.



The Comments and Suggestions of the Educators

The educators' feedback (Multimedia Appendix 14) was also analyzed under prototyped game ideas, potential game ideas, and general feedback. For the first segment, educators were asked to comment on current games, game difficulty levels, graphics, and game-related features. Instead of words, one of the educators proposed that game pictures could be used for the first part of the category. Another educator suggested that visuals should enrich the Math game without creating distractions. Other teachers suggested that creating a shape by collecting a gem in the Space game would support children's sketching ability. She also said that a star or a diamond should be gained instead of points since these symbols correspond to the special education center's current motivation system.

Educators were also asked if there was any game they want to see in this set of games. One educator suggested a game that enables the players to write letters by tracing a line or an area with fingers. Another educator said that a game based on finding synonyms or antonyms would be helpful to children with dyslexia. Rhythmic counting is indicated as another possible addition to the Math game. A "Spot 7 differences" type of puzzle and jigsaw puzzles were also suggested. Finally, an educator said that a game to improve players' auditory memory would be beneficial. Three educators said the lack of feedback makes it difficult to feel like they are making progress. Two educators believed that the degrees of difficulty should be varied. Two other educators indicated that holding scores on a leaderboard list and displaying previous scores would be helpful and encouraging to children. One of the educators indicated that a guide for educators or parents would be helpful. Finally, although most educators favored using this method in the classroom, as homework or as auxiliary material, only one educator said she would not use the games until the parents were sufficiently aware of this system and its benefits. Multimedia Appendix 15 shows the most common words or phrases in the comments used by the educators. The most common five words were *feedback*, *audial feedback*, *well-thought*, *visual elements*, and *reward*. The only negative feedback was "Will not use," which was said due to the educator's concern about parents' unfamiliarity with tablet use in special education.

Discussion

Principal Findings

In this study, five different serious games and their prototypes were developed to provide training material for children with any type of SpLD. The proposed games were evaluated in a two-phase study consisting of the SUS and TAM questionnaires, and open-ended questions for 23 children with SpLD and 10 educators. The main findings of this study showcase the potentials of using serious games in SpLD training, the impact of adding adaptive difficulty systems that enable personalized training experiences, and the outcomes of player-oriented design where each prototype was structured based on the player statistics and suggestions.

Results on comments and technology acceptance of the students (Table S1 in Multimedia Appendix 9) show that children evaluated the games as easy-to-use and easy-to-learn. *Easy-to-use* was a design criterion, and the results demonstrate that it was accomplished. Since the game was designed with a single control mechanism and a single goal, children did not feel uncomfortable or overwhelmed, even when the game itself was challenging (like the Space game). Children with SpLD aged 7 to 11 years were comfortable with different input types (dragging, tapping, and joystick) when using only one of them during the game. The single goal also mattered because most of these children did not experience more complex games before these interventions (Multimedia Appendix 16). To habituate the children with SpLD to serious games, not only difficulty levels but also the complexity of games were increased in a stepwise fashion.

The results of the questionnaires showed that the outcomes of the feedback-related questions were relatively low. There was no audio feedback in the proposed games, and visual feedback was reduced to prevent any possible distraction. After this result, the final version of the games was modified to have slightly more visual feedback. The questionnaire results of the educators supported these initial findings. Educators remarked that the games were easy to learn and play, but there was a lack of feedback in the final version of the games, although the games included several visual feedbacks. The children's average points indicated that verbal games (the Word and Category games) had the lowest scores (8.36 and 8.08 out of 10, respectively), which may stem from the duration of the games. Both the Word and Category games took more than four times longer than other games during game play. Since playtime did not include any level mechanism in these games, it became boring for children after a while. In addition, children gave a correct answer in 54.8 seconds (Word game) and 38 seconds (the second part of the Category game), which can be accepted as exceedingly long time durations for students with dyslexia. The point difference between the Category and Word games is based on the visual differences. The Word game has a child avatar and a kitchen illustration, which is a more child-oriented visual theme. The Word game, which only includes a blackboard and some writings, was liked by the children, and they completed the game in their given time period. This result conflicts with Shabbir et al's [19] rules of serious games for children with dyslexia. Shabbir et al [19] claim that not using a plain background can cause a distraction for children. However, the children's questionnaires and comments show that game-related static background images or avatars increased the immersion. The kitchen image in the background or the child avatar that asks questions was not mandatory for the Word game, but children stated that they loved both versions. Some of the children suggested a similar visual design for the Category game.

The average points by both participant groups were between 7.9 and 8.84 out of 10. When these results were combined with the questionnaire results, it can be easily said that both participant groups liked all games, and they can accept all parts of the system as a training source. Unlike other studies on serious games for children with SpLD, this study proposes a collection of games to educate children with any type of SpLD. All of the games can be played by children regardless of their SpLD types. A specially designed set of games for these children can provide a complete training source and better experience rather than only playing a game directly related to their disorder or difficulty. The average number of days when students want to play these games was 3.6 per week. This average can be accepted as high since the students' permission to play games was limited. Almost half of the kids indicated that they were allowed to play games only during the weekend or every other day, and they said that they would play these games in that allocated limited time.

A two-phase SUS and TAM evaluation was provided in this study, although the SUS scores between initial and final versions of the games were demonstrating similar adjective ratings (ie, best imaginable, in our case) as mentioned by Bangor et al [20]. However, it was seen that in the second interview, the average of the SUS was higher, and the SD was lower, which can be considered to demonstrate the condensed results regarding the outcome of the games, where the hesitations and fluctuations from the data were less than the initial versions.

A rule-based adaptive difficulty system was also proposed in serious games that adjust difficulty during the game, in contrast to other studies in the literature that used a system that changes difficulty between two play sessions [19] and used predefined levels, which is limited [21]. The Space game's adaptive version did not change the results considerably for the students. Moreover, the game's speed change was practically equal, which means the users nearly had the same statistics at the same difficulty level. It can be interpreted in a couple of ways. The first possible reason for this result can be the adaptive rules or effects not being well-adjusted, which is too difficult to achieve. In that case, the players will usually play in normal difficulty. If adaptive changes affect the game difficulty, too many players will probably increase or decrease their pace to the standard difficulty level. The other possible reason for this result can be the game difficulty being too balanced. The game could have such a difficulty that players can collect the gem and crash into enemies at almost the same rate. However, this is not a desired outcome since children should encounter this situation after a successful game period to feel their progress. The Wilcoxon test for the Memory game results showed that children had better performance in the game's adaptive version, but it was not statistically significant. Since the adaptive version increases its difficulty while players show good performance, children played a more challenging version and showed better performance, which means that the flow strategy worked. Due to the increased difficulty level, children's performance could not be statistically different but was still better.

Especially for adaptive games, more tests should be done for different variables. This study includes the modification of only one variable—enemy spawn frequency for the Space game and question screen time for the Memory game—which gives promising results for the future, especially in the Memory game. Children did not realize they were playing faster, and they showed slightly better performance in adaptive versions. More tests can be applied by changing grid numbers or changing the predefined letters in the left panel. During the test of the adaptive games, only 1 student said that he understood the change in the Memory game and explained the correct reason, while others could not explain the change they felt in both games, misjudged them, or did not feel any change at all. Many students showed that the games influenced them by expressing sadness, joy, or surprise during the games—mainly in the Space game. Except for 1 student, the students were eager to replay the games, and they demanded to continue to play after the interventions were complete. In the first interview, 1 student stated that playing the Space game for the second time would be unnecessary, so she did not want to replay the game. The same student did not make such a request in other games or during the second interview. The students wanted different visual themes in the Math game, showing that the game could be diversified visually, not mechanically. The "boss creatures" concept was the most highlighted suggestion of the students, which was proposed as an improvement several times. It may be good to add them as positive feedback in the games that are designed in levels instead of endless games. Besides, although the game ends due to the player's mistake in endless games, in level-based games, the player's success leads to a new stage, which changes the feeling at the end of the game. Thus, boss creatures can be good for strengthening the feeling at the end of the episode. Students' comments on serious games were usually game-related suggestions, and they were not interested in the games' educational purpose. This is because children did not see these games as educational material, but they accepted these games as standard games, which can be considered in line with the fundamental purpose of the serious game concept.

Conclusion

This study proposed and tested five serious games, each having a different number of prototypes, specifically developed for children with SpLD. Both educators and students participated in the game-based interventions, and their feedback was recorded. Both participant groups were enthusiastic about the proposed "all-in-one" scheme. Besides, different design requirements that have been listed in previous studies were investigated with an additional adaptive difficulty level. These adjustments and personalized approaches increased children with dyslexia's immersion in the games. One of the study's contributions was the use of serious games to target various SpLD types. Each child participant had more than one form of SpLD, and playing games that focus on dyslexia or dyscalculia has been beneficial to them instead of playing games that focus on only one of these difficulties.

Furthermore, playing games like the Memory and Space games and education-oriented question-and-answer games resulted in increased positive feedback and excitement. Furthermore, the students mentioned that they preferred to play in a general system rather than a system focusing only on an area where they already had difficulties. The teachers' overall impressions were highly positive, and most of them intend to use these games as part of their curriculum's supplementary material. This integration will enable the games to be tested for a longer period so that new enhancements can be added to the proposed games with additional evaluations to measure their educational impact.

Acknowledgments

The authors thank Dislev Türkiye Disleksili Çocuklar Vakfı (Dislev Turkish Association of Children with Dyslexia) for their tremendous help during data acquisition. The authors also thank the children and their educators for participating in this study and giving valuable and detailed feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The workflow diagram of this study.
[PNG File , 29 KB - games_v9i2e25997_app1.png]

Multimedia Appendix 2
Screenshot of the picture version of the Memory game from Prototype IV.
[PNG File , 83 KB - games_v9i2e25997_app2.png]

Multimedia Appendix 3
Screenshot of the first part of the Category game from Prototype II.
[PNG File , 564 KB - games_v9i2e25997_app3.png]

Multimedia Appendix 4
Screenshot of the second part of the Category game from Prototype II.
[PNG File , 439 KB - games_v9i2e25997_app4.png]

Multimedia Appendix 5
Screenshot of the second part of the Category game with an alternative keyboard layout from Prototype III for the theme "Clothes".
[PNG File , 477 KB - games_v9i2e25997_app5.png]

Multimedia Appendix 6
The optimal game play corridor and the golden path figure are adapted from Thomas 2010.
[PNG File , 109 KB - games_v9i2e25997_app6.png]

Multimedia Appendix 7
Visual diagram of the adaptive difficulty system.
[PNG File , 33 KB - games_v9i2e25997_app7.png]

Multimedia Appendix 8
Screenshot of the letter version of the Memory game from Prototype III.
[PNG File , 123 KB - games_v9i2e25997_app8.png]

Multimedia Appendix 9
Supplementary files and tables on game statistics and results of the questionnaires.
[DOCX File , 20 KB - games_v9i2e25997_app9.docx]

Multimedia Appendix 10
Student feedback illustration map.
[PNG File , 53 KB - games_v9i2e25997_app10.png]

Multimedia Appendix 11
Student feedback word cloud.
[PNG File , 112 KB - games_v9i2e25997_app11.png]

Multimedia Appendix 12
Individual plot of System Usability Scale scores of educators from the first (October) and second (March) sessions.
[PNG File , 13 KB - games_v9i2e25997_app12.png]

Multimedia Appendix 13
Individual plot of Technology Acceptance Model questionnaire answers of educators from the first (October) and second (March) sessions.
[PNG File , 14 KB - games_v9i2e25997_app13.png]

Multimedia Appendix 14
Educator feedback illustration map.
[PNG File , 47 KB - games_v9i2e25997_app14.png]

Multimedia Appendix 15
Educator feedback word cloud.
[PNG File , 96 KB - games_v9i2e25997_app15.png]

Multimedia Appendix 16
Photo of a child during the intervention.
[PNG File , 313 KB - games_v9i2e25997_app16.png]

References

1. Lipkin PH, Okamoto J, Council on Children with Disabilities, Council on School Health. The Individuals With Disabilities Education Act (IDEA) for children with special educational needs. Pediatrics 2015 Dec;136(6):e1650-e1662. [doi: 10.1542/peds.2015-3409] [Medline: 26620061]
2. Balcı E. Disleksi hakkında öğretmen görüşleri ve karşılaştıkları sorunlar. Ege Eğitim Dergisi 2019;20(1):162-179. [doi: 10.12984/egeefd.453922]
3. Bellotti F, Ott M, Arnab S, Berta R, de Freitas S, Kiili K, et al. Designing serious games for education: from pedagogical principles to game mechanisms. 2011 Oct 1 Presented at: The 5th European Conference on Games Based Learning; 2011; Athens, Greece.
4. Franceschini S, Gori S, Ruffino M, Viola S, Molteni M, Facoetti A. Action video games make dyslexic children read better. Curr Biol 2013 Mar 18;23(6):462-466 [FREE Full text] [doi: 10.1016/j.cub.2013.01.044] [Medline: 23453956]

5. Bartolomé NA, Zorrilla AM, Zapirain BG. Dyslexia diagnosis in reading stage though the use of games at school. 2012 Jul 30 Presented at: 17th International Conference on Computer Games; July 30-August 1, 2012; Louisville, KY. [doi: 10.1109/cgames.2012.6314544]
6. Rauschenberger M, Rello L, Baeza-Yates R, Bigham JP. Towards language independent detection of dyslexia with a web-based game. In: Proceedings of the 15th International Web for All Conference. 2018 Apr 01 Presented at: W4A '18; April 2018; Lyon, France p. 1-10. [doi: 10.1145/3192714.3192816]
7. Avila-Pesantez DF, Vaca-Cardenas LA, Avila RD, Padilla NP, Rivera LA. Design of an augmented reality serious game for children with dyscalculia: a case study. In: Botto-Tobar M, Pizarro G, Zúñiga-Prieto M, D'Armas M, Sánchez MZ, editors. Technology Trends: 4th International Conference, CITT 2018, Babahoyo, Ecuador, August 29–31, 2018, Revised Selected Papers. Cham: Springer; 2019:165-175.
8. Holz H, Brandelik K, Beuttler B, Brandelik J, Ninaus M. How to train your syllable stress awareness - a digital game-based intervention for German dyslexic children. Int J Serious Games 2018 Sep 25;5(3):37-59. [doi: 10.17083/ijsg.v5i3.242]
9. Sood M, Toornstra A, Sereno M, Boland M, Filaretti D, Sood A. A digital app to aid detection, monitoring, and management of dyslexia in young children (DIMMAND): protocol for a digital health and education solution. JMIR Res Protoc 2018 May 17;7(5):e135 [FREE Full text] [doi: 10.2196/resprot.9583] [Medline: 29773528]
10. Flogie A, Aberšek B, Kordigel Aberšek M, Sik Lanyi C, Pesek I. Development and evaluation of intelligent serious games for children with learning difficulties: observational study. JMIR Serious Games 2020 Apr 16;8(2):e13190 [FREE Full text] [doi: 10.2196/13190] [Medline: 32297864]
11. Yıldırım O. Development of adaptive serious games for the training of children with specific learning disabilities. OpenMETU. 2020 Sep 10. URL: https://open.metu.edu.tr/handle/11511/69207 [accessed 2021-05-27]
12. Tobler-Ammann BC, Surer E, de Bruin ED, Rabuffetti M, Borghese NA, Mainetti R, et al. Exergames encouraging exploration of hemineglected space in stroke patients with visuospatial neglect: a feasibility study. JMIR Serious Games 2017 Aug 25;5(3):e17 [FREE Full text] [doi: 10.2196/games.7923] [Medline: 28842388]
13. Tobler-Ammann BC, Surer E, Knols RH, Borghese NA, de Bruin ED. User perspectives on exergames designed to explore the hemineglected space for stroke patients with visuospatial neglect: usability study. JMIR Serious Games 2017 Aug 25;5(3):e18 [FREE Full text] [doi: 10.2196/games.8013] [Medline: 28842390]
14. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. United Kingdom: Taylor & Francis; 1996.
15. Davis, Jr FD. A technology acceptance model for empirically testing new end-user information systems: theory and results. DSpace@MIT. 1985. URL: https://dspace.mit.edu/bitstream/handle/1721.1/15192/14927137-MIT.pdf [accessed 2021-05-27]
16. Czikszentmihalyi M. Flow: The Psychology of Optimal Experience. New York, NY: Harper & Row; 1990.
17. Thomas JM, Young RM. Annie: automated generation of adaptive learner guidance for fun serious games. IEEE Trans Learning Technologies 2010 Oct;3(4):329-343. [doi: 10.1109/tlt.2010.32]
18. Hendrix M, Bellamy-Wood T, McKay S, Bloom V, Dunwell I. Implementing adaptive game difficulty balancing in serious games. IEEE Trans Games 2019 Dec;11(4):320-327. [doi: 10.1109/tg.2018.2791019]
19. Shabbir N, Bhatti Z, Hakro DN. Serious game user interface design rules for dyslexic children. 2019 Dec 14 Presented at: 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics; December 14-15, 2019; Karachi, Pakistan p. 1-6. [doi: 10.1109/macs48846.2019.9024786]
20. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 May 01:114-123.
21. Cuschieri T, Khaled R, Farrugia VE, Martinez HP, Yannakakis GN. The iLearnRW game: support for students with dyslexia in class and at home. 2014 Presented at: 6th International Conference on Games and Virtual Worlds for Serious Applications; September 9-12, 2014; Valletta, Malta p. 1-2. [doi: 10.1109/vs-games.2014.7012167]

Abbreviations

IDEA: Individuals with Disabilities Education Act
LD: learning disorder
SpLD: specific learning difficulties
SUS: System Usability Scale
TAM: Technology Acceptance Model

Edited by N Zary; submitted 07.04.21; peer-reviewed by CW Ong, J Brooke; comments to author 22.04.21; revised version received 30.04.21; accepted 21.05.21; published 31.05.21.

Please cite as:
Yildirim O, Surer E
Developing Adaptive Serious Games for Children With Specific Learning Difficulties: A Two-phase Usability and Technology Acceptance Study
JMIR Serious Games 2021;9(2):e25997
URL: https://games.jmir.org/2021/2/e25997
doi:10.2196/25997
PMID:34057415

Original Paper

Virtual Reality Human–Human Interface to Deliver Psychotherapy to People Experiencing Auditory Verbal Hallucinations: Development and Usability Study

Mischa Brander[1*], MSc; Stephan T Egger[2,3*], MD; Noa Hürlimann[2], MSc; Erich Seifritz[2], MD; Robert W Sumner[1], PhD; Stefan Vetter[2], MD; Stéphane Magnenat[1], PhD

[1]Game Technology Center, Department of Computer Science, Eidgenössische Technische Hochschule Zürich, Zurich, Switzerland

[2]Psychotherapy and Psychosomatics, Department of Psychiatry, Psychiatric University Hospital of Zurich, University of Zürich, Zurich, Switzerland

[3]Department of Psychiatry, Faculty of Medicine, University of Oviedo, Oviedo, Spain

[*]these authors contributed equally

Corresponding Author:
Stephan T Egger, MD
Psychotherapy and Psychosomatics
Department of Psychiatry
Psychiatric University Hospital of Zurich, University of Zürich
Lenggstrasse 31
Zurich, 8032
Switzerland
Phone: 41 523049340
Email: stephan.egger@pukzh.ch

Abstract

Background: Digital technologies have expanded the options for delivering psychotherapy, permitting for example, the treatment of schizophrenia using Avatar Therapy. Despite its considerable potential, this treatment method has not been widely disseminated. As a result, its operability and functionality remain largely unknown.

Objective: We aimed to study the usability of a therapeutic virtual reality human–human interface, created in a game engine.

Methods: Participants were psychiatric hospital staff who were introduced to the therapeutic platform in a hands-on session. The System Usability Scale (SUS) was employed for evaluation purposes. Statistical evaluation was conducted using descriptive statistics, the chi-square test, analysis of variance, and multilevel factor analysis.

Results: In total, 109 staff members were introduced to the therapeutic tool and completed the SUS. The mean SUS global score was 81.49 (SD 11.1). Psychotherapists (mean 86.44, SD 8.79) scored significantly higher ($F_{2,106}$=6.136; P=.003) than nursing staff (mean 79.01, SD 13.30) and administrative personnel (mean 77.98, SD 10.72). A multilevel factor analysis demonstrates a different factor structure for each profession.

Conclusions: In all professional groups in this study, the usability of a digital psychotherapeutic tool developed using a game engine achieved the benchmark for an excellent system, scoring highest among the professional target group (psychotherapists). The usability of the system seems, to some extent, to be dependent on the professional background of the user. It is possible to create and customize novel psychotherapeutic approaches with gaming technologies and platforms.

Trial Registration: Clinicaltrials.gov NCT04099940; http://clinicaltrials.gov/ct2/show/NCT04099940

(JMIR Serious Games 2021;9(2):e26820) doi:10.2196/26820

KEYWORDS

system usability; virtual reality psychotherapy; verbal auditory hallucinations

Introduction

Psychotherapy is an effective and cost-efficient method for the treatment of psychiatric and psychological disorders [1]. Over the last few decades, it has been evolving continuously, demonstrating both feasibility and efficacy in practically all diagnostic categories. Indeed, in several categories, it has become the first line of treatment [2,3]. In patients with schizophrenia, psychotherapy as a treatment option has been largely neglected. Recently, however, it has gained recognition as an effective treatment when used in conjunction with pharmacotherapy [4,5]. Furthermore, current guidelines now recommend the early implementation of psychotherapy in the treatment process [6,7].

Psychotherapeutic treatment using digital technologies, virtual reality in particular, has been shown to be at least as efficacious as other treatments [8]. In some fields, particularly schizophrenia, digital technologies have considerably extended therapeutic options [5,9,10], for example, with the novel implementation of Avatar Therapy, whereby psychotherapy is delivered through a computer interface [11,12]. Patients with auditory verbal hallucinations create an avatar of a human entity, to which they attribute the voices. With the help of a therapist, they progressively gain control over the voices, which leads to a reduction of symptoms and distress while increasing quality of life [12].

Despite encouraging early studies and its vast potential, psychotherapeutic treatment using digital technologies has still not been widely disseminated in research or clinical practice [10]. We attribute the limited deployment partially to unavailability as off the shelf tools, making implementation difficult [13,14]. From previous research, it is known that for the optimal delivery of therapy through digital technologies, besides availability, the operability and functionality of the technology are crucial; only once these are well established can the therapist confidently utilize digital technology [14,15]. Moreover, the proper use of such technology is essential for the optimal delivery of the therapy, thus allowing the therapist to develop their therapeutic skills [10].

In this paper, we present a human–human interface that we developed for use in the treatment of patients experiencing verbal acoustic hallucinations. As the usability of the system is a prerequisite for its clinical application, we systematically sought input from nonpatient users [16]. Within mixed skill-grade users, we sought to determine what influence professional background and therapeutic skills have in relation to the use of the therapeutic system.

Methods

Virtual Reality Human–Human Interface

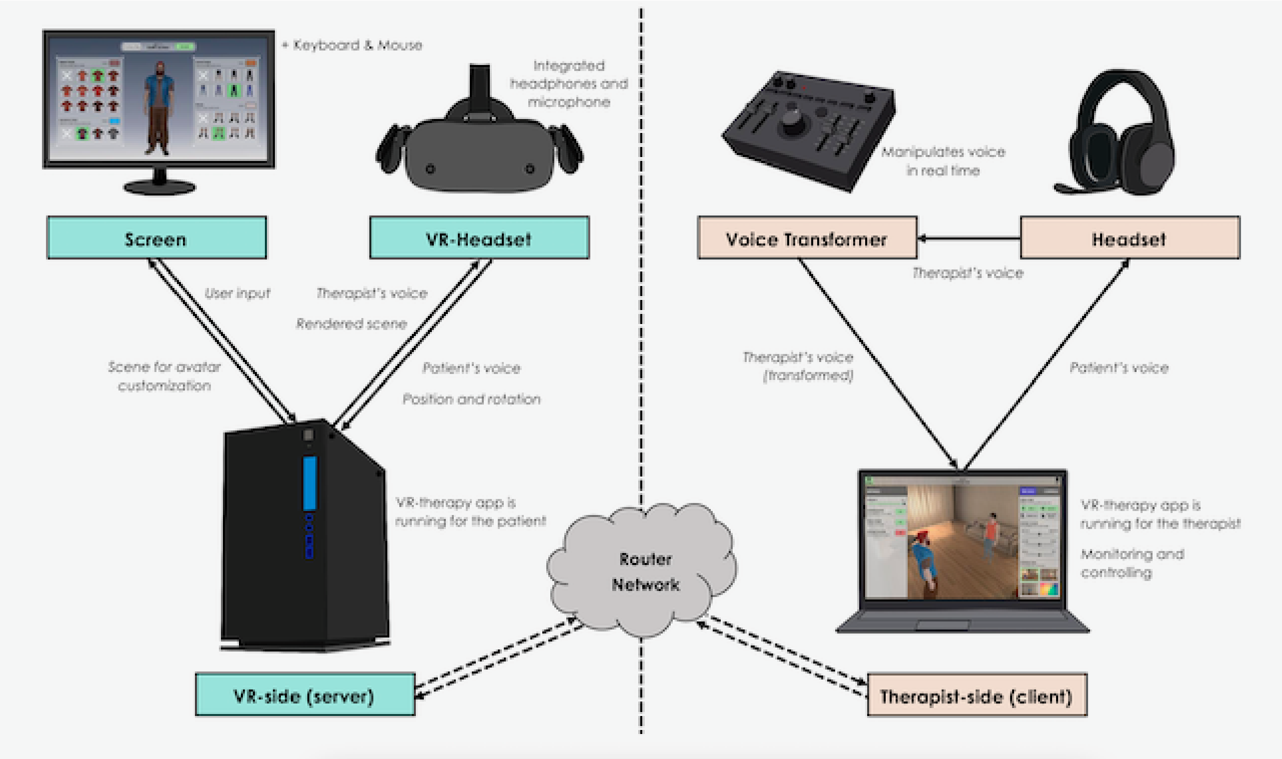
Building upon previous studies [11,17,18], we have created a virtual reality human–human interface using the Unity game engine (Unity Technologies) to deliver Avatar Therapy for people experiencing auditory verbal hallucinations. The basic design employs 2 separate apps running on different devices connected via a network, including bidirectional audio (full-duplex voice over internet protocol connection) communication. The first computer hosts a personal avatar creation tool (Virtual Reality Avatar-Creation Tool, VRAT-CT) to design and customize a humanoid avatar, to which patients attribute their auditory verbal hallucinations. This computer also renders the virtual reality through a head-mounted display for the therapeutic session. The voice of the therapist is modulated through a voice transformer (Roland VT-4) to match the auditory verbal hallucination. The therapeutic session is initialized and controlled from the second computer with a Control Center (VRAT-CC) which allows the therapist to control the Avatar and to speak through the Avatar using lip synchronization.

Software and Hardware

The Unity game engine is a freely available platform for game development. Over 60% of current virtual and augmented reality content have been created with Unity [19]. It provides a 3D editor, a scripting application programming interface written in C# which allows all components to be brought together, and supports 3D graphics, socket communication, and virtual reality (VR). Both apps are created with Unity (version 2019.3.7) and the associated scripts are written in C# with Microsoft Visual Studio 2017, an integrated development environment. For the 3D character, the Multipurpose Avatar package (version 2.11.5; Unity Technologies) was used. To increase the impression that the avatar is speaking, the SALSA LipSync Suite package (version 2.5.0.92; Crazy Minnow Studio LLC) is used to synchronize lip and mouth movements to the voice input of the therapist.

The Reverb (Hewlett Packard) headset is used as a head-mounted display. The headset provides a resolution of 2160×2160 per eye at 90 Hz and with a 114° field of view; however, this comes with a series of minimum computational requirements. The producer recommends, at minimum, a Nvidia GeForce GTX 1080 graphic card or an AMD Radeon Pro WX 8200, an Intel Core i7 processor, and 16 GB of RAM. For the operating system, Windows 10 (version 1809 or later) is needed. For this project, a Roland VT-4 voice transformer was used, which provided a set of options for manipulating a voice in real-time. Pitch and format frequency, which can be set with sliders for a deeper or higher voice, were relevant for creating the avatar voice. Figure 1 shows an overview of the software and hardware setup (Multimedia Appendix 1).

Figure 1. Virtual reality (VR) human–human interface.



System Usability Scale

The System Usability Scale (SUS) is a tool for measuring the usability of a wide variety of products and services including hardware, software, mobile devices, websites, and apps [18,20]. It is a 10-item questionnaire, with a 5-point Likert Scale from 1 (strongly disagree) to 5 (strongly agree) (Textbox 1). Scale items alternate between positive and negative statements; therefore, correction is required for scoring. For odd-numbered items, the value 1 is subtracted from the user's response whereas for even-numbered items, the user's response is subtracted from 5, yielding a score from 0 to 4 for each item. For interpretation, scores are summed and multiplied by a factor of 2.5. The final score ranges from 0 to 100 [20].

Textbox 1. System Usability Scale.

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Participants and Assessment

Employees (irrespective of professional background and occupation) of the Psychiatric University Hospital of Zürich were invited to view and test the virtual reality human–human interface used to deliver Avatar Therapy to people experiencing auditory verbal hallucinations. Basic demographic characteristics (age, gender, and occupation) were gathered. Participants were divided into 3 categories: psychotherapists (either psychiatrists or psychologists), nursing staff, and administrative personnel.

Procedure

Participants were individually informed about the nature of the study and introduced to the therapeutic platform in a hands-on session. They were provided with information about the theoretical background of the therapy and the design and implementation process of the virtual reality human–human interface in practice. Each step of the therapeutic process was explained. Afterward, they created an avatar and customized its voice before experiencing it through VR. Thus, they first created and customized their Avatar for therapy in the patient's role, after which they carried out a session in the therapeutic role. After the session, participants completed the SUS for each component.

Statistical Analysis

Descriptive statistics (percentages, means, standard deviations) were used to represent the demographic characteristics of the sample. Differences in the sample were calculated using the chi-square test for proportions. An analysis of variance was performed on continuous variables. The SUS score for the system was calculated. Scores for the avatar creation tool (VRAT-CT) and the control center (VRAT-CC) were evaluated separately. The SUS was evaluated at both item level and global level.

Additionally, a multilevel factor analysis was conducted. Statistical analysis was performed using the statistical language program (version 4.0.3, The R Project).

Ethics

The study was designed to comply with current ethical standards and local regulations. The ethics committee of the Canton of Zürich approved the study protocol (BASEC 2019-01386). The study was registered at Clinicaltrials.gov (NCT04099940).

Results

Sample Demographics

In total, 109 staff members were introduced to the therapy. The sample comprised psychotherapists (n=40), nursing staff (n=43), and administrative personnel (n=26); with a mean age of 34.76 (SD 12.69); 74 participants were female (67.9%). There were no statistically significant differences regarding age or gender distribution among the different professions (Table 1).

Table 1. Sample characteristics and outcome evaluation.

Results	Profession			Statistics	
	Psychotherapists (n=40)	Nursing staff (n=43)	Administrative personnel (n=26)	*F* test (*df1*,*df2*)	*P* value
Age, mean (SD)	33.25 (9.00)	33.51 (15.85)	30.15 (10.60)	2.09 (2, 106)	.13
Gender, n (%)				4.451 (2, 109)[a]	.11
Male	19 (47)	12 (28)	7 (27)		
Female	21 (53)	31 (72)	19 (73)		
SUS[b] score, mean (SD)					
Global	86.44 (8.79)[c]	79.01 (13.30)	77.98 (10.72)	6.136 (2, 106)	.003
Virtual reality avatar	87.00 (9.83)[c]	79.71 (13.56)	78.08 (12.50)	5.597 (2, 106)	.005
Control center	85.88 (9.53)[c]	78.31 (14.54)	77.88 (11.68)	5.064 (2, 106)	.008

[a]Chi-square test statistic (*df1*,*df2*).

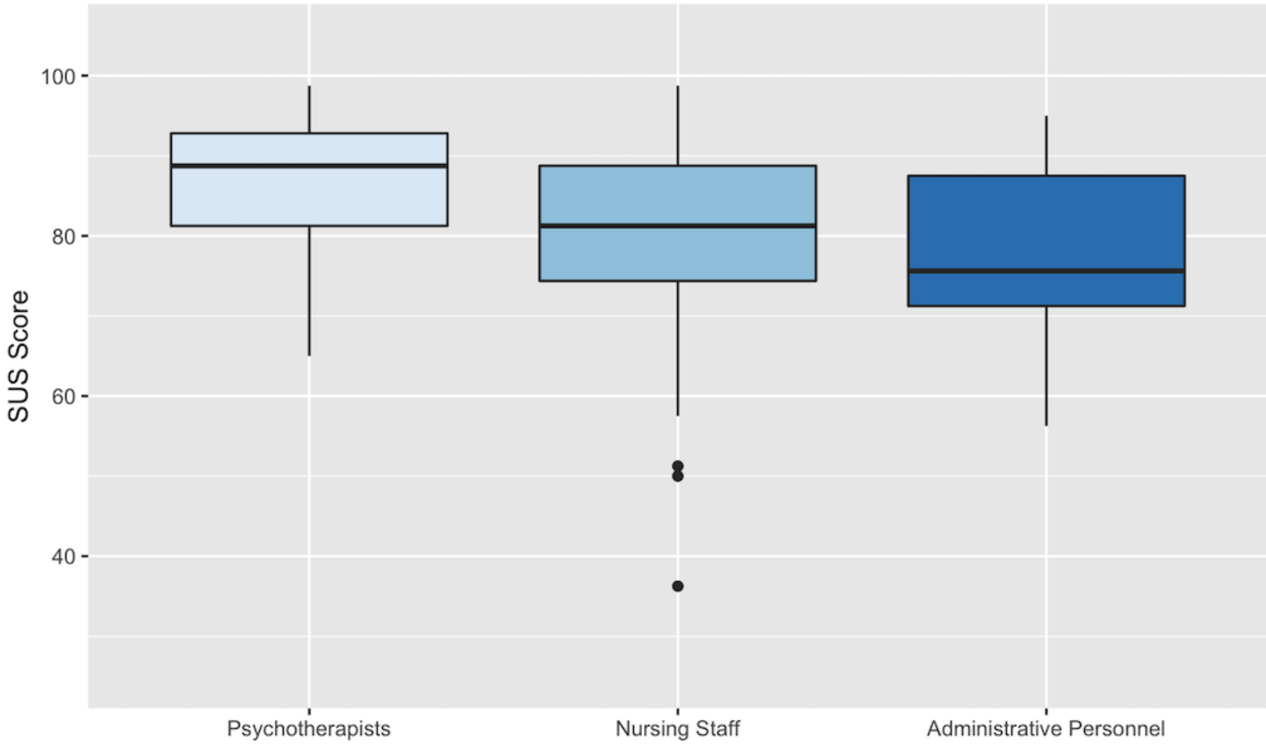
[b]SUS: System Usability Scale.

[c]Posthoc analysis with Bonferroni correction: psychotherapists scores were greater than nursing staff scores and administrative personnel scores.

Evaluation Outcomes, System Usability Scale

There were no missing items; therefore, no imputation of values was necessary. The SUS scores were normally distributed with few outliers. The mean SUS global score was 81.49 (SD 11.10). The mean score for the VRAT-CT was 82.00 (SD 12.55), and the mean for the VRAT-CC 80.99 (SD 12.67). Male participants scored slightly higher (mean 81.71, SD 15.24) than female participants (mean 81.39, SD 11.19), but this difference was not statistically significant (*P*=.11). Among the professional groups, psychotherapists (mean 86.44, SD 8.79) scored higher than nursing staff (mean 79.01, SD 13.30) and administrative personnel (mean 77.98, SD 10.72). The difference between psychotherapists and other professional groups reached statistical significance ($F_{2,106}$=6.136; *P*=.003) (Table 1 and Figure 2).

Figure 2. System Usability Scale (SUS) scores. Psychotherapists scored significantly higher than nursing staff and administrative personnel (P=.003).



Multilevel Factorial Analysis

The System Usability Scale produced a Cronbach α=.80 with good correlation between single items. Item loadings ranged from 0.2 to 0.8 (chi-square P<.001; comparative fit index 0.905; sample-size adjusted Bayesian information criterion 5028.08; root mean square error of approximation 0.075) and demonstrated different factor structures for each profession (Table 2).

Table 2. Mean score on the System Usability Scale and loadings for each item.

Item	Profession						Statistics	
	Psychotherapists		Nursing staff		Administrative personnel			
	Loadings	Mean (SD)	Loadings	Mean (SD)	Loadings	Mean (SD)	*F* test (*df1*,*df2*)	*P* value
1	0.489	4.42 (0.67)[b]	0.356	4.16 (0.79)	0.396	3.92 (1.04)	6.048 (2, 215)	.003
2	0.591	1.25 (0.46)[b]	0.670	1.35 (0.59)[c]	0.687	1.60 (0.66)	5.993 (2, 215)	.003
3	0.672	4.45 (0.57)	0.665	4.28 (1.01)	0.565	4.33 (0.86)	1.929 (2, 205)	.15
4	0.520	2.00 (1.07)[a,b]	0.428	2.49 (1.26)	0.328	2.81 (1.28)	97.663 (2, 215)	<.001
5	0.500	4.44 (0.65)	0.405	4.31 (0.58)	0.575	4.38 (0.53)	0.896 (2, 215)	.41
6	0.523	1.30 (0.54)[a,b]	0.513	1.73 (0.90)	0.579	1.79 (0.64)	10.070 (2, 215)	<.001
7	0.335	4.31 (0.76)	0.805	4.24 (0.85)	0.554	4.08 (0.62)	1.351 (2, 215)	.22
8	0.501	1.30 (0.58)[a]	0.506	1.71 (1.13)	0.793	1.40 (0.57)	4.946 (2, 215)	.008
9	0.561	4.22 (0.73)[b]	0.523	3.90 (1.01)	0.5552	3.81 (0.97)	4.219 (2, 215)	.02
10	0.291	1.50 (0.78)[b]	0.566	2.02 (1.13)[c]	0.535	1.73 (0.79)	6.519 (2, 215)	.002

[a]Posthoc analysis with Bonferroni correction: psychotherapists scores were greater than nursing staff scores and administrative personnel scores.

[b]Posthoc analysis with Bonferroni correction: psychotherapists scores were greater than administrative personnel scores.

[c]Posthoc analysis with Bonferroni correction: nursing staff scores were greater than administrative personnel scores.

Discussion

The usability of a digital psychotherapeutic tool developed using a game engine was studied in different professions. The SUS score obtained for the virtual reality human–human interface achieved the benchmark for an excellent system [21,22], scoring highest among the professional target group. The sample's demographic characteristics did not affect these results: the SUS scores were similar regardless of age or gender. To the best of our knowledge, this is the first study to assess the usability of a psychotherapeutic VR treatment tool for people experiencing acoustic verbal hallucinations.

Our study's main strengths are the large sample size and the naturalistic design, particularly the personalized introduction and the practical hands-on approach to the system [14,15]. We chose this approach to emulate the introduction and instruction used by psychotherapists in research and clinical practice. Furthermore, through the personalized introduction to the system, we sought to compensate for differences in the theoretical backgrounds of the professional groups.

The SUS was originally developed to evaluate the usability of products and services, including hardware, software, mobile devices, websites, and apps. Because the product was a combination of these elements and has previously been used to test medical devices and products [23-25], we selected the SUS to enable easy comparison with both similar and dissimilar products or devices [21]. The digital therapeutic system tested yielded a score ranging from good to excellent, depending on professional background [21,22].

The virtual reality human–human interface achieved a higher score among professionals with psychotherapeutic backgrounds (ie, psychiatrists and psychologists). Since all participants were naïve to the system, differences cannot be attributed to user experience [26]. In our opinion, these differences underscore the need for relevant training and professional background in order to fully understand and use the virtual reality therapeutic tool that we have created [14]. Posthoc analysis revealed no differences between psychiatrists and psychologists. We, therefore, consider both to form a uniform group with psychotherapeutic training as the common factor (in Switzerland) [27-29]. In addition, the similarities between psychologists and psychiatrists, regarding educational level and awareness of relevant research, should also be taken into account [30].

The SUS scale was designed as a global measure of perceived usability. Attempts thus far to identify an underlying factor analysis have been misleading and mainly reflected its alternating structure [31]. Nonetheless, we analyzed the SUS at an item level to discern differences potentially attributable to the skill-grade mix of the participants. The SUS scale yielded similar scores between psychotherapists and nonpsychotherapists only for items 3, 5, and 7. These items are more closely related to the handling of the system than to its actual implementation and use in research and clinical practice. The SUS scale also has a different factor structure for each profession, indicating different evaluation patterns for the usability of the system. This leads us to believe that the system is generally easy to use, allowing therapists to quickly become familiar with it and develop confidence, thereby increasing the likelihood of incorporating this system into their therapeutic repertoire and using it to deliver therapy [15,32].

Our study has several limitations that must be acknowledged. First, we did not include a clinical population. Although people experiencing verbal acoustic hallucinations were involved in the development process [16], they were not systematically involved in evaluating the system. At this stage, we chose to focus on the therapeutic end user since they would be responsible for introducing and guiding patients through the system and conducting the therapy sessions afterward. Another factor in our study was the use of only a single session for evaluation. This approach was chosen with the intention of assessing the intuitive usability of the system and to avoid learning effects. We did not compare our therapeutic system with those based on other technical possibilities, such as a non-VR presentation of the Avatar or the use of mobile or handheld devices. It is possible that such technical alternatives may yield a higher usability score. Finally, although no discomfort or side effects were reported, we did not systematically assess these important issues related to the use of VR technology [16].

We were able to demonstrate that a virtual reality human–human interface for research and clinical practice can be developed using an existing and widely available game engine. The results show that the usability of the digital therapeutic tools depends not only on the system itself but also on the user's professional background. We believe this system may enable and encourage psychotherapists to expand their therapeutic skills, to routinely using this technology in research and clinical practice [13,33-35]. In summary, given the high usability scores, gaming technology and platforms seem to be suitable for the creation and customization of novel therapeutic approaches in psychiatry.

Acknowledgments

We thank Mrs Lorna McBroom for manuscript proofreading and language editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Virtual reality human–human interface.
[MP4 File (MP4 Video), 41377 KB - games_v9i2e26820_app1.mp4]

References

1. Lazar SG. The cost-effectiveness of psychotherapy for the major psychiatric diagnoses. Psychodyn Psychiatry 2014 Sep;42(3):423-457. [doi: 10.1521/pdps.2014.42.3.423] [Medline: 25211432]
2. Schramm E, Gerardi M, Rothbaum B, Berger M. Role of psychotherapy in the management of psychiatric diseases. Handb Clin Neurol 2012;106:633-641. [doi: 10.1016/B978-0-444-52002-9.00037-1] [Medline: 22608648]
3. Craighead WE, Craighead LW. The role of psychotherapy in treating psychiatric disorders. Med Clin North Am 2001 May;85(3):617-629. [doi: 10.1016/s0025-7125(05)70332-1] [Medline: 11349476]
4. Bighelli I, Huhn M, Schneider-Thoma J, Krause M, Reitmeir C, Wallis S, et al. Response rates in patients with schizophrenia and positive symptoms receiving cognitive behavioural therapy: a systematic review and single-group meta-analysis. BMC Psychiatry 2018 Dec 04;18(1):380 [FREE Full text] [doi: 10.1186/s12888-018-1964-8] [Medline: 30514268]
5. Bighelli I, Salanti G, Huhn M, Schneider-Thoma J, Krause M, Reitmeir C, et al. Psychological interventions to reduce positive symptoms in schizophrenia: systematic review and network meta-analysis. World Psychiatry 2018 Oct;17(3):316-329 [FREE Full text] [doi: 10.1002/wps.20577] [Medline: 30192101]
6. NICE. Psychosis and schizophrenia in adults: prevention and management. NICE Guidelines CG178 2014:1-46. [Medline: 32207892]
7. American Psychiatric Association. The American Psychiatric Association Practice Guideline for the Treatment of Patients With Schizophrenia. Washington, DC: American Psychiatric Association Publishing; 2020.
8. Riva G. Virtual reality in psychotherapy: review. Cyberpsychol Behav 2005 Jun;8(3):220-30; discussion 231. [doi: 10.1089/cpb.2005.8.220] [Medline: 15971972]
9. Rizzo AS, Koenig ST, Talbot TB. Clinical virtual reality: emerging opportunities for psychiatry. Focus (Am Psychiatr Publ) 2018 Jul;16(3):266-278 [FREE Full text] [doi: 10.1176/appi.focus.20180011] [Medline: 31975921]
10. Bell IH, Nicholas J, Alvarez-Jimenez M, Thompson A, Valmaggia L. Virtual reality as a clinical tool in mental health research and practice . Dialogues Clin Neurosci 2020 Jun;22(2):169-177 [FREE Full text] [doi: 10.31887/DCNS.2020.22.2/lvalmaggia] [Medline: 32699517]
11. Leff J, Williams G, Huckvale MA, Arbuthnot M, Leff AP. Computer-assisted therapy for medication-resistant auditory hallucinations: proof-of-concept study. Br J Psychiatry 2013 Jun;202:428-433 [FREE Full text] [doi: 10.1192/bjp.bp.112.124883] [Medline: 23429202]
12. Craig TK, Rus-Calafell M, Ward T, Leff JP, Huckvale M, Howarth E, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. Lancet Psychiatry 2018 Dec;5(1):31-40 [FREE Full text] [doi: 10.1016/S2215-0366(17)30427-3] [Medline: 29175276]
13. Mishkind MC, Norr AM, Katz AC, Reger GM. Review of virtual reality treatment in psychiatry: evidence versus current diffusion and use. Curr Psychiatry Rep 2017 Sep 18;19(80):1-8. [doi: 10.1007/s11920-017-0836-0] [Medline: 28920179]
14. Rus-Calafell M, Garety P, Sason E, Craig TJK, Valmaggia LR. Virtual reality in the assessment and treatment of psychosis: a systematic review of its utility, acceptability and effectiveness. Psychol Med 2018 Feb;48(3):362-391. [doi: 10.1017/S0033291717001945] [Medline: 28735593]
15. Schwartzman D, Segal R, Drapeau M. Perceptions of virtual reality among therapists who do not apply this technology in clinical practice. Psychol Serv 2012 Aug;9(3):310-315. [doi: 10.1037/a0026801] [Medline: 22867123]
16. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]
17. Leff J, Williams G, Huckvale M, Arbuthnot M, Leff AP. Avatar therapy for persecutory auditory hallucinations: what is it and how does it work? Psychosis 2014 Jun;6(2):166-176 [FREE Full text] [doi: 10.1080/17522439.2013.773457] [Medline: 24999369]
18. du Sert OP, Potvin S, Lipp O, Dellazizzo L, Laurelli M, Breton R, et al. Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: a pilot clinical trial. Schizophr Res 2018 Feb 24:176-181. [doi: 10.1016/j.schres.2018.02.031] [Medline: 29486956]
19. Bonfiglio N. DeepMind partners with gaming company for AI research. The Daily Dot. 2018. URL: https://www.dailydot.com/debug/unity-deempind-ai/ [accessed 2020-11-20] [WebCite Cache ID https://www.dailydot.com/debug/unity-deempind-ai/]
20. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation in Industry. Boca Raton, Florida: CRC Press; 1996:189.
21. Kortum PT, Bangor A. Usability ratings for everyday products measured with the System Usability Scale. Int J Hum Comput Interact 2013 Jan;29(2):67-76. [doi: 10.1080/10447318.2012.681221]
22. Lewis JR, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018;13(3):158-167.
23. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Dec;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]

24. Zhang T, Booth R, Jean-Louis R, Chan R, Yeung A, Gratzer D, et al. A primer on usability assessment approaches for health-related applications of virtual reality. JMIR Serious Games 2020 Oct 28;8(4):e18153 [FREE Full text] [doi: 10.2196/18153] [Medline: 33112243]
25. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1-19. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
26. McLellan S, Muddimer A, Peres S. The effect of experience on System Usability Scale ratings. J Usability Stud 2012;7(2):56-67.
27. Caspar F. The current status of psychotherapy integration in Germany and Switzerland. J Psychother Integr 2008;18(1):74-78. [doi: 10.1037/1053-0479.18.1.74]
28. Rubo M, Martin-Soelch C, Munsch S. Education and training in clinical psychology and psychological psychotherapy in Switzerland. Clin Psychol Eur 2020 Sep 30;2(3):1-11. [doi: 10.32872/cpe.v2i3.2991]
29. Oakley C, Malik A. Psychiatric training in Europe. Psychiatrist 2018 Jan 02;34(10):447-450. [doi: 10.1192/pb.bp.109.026062]
30. Swanson JA, Walker E. Academic versus non-academic emerging adult college student technology use. Tech Know Learn 2015 May 7;20(2):147-158. [doi: 10.1007/s10758-015-9258-4]
31. Lewis J, Sauro J. Revisiting the factor structure of the System Usability Scale. J Usability Stud 2017;12(4):183-192. [doi: 10.1007/978-3-642-02806-9_12]
32. Spurgeon JA, Wright JH. Computer-assisted cognitive-behavioral therapy. Curr Psychiatry Rep 2010 Dec;12(6):547-552. [doi: 10.1007/s11920-010-0152-4] [Medline: 20872100]
33. Segal R, Bhatia M, Drapeau M. Therapists' perception of benefits and costs of using virtual reality treatments. Cyberpsychol Behav Soc Netw 2011;14(1-2):29-34. [doi: 10.1089/cyber.2009.0398] [Medline: 21329440]
34. Wampold BE. How important are the common factors in psychotherapy? an update. World Psychiatry 2015 Oct;14(3):270-277 [FREE Full text] [doi: 10.1002/wps.20238] [Medline: 26407772]
35. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2006 Nov 20;37(03):319-328. [doi: 10.1017/s0033291706008944]

Abbreviations

SUS: System Usability Scale
VR: virtual reality
VRAT: Virtual Reality Avatar-Creation Tool

Edited by N Zary; submitted 29.12.20; peer-reviewed by J Brooke, R Lundin; comments to author 08.02.21; revised version received 05.03.21; accepted 19.03.21; published 01.06.21.

Please cite as:
Brander M, Egger ST, Hürlimann N, Seifritz E, Sumner RW, Vetter S, Magnenat S
Virtual Reality Human–Human Interface to Deliver Psychotherapy to People Experiencing Auditory Verbal Hallucinations: Development and Usability Study
JMIR Serious Games 2021;9(2):e26820
URL: https://games.jmir.org/2021/2/e26820
doi:10.2196/26820
PMID:33769295

Original Paper

Impact of Gamification on the Self-Efficacy and Motivation to Quit of Smokers: Observational Study of Two Gamified Smoking Cessation Mobile Apps

Nikita B Rajani[1], MPH; Nikolaos Mastellos[1], PhD; Filippos T Filippidis[1], PhD

Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Nikita B Rajani, MPH
Department of Primary Care and Public Health
Imperial College London
School of Public Health
St. Dunstan's Road
London, W6 8RP
United Kingdom
Phone: 44 7427615928
Email: nikita.rajani14@imperial.ac.uk

Abstract

Background: The proportion of smokers making quit attempts and the proportion of smokers successfully quitting have been decreasing over the past few years. Previous studies have shown that smokers with high self-efficacy and motivation to quit have an increased likelihood of quitting and staying quit. Consequently, further research on strategies that can improve the self-efficacy and motivation of smokers seeking to quit could lead to substantially higher cessation rates. Some studies have found that gamification can positively impact the cognitive components of behavioral change, including self-efficacy and motivation. However, the impact of gamification in the context of smoking cessation and mobile health has been sparsely investigated.

Objective: This study aims to examine the association between perceived usefulness, perceived ease of use, and frequency of use of gamification features embedded in smoking cessation apps on self-efficacy and motivation to quit smoking.

Methods: Participants were assigned to use 1 of the 2 mobile apps for a duration of 4 weeks. App-based questionnaires were provided to participants before app use and 2 weeks and 4 weeks after they started using the app. Gamification was quantitatively operationalized based on the Cugelman gamification framework and concepts from the technology acceptance model. The mean values of perceived frequency, ease of use, and usefulness of gamification features were calculated at midstudy and end-study. Two linear regression models were used to investigate the impact of gamification on self-efficacy and motivation to quit.

Results: A total of 116 participants completed the study. The mean self-efficacy increased from 37.38 (SD 13.3) to 42.47 (SD 11.5) points and motivation to quit increased from 5.94 (SD 1.4) to 6.32 (SD 1.7) points after app use. *Goal setting* was perceived to be the most useful gamification feature, whereas *sharing* was perceived to be the least useful. Participants self-reported that they used the progress dashboards the most often, whereas they used the sharing feature the least often. The average perceived frequency of gamification features was statistically significantly associated with change in self-efficacy (β=3.35; 95% CI 0.31-6.40) and change in motivation to quit (β=.54; 95% CI 0.15-0.94) between baseline and end-study.

Conclusions: Gamification embedded in mobile apps can have positive effects on self-efficacy and motivation to quit smoking. The findings of this study can provide important insights for tobacco control policy makers, mobile app developers, and smokers seeking to quit.

(JMIR Serious Games 2021;9(2):e27290) doi:10.2196/27290

KEYWORDS

gamification; smoking cessation; mobile applications; self-efficacy; motivation to quit; mHealth; mobile phone

Introduction

Smoking is the second leading risk factor for early death and disability worldwide, with approximately 8 million deaths annually [1,2]. In the United Kingdom, 16% of all deaths in 2016 were attributed to smoking [3]. Despite most smokers wanting to quit, research shows that the number of smokers in the United Kingdom who tried to quit in the past year dropped by 7% between 2008 and 2017 [4]. Among the smokers making attempts, only 3%-5% are successful in staying quit after a year [5]. Exploring methods that can improve the success rates of long-term quitting can lead to a decreased prevalence of chronic diseases and premature mortality. Previous studies have shown that high motivation to quit and self-efficacy, 2 factors often integrated into face-to-face behavioral support interventions, have been found to increase the likelihood of attempting to quit and successfully quitting [6-9].

Motivation to quit refers to the level of determination and importance placed by an individual on quitting smoking [10]. On the other hand, self-efficacy, a theoretical construct first termed by Bandura [11], in the context of smoking cessation refers to one's confidence in their ability to refrain from smoking when faced with internal and external stimuli [11,12]. Several strategies are adopted by health behavior change interventions to influence self-efficacy or motivation to quit. For example, motivational interviewing is often used to help smokers explore reasons for quitting and "make them feel more willing and able to stop smoking" [13]. Similarly, vicarious experience and performance attainment are strategies to influence self-efficacy. Vicarious experience involves "exposing the individual to successful behaviour performances or gaining experience through practice" [14]. In the context of smoking cessation, this could mean showing smokers examples of other smokers who successfully quit after attending cessation programs. Performance attainment refers to having successful experiences [11,14]; for a smoker trying to quit, this could mean staying abstinent for a day and recognizing this as a success. Such strategies have been integrated into both face-to-face and digital interventions, such as mobile health (mHealth) solutions, which have become increasingly important to improve access, knowledge, and behavior across different contexts and population groups [15].

One strategy that has been frequently applied across both physical and remote interventions for behavioral change is gamification, also known as the use of game elements in a nongame context [16]. Some examples of game elements include achievement badges, goal setting, progress tracking, sharing progress, and levels [17]. Past studies have found that gamification can positively impact cognitive components of behavioral change, including self-efficacy and motivation. For example, self-efficacy was the most cited advantage of gamified classrooms, as it improved the confidence levels of students [18]. Similarly, Thorsteinsen et al [19] found that gaming elements significantly increased the motivation of individuals to engage in physical activity. The use of gamification has gradually become more popular as it appears to share key components with several behavioral change theories and techniques [17,20]. For example, self-determination theory, a dominant theory of motivation, suggests that gamification elements such as points and badges serve as informational feedback instilling a sense of intrinsically motivating competence in the user [21]. Similarly, goal-setting theory, another prominent theory of motivation, has been associated with gamification; for example, elements such as leaderboards and levels provide individuals with smaller, more immediate goals that can improve task performance and, in turn, increase one's confidence in their ability to complete tasks [21,22].

Although gamification seems promising, the impact of gamification in the field of health behavior change has focused primarily on improving physical activity levels [23-25]. More specifically, there is room to further study the use of gamification in the context of other behaviors, such as smoking. With the proliferation of smartphone use and the increased availability of digitalized interventions, it is vital to investigate novel strategies, such as gamification, that complement mHealth solutions and are in line with current digitalization trends. The majority of studies that have explored the impact of gamification in the context of smoking cessation have been purely qualitative and consequently have not operationalized gamification quantitatively. For example, Pløhn and Aalberg [26] interviewed participants after they used a digital smoking cessation intervention with gamification features and found positive perceptions of gamification as an important motivational factor to aid quitting. However, the smoking cessation intervention was not delivered via a mobile app. Another study by El-Hilly et al [27] found promising results on the effect of gamification on the motivation and engagement of smokers within the context of mHealth. Similar to the study by Pløhn and Aalberg [26], the study by El-Hilly et al [27] was also qualitative and had a small sample size (n=16), hindering the external validity of the findings. On the other hand, Lin et al [28] investigated gamification quantitatively and found that program progress or step completion as a gamification element in a smoking cessation mobile app had a positive impact on psychological factors such as user well-being, inspiration, and empowerment [28]. It would be worthwhile to investigate whether gamification in smoking cessation apps can also positively impact essential cognitive factors such as self-efficacy and motivation to quit.

Exploring whether gamification in the context of mHealth can increase the self-efficacy and motivation of smokers could lead to the design of more tailored interventions, which, in turn, could improve cessation rates and reduce the health burden of tobacco consumption. The findings of this study could also provide insights into the effective design of mobile apps. Moreover, because of the wide reach and low cost of mHealth solutions, understanding the effects of gamification on mHealth could have considerable effects on helping disadvantaged groups and reducing health inequalities [29]. On the basis of the limitations of prior research, this study's aim of exploring gamified smoking cessation apps can also help enhance the current understanding of the effectiveness of mHealth interventions for behavior change and extend our knowledge of novel methods to promote healthy lifestyle changes. Specifically, our study aims to quantitatively assess the association between overall perceived usefulness, ease of use, and frequency of use of gamification features and the self-efficacy and motivation of smokers seeking to quit.

Methods

Sampling and Eligibility

On the basis of an a priori analysis using a power level of 1−β=.80 and a significance level α=.05, we aimed to recruit 140 participants. The results of a previous study examining the impact of gamification elements in a fitness app on perceived competence, a proxy for self-efficacy, indicated that 112 participants were required to complete the study [25]. Similar to previous mobile app studies, we considered a dropout rate of approximately 20% [30-32], which resulted in the recruitment of 140 participants. The sample size calculation also assumes that both apps are similar based on a prior mobile app review [33].

Participants were required to be at least 18 years old and current smokers (at least one cigarette a day and 100 cigarettes smoked in their lifetime) to be eligible. Moreover, to take part in the study, individuals had to report that they were trying or willing to quit smoking in the next 30 days and were not using other forms of smoking cessation treatments. Individuals diagnosed with mental health conditions were excluded from the study.

Study Design

A 4-week observational study assessing the association between gamification, self-efficacy, and motivation to quit was conducted from June 2019 to July 2020. No face-to-face contact was required, and the study was conducted on the internet. Participants were recruited via social media, and posters were displayed in public places in London. Initially, participants who expressed interest in the study (N=326) were screened to assess their eligibility. Eligible participants provided informed consent (n=170) and were assigned a participant identification number (PID). They were then requested to complete a baseline questionnaire that asked about general demographics (age, gender, education, marital status, education, country of residence, etc), smoking habits (number of cigarettes smoked, nicotine dependence, etc), self-efficacy, and motivation to quit.

In total, 154 participants completed the baseline assessment and were provided instructions on how to download and start using the app. Even-numbered PIDs were assigned to the mobile app Quit Genius, and odd-numbered PIDs were assigned to the mobile app Kwit. This deterministic method was used to ensure an equal split of participants between the 2 apps. Participants were asked to use the assigned mobile app for a total of 4 weeks. A midstudy questionnaire after 2 weeks of using the app and an end-study questionnaire after 4 weeks of using the app were given to participants.

Of those participants who completed the baseline assessment, 138 installed the app and 116 completed all 4 weeks of the study. Midstudy and end-study assessments included questions regarding gamification, self-efficacy, and motivation to quit. Participants were incentivized via free access to all features of the app and a chance to win a £50 (US $68) Amazon voucher. An overview of the number of participants at each stage of the study is presented in Multimedia Appendix 1.

Mobile Apps

Mobile apps for the study were selected based on a mobile app review that found these 2 apps to have a high embedment of gamification features and a high adherence to smoking cessation guidelines in the United Kingdom [33,34]. Screenshots of both apps are shown in Multimedia Appendix 2.

Kwit

Kwit is a smoking cessation mobile app that helps individuals starting their quit journey and individuals trying to stay quit [35]. The app includes several features such as a calculator, a smoking diary that helps smokers log and analyze cravings and triggers, motivation cards, social media sharing, levels, and achievement cards. The app is based on cognitive behavioral therapy (CBT) and gamification strategies. The versions of Kwit used during the study period included those released from June 2019 (v.4.1) to July 2020 (v.4.4).

Quit Genius

Quit Genius is a gamified smoking cessation mobile app based on CBT [36]. It delivers personalized support to individuals seeking to quit smoking and helps quitters maintain their quit status. It includes several features such as a tracker, a daily diary that allows quitters to log their cravings and triggers, a quitting toolbox, a goal-setting feature, achievement badges, stages of information that build upon each other, and a quit coach who provides continuous personalized support. The versions downloaded by participants were those released from June 2019 (v.1.1) to July 2020 (v.1.9).

Measures

Sociodemographic Factors

Common sociodemographic factors were assessed at baseline. Age in years was categorized as 18-29, 30-41, 42-53, or 54-65 years, and gender was categorized as male or female. Marital status was categorized as single or married or civil partnered. Education was based on United Nations Educational, Scientific and Cultural Organization's classification into 3 categories: low if primary school was completed, medium if secondary school was completed, and high if a college or university degree was attained [37]. Employment status was categorized as unemployed (individuals who are willing and able to work but have no employment), employed, or nonemployed (individuals who are unable to work, including students and homemakers). Residence was categorized based on the World Health Organization regions: Western Pacific, Americas, Southeast Asia, Europe, Africa, and Eastern Mediterranean [38].

Nicotine Dependence

The Fagerström test with 6 items was used to measure participants' tolerance of and dependence on nicotine [36]. On the basis of the responses, participants were categorized into 3 levels: low (0-4 points), moderate (5-7 points), and high (8-10 points) [39,40].

Self-Efficacy

The self-efficacy of a participant was measured using a 12-item scale called The Smoking Self-Efficacy Questionnaire [12]. The scale assesses an individual's confidence in their ability to refrain from smoking when faced with internal and external stimuli. Response options included not at all sure, not very sure, more or less sure, fairly sure, and absolutely sure. A total score ranging from 12 to 60 was computed for each participant, with higher scores signifying higher self-efficacy.

Motivation to Quit

Participants were asked 2 items frequently used in cessation studies to measure their motivation to quit smoking [10,41,42]. The first item asked was as follows: How important is it to you to give up smoking altogether at this attempt? Responses included the following: desperately important, very important, quite important, and not all that important. The second item asked was as follows: How determined are you to give up smoking at this attempt? Response options included the following: extremely determined, very determined, quite determined, and not all that determined. A total score ranging from 2 to 8 was calculated for each participant, with higher scores signifying higher motivation.

Gamification

Gamification features for each app were identified using Cugelman framework for gamification strategies and tactics and are displayed in Multimedia Appendix 3 [17]. For each identified gamification feature, participants were asked how useful and easy to use they found it during their quit attempt. Participants were provided with 5-point Likert scale responses: strongly agree, agree, neither agree nor disagree, disagree, and strongly disagree. Perceived usefulness and ease of use are 2 vital components of the technology acceptance model, which has been widely used in existing literature to better understand user acceptance and attitudes toward mobile apps and app features [43,44].

Participants were also asked how frequently they engaged with each gamification element or feature during their quit attempt. Participants were provided with 5-point Likert scale responses: almost always, often, sometimes, rarely, and never. Responses were assigned points ranging from 1 to 5 for each gamification feature. A pooled mean was calculated for all features, with a higher mean (from 1 to 5) indicating greater overall engagement with gamification.

Statistical Analysis

The statistical software STATA 13.1 (StataCorp), was used for the analyses. Box and whisker plots were created to present an overview of self-efficacy and motivation to quit levels of participants at various study time points. The mean values of perceived frequency, ease of use, and usefulness of gamification features were calculated at midstudy and end-study. Two-way paired sample *t* tests were used to test whether differences in self-efficacy and motivation to quit at various time points of the study were statistically significant. In addition, we explored various linear regression models to examine whether gamification was associated with changes in self-efficacy and motivation to quit. On the basis of an iterative process that considered the fit of the data with our model (ie, comparing the Akaike information criterion and Bayesian information criterion), 2 linear regression models were performed. The first tested the association between perceived ease of use, frequency of use, and usefulness of gamification with change in self-efficacy, and the second tested the association between perceived ease of use, frequency of use, and usefulness of gamification with change in motivation to quit. Both models controlled for age, gender, education, marital status, nicotine dependence, baseline self-efficacy, and baseline motivation to quit. Significance at the 5% level (0.05), along with 95% CIs for all included coefficients, is presented in the *Results* section.

Results

Study Participants

As shown in Table 1, there was an equal split of participants who used the apps Kwit (58/116, 50%) and Quit Genius (58/116, 50%). The majority of participants were male (71/116, 61.2%), highly educated (87/116, 75%), single (77/116, 66.4%), employed (76/116, 65.6%), and living in Europe (67/116, 57.8%). More than half of the participants smoked 10 or fewer cigarettes on a daily basis (63/116, 54.3%), and the majority had low to moderate dependence on nicotine (107/116, 92.2%).

Table 1. Sociodemographic and general characteristics of participants (n=116).

Characteristics	Respondents, n (%)
Assigned mobile app	
Kwit	58 (50)
Quit Genius	58 (50)
Age (years)	
18-29	49 (42.2)
30-41	41 (35.3)
42-53	15 (12.9)
54-65	11 (9.5)
Gender	
Male	71 (61.2)
Female	45 (38.8)
Education	
Low (primary school)	8 (6.9)
Medium (secondary school)	21 (18.1)
High (university or college degree)	87 (75)
Marital status	
Single	77 (66.4)
Married or civil partnered	39 (33.6)
Employment status	
Employed	76 (65.6)
Nonemployed	31 (26.7)
Unemployed	6 (5.2)
Prefer not to answer	3 (2.6)
World Health Organization regions	
Western Pacific	4 (3.4)
Americas	10 (8.6)
Southeast Asia	16 (13.8)
Europe	67 (57.8)
Africa	17 (14.7)
Eastern Mediterranean	2 (1.7)
Daily smoking (number of cigarettes)	
10 or less	63 (54.3)
11-20	43 (37.1)
21-30	8 (6.9)
31 or more	2 (1.7)
Fagerström nicotine dependence	
Low (0-4)	62 (53.4)
Moderate (5-7)	45 (38.8)
High (8-10)	9 (7.8)

Self-Efficacy and Motivation to Quit

Figure 1 shows that the mean motivation to quit increased from 5.94 at baseline to 6.20 after 2 weeks of app use and to 6.32 points after 4 weeks of app use. The median motivation to quit (and IQR) at baseline was 6.00 (IQR 5-7), 6.00 (IQR 5-8) at 2 weeks, and 7.00 (IQR 5-8) at 4 weeks. Similarly, the mean self-efficacy score increased from 37.38 points at baseline to 41.37 points after 2 weeks of using the app and then to 42.47 points after 4 weeks of using the app. Median self-efficacy (and IQR) at baseline was 37.00 (IQR 27-39), which increased to 42.50 points (IQR 33-50) at 2 weeks and increased further to 44.00 (IQR 35-51) at 4 weeks. Paired *t* tests found that increases in mean self-efficacy and motivation to quit between baseline and midstudy and baseline and end-study were statistically significant (Multimedia Appendix 4).

Figure 1. Self-efficacy and motivation to quit at baseline, midstudy, and end-study (n=116).

Gamification

Table 2 displays the average midstudy and end-study perceived usefulness, ease of use, and frequency of use of overall gamification and specific gamification features embedded in the apps. At midstudy and end-study, goal setting was perceived to be the most useful gamification feature (4.14 score out of 5), whereas sharing was perceived to be the least useful feature (3.72 at midstudy and 3.28 at end-study out of 5). Participants also perceived goal setting to be the easiest to use feature at both midstudy and end-study (4.31 and 4.36, respectively). In terms of frequency of use, participants self-reported that they used the progress dashboards the most often during both the midstudy and end-study assessments (3.23 and 3.30, respectively). The feature reported to be used the least frequently was the sharing feature, as not many participants shared their progress or results with others.

The linear regression model results presented in Table 3 show that a 1-point increase in average perceived frequency of gamification features was statistically significantly associated with a 3.35-point increase in self-efficacy from baseline to end-study (β=3.35; 95% CI 0.31-6.40). The average perceived ease of use and usefulness of gamification were not associated with changes in self-efficacy. In addition, a 1-point increase in baseline self-efficacy was associated with a 1.06-point decrease in self-efficacy between baseline and end-study (β=−1.06; 95% CI −1.22 to −0.90). Gender, marital status, nicotine dependence, baseline motivation, average perceived ease of use, and usefulness of gamification features were not statistically significantly associated with changes in self-efficacy from baseline to end-study. The second linear regression model presented in Table 3 shows that individuals with medium or high education had, on average, a 1.31 point (95% CI −2.60 to −0.01) and 1.21 point (95% CI −2.32 to −0.10) lower motivation to quit than individuals with a low level of education. Moreover, a 1-point increase in average perceived frequency of use of gamification features was statistically significantly associated with a 0.54-point increase in motivation to quit at end-study compared with baseline (β=.54; 95% CI 0.15-0.94). Similarly, there was some indication that the average usefulness of gamification and baseline self-efficacy are associated with change in motivation to quit. Finally, a 1-point increase in baseline motivation to quit was statistically significantly associated with a 0.69-point decrease in motivation to quit (β=−.69; 95% CI −0.90 to −0.49).

Table 2. Overview of perceived frequency of use, ease of use, and usefulness of gamification features embedded in the Kwit and Quit Genius apps (n=116).

Gamification features	Midstudy, mean (SD)			End-study, mean (SD)		
	Perceived usefulness[a]	Perceived ease of use[a]	Perceived frequency of use[a]	Perceived usefulness[a]	Perceived ease of use[a]	Perceived frequency of use[a]
Logging diaries	3.78 (0.99)	4.08 (0.95)	3.13 (1.21)	3.85 (0.98)	3.85 (0.97)	3.19 (1.20)
Achievements and badges	3.64 (1.11)	3.91 (0.96)	2.90 (1.27)	3.78 (1.06)	3.97 (1.07)	2.96 (1.23)
Progress tracking	3.91 (0.94)	4.07 (0.86)	3.23 (1.11)	4.04 (0.93)	4.07 (0.96)	3.30 (1.21)
Unlocking levels or competing stages	3.93 (0.89)	4.01 (0.94)	3.03 (1.02)	3.94 (0.93)	4.18 (0.79)	3.09 (1.08)
Sharing feature	3.08 (1.15)	3.72 (0.87)	1.86 (1.13)	3.28 (1.17)	3.72 (0.95)	1.93 (1.16)
Motivation cards[b]	3.64 (0.95)	4.10 (0.79)	2.95 (1.26)	3.71 (1.12)	4.08 (0.98)	3.14 (1.23)
Goal setting[c]	4.14 (0.85)	4.31 (0.71)	2.64 (0.91)	4.14 (0.80)	4.36 (0.81)	2.97 (1.03)
Overall	3.71 (0.75)	4.00 (0.64)	2.83 (0.80)	3.80 (0.78)	4.04 (0.72)	2.92 (0.87)

[a]Range: 1-5.
[b]Only applicable to Kwit.
[c]Only applicable to Quit Genius.

Table 3. Linear regression model examining the association between perceived usefulness, ease of use, and frequency of use of gamification features with change in self-efficacy and change in motivation to quit (n=116).

Variables	Change in self-efficacy		Change in motivation to quit	
	β	95% CI	β	95% CI
Age (years)	−.05	−0.26 to 0.17	−.01	−0.04 to 0.02
Gender				
Male (referent)	Reference	Reference	Reference	Reference
Female	1.89	−2.48 to 6.26	.19	−0.37 to 0.76
Nicotine dependence				
Low (referent)	Reference	Reference	Reference	Reference
Moderate	−1.02	−5.50 to 3.46	−.08	−0.66 to 0.50
High	5.93	−2.15 to 14.02	.42	−0.61 to 1.46
Education				
Low (referent)	Reference	Reference	Reference	Reference
Medium	−4.95	−15.01 to 5.16	−1.31[a]	−2.60 to −0.01
High	−8.01	−16.63 to 0.61	−1.21[a]	−2.32 to −0.10
Marital status				
Single (referent)	Reference	Reference	Reference	Reference
Married	−.10	−5.35 to 5.17	−.03	−0.70 to 0.65
Mean frequency of gamification use	3.35[a]	0.31 to 6.40	.54[a]	0.15 to 0.94
Mean ease of use of gamification	−1.21	−5.16 to 2.74	−.03	−0.54 to 0.48
Mean usefulness of gamification	1.63	−2.53 to 5.79	.51	−0.03 to 1.04
Baseline self-efficacy	−1.06[a]	−1.22 to −0.90	−.02	−0.04 to −0.00
Baseline motivation to quit	1.14	−0.44 to 2.71	−.69[a]	−0.90 to −0.49
Constant	35.79[a]	15.68 to 53.89	3.19	0.73 to 5.65

[a]*P*<.05.

Discussion

Principal Findings

We found that the use of Kwit and Quit Genius was associated with increased self-efficacy and motivation to quit levels 4 weeks after app use compared with baseline. Our study also found that the perceived frequency of use of gamification features was associated with an increase in self-efficacy and motivation to quit. Finally, higher baseline self-efficacy and motivation to quit were both associated with smaller increases in self-efficacy and motivation to quit levels 4 weeks after using the mobile apps compared with preapp use.

The key finding from our analyses showed that the frequency of gamification use was associated with increased levels of self-efficacy and motivation to quit after app use compared with before app use. One possible reason for this could be that the frequency of gamification use has an effect on the overall user engagement with the app, which in turn influences the self-efficacy and motivation to quit levels. Some studies in the existing literature have found positive effects of gamification on user engagement. For example, Othman et al [45] found that gamification had a positive impact on user engagement with Play4fit, a fitness smartphone app. Similarly, Looyestyn et al [46] found that gamification was effective in increasing engagement levels with app-based programs. Therefore, higher overall app engagement as a result of gamification could have increased smokers' confidence in their ability to quit and the level of motivation to quit. Moreover, as higher engagement has been found to be positively associated with intervention effectiveness [47,48], it is possible that apps with gamification features that influence the overall engagement levels are associated with better cessation outcomes than those without gamification features.

Although not explicitly investigated in our study, it is also possible that the frequency of gamification use had an effect on user enjoyment, which in turn affected the motivation to quit and self-efficacy levels. Higher levels of user engagement could intrinsically influence motivation levels, as the use of the app could be rewarding or enjoyable for the user regardless of the final outcome. The theory of flow suggests that people can experience the state of *flow* in which they are highly involved in an activity because it is so enjoyable that they would engage in it even at a cost [49]. Research shows that gamification elements can enhance the level of enjoyment experienced, leading to higher levels of motivation [49]. Another possible theory-based explanation for why the frequency of gamification use was found to be associated with increased self-efficacy levels could be that gamification may influence a user's competence and confidence. Certain gamification features such as providing immediate feedback on performance, incremental levels, and providing badges of achievement could have provided a low risk way to attempt a task while also increasing confidence levels that their set goals are attainable [50]. According to the self-determination theory, the fulfillment of 3 types of psychological needs (autonomy, competence, and relatedness) can foster motivation [51]. By providing immediate feedback on performance through game elements such as badges, level advancements, and progress tracking, gamification may help fulfill competence needs and enhance self-efficacy. On the other hand, elements such as the sharing feature could help support and enhance the feeling of relatedness and in turn boost motivation levels.

The association between perceived frequency of use of gamification features and increases in motivation to quit and self-efficacy can have important implications for the use of gamification and game design principles in nongame contexts such as health behavior change. We found progress tracking to be the most frequently used gamification feature after 4 weeks of app use. According to a review of smoking cessation mobile apps, this feature was found to be most commonly integrated into apps by app developers [52]. However, we also found that one of the gamification features that users interacted with frequently (unlocking levels or completing stages) was also the feature that was adopted by only 20% of the smoking cessation apps investigated in the review. It could be valuable for app developers to investigate the impact of such gamification features as they are not often integrated into mobile apps but could have the potential to improve user engagement and thereby self-efficacy and motivation to quit. This also highlights the importance and need for collaboration between mobile app developers, researchers, and behavior change specialists to create interventions that can effectively target and influence vital cognitive factors via strategies such as gamification.

Our study also found some indications that the average perceived usefulness of gamification was associated with increased levels of self-efficacy and motivation to quit after 4 weeks of app use compared with baseline. This finding is in accordance with previous studies that have explored the impact of gamification on motivation to quit. For example, Pløhn and Aalberg [26] found that participants who managed to quit smoking after using a gamified app-based cessation program reported the effectiveness of gamification as a motivational factor [26]. Similarly, a study of 16 participants found that individuals who engaged in a gamified cessation intervention had higher levels of motivation than those who engaged with a nongamified cessation intervention [27]. Our findings, supported by other studies, highlight the value of further investigating the usefulness of specific gamification features to better design mHealth solutions geared toward facilitating health behavior change.

In addition to our findings on gamification, a general finding of our analyses was the increase in self-efficacy between baseline and 4 weeks after app use. This implies that participants experienced increased levels of confidence to refrain from smoking not only when faced with both internal stimuli such as cravings and emotions but also when faced with external stimuli such as being surrounded by other smokers or social situations that trigger smoking cravings. Likewise, the increase in motivation to quit between baseline and end-study suggests that participants experienced higher determination to quit and placed greater importance in successfully quitting on the current quit attempt. The association between high self-efficacy and motivation to quit with better cessation outcomes has been established in a large number of previous studies [6-9]. Although we did not assess quit outcomes in our study, the evidence that increased self-efficacy and motivation to quit can lead to better quitting rates is an encouraging finding for gamified smoking cessation apps. Such apps could be considered as possible cessation interventions available to smokers seeking to quit. Increased use of mobile apps for smoking cessation could also have wide-reaching consequences for alleviating health inequalities, as mHealth solutions are able to reach a large number of people at low cost [15].

It is important to note that the majority of increases in both self-efficacy and motivation to quit are evident during the first 2 weeks of app use. This could imply that the gamified mobile apps have a saturated effect after an initial period of using the app, after which they help participants maintain their self-efficacy and motivation levels. Past research has found that an increase in self-efficacy during the course of an intervention can lead to greater likelihood of long-term success [53]. Another study showed that participants who experienced an increase in self-efficacy during the first 2 weeks of a 12-week smoking cessation intervention were significantly more likely to stay quit after treatment [54]. The study also sheds light on the importance of promoting a smoker's early sense of confidence in their ability to quit to increase the odds of successfully quitting long after the intervention is completed [54]. These findings could have possible implications for future smoking cessation interventions to adopt strategies to raise self-efficacy and motivation levels over the course of the intervention, especially *early on*.

Finally, our analyses also showed that having higher education, baseline self-efficacy, and motivation to quit were associated with smaller improvements in self-efficacy and motivation to quit. This suggests that the gamified mobile apps in our study have a greater benefit for individuals with lower levels of confidence in their ability to quit, individuals with lower determination to quit, and individuals with lower education levels or individuals with lower socioeconomic status. As socioeconomic differences are present in both smoking prevalence and successful cessation, this finding could be used to inform future interventions to help disadvantaged groups and thereby reduce inequalities.

Limitations

By examining the impact of gamified smoking cessation mobile apps quantitatively, we attempt to address a gap in the existing literature. However, to do so, we developed a questionnaire to quantitatively assess gamification that requires participant self-report. Self-reporting may have led to different or inaccurate perceptions, particularly when answering questions regarding frequency of use. Moreover, the developed questionnaire has not yet been scientifically validated. Future research could test the validity and reliability of the questionnaire developed to assess gamification more rigorously. Another drawback of our research was that the majority of participants reported having a low to moderate dependence on nicotine. It could be that the findings differed among individuals with high nicotine dependence. Therefore, future research could explore the differences between the various types of smokers. Similarly, participants with mental health conditions were not eligible to participate in the study, and it could be that the findings are not generalizable to all members of the population. Finally, our study comprised motivated volunteers, which could subject the findings to volunteer bias.

To address the natural limitations of this study design, future research could consider running randomized controlled trials with 2 smoking cessation apps that are as similar as possible, differing only in the number of gamification elements or, more importantly, the type of gamification elements to robustly test the impact of gamification. Our study was underpowered to investigate the impact of individual game elements; therefore, we were unable to explore the effects of and differences between game elements. Future studies could try to isolate and test individual game elements within behavior change interventions, as not all gamification elements may have the same impact or function in the same way. It could be that certain gamification elements interact with other elements or with individual dispositions, situational circumstances, and the characteristics of particular target activities differently than others [20]. Consequently, it is vital for future research to focus on theory-driven studies that investigate how individual gamification elements work.

Conclusions

In conclusion, our research found that more frequent engagement with gamification features in smoking cessation apps was associated with higher self-efficacy and motivation to quit. The findings of this study provide a good platform for further investigation into the role of gamification in improving important cognitive factors essential for the quitting process of smokers. Future studies should continue to explore the impact and usefulness of gamification in the context of mHealth. On the basis of our findings and existing literature, we recommend that mobile app developers collaborate with behavior change specialists to develop more tailored, evidence-based, and theory-driven interventions. At the same time, app developers should be encouraged to work together with scientists to explore and test strategies, such as gamification, that could target vital psychological components of behavior change while possibly improving engagement with the app.

Acknowledgments

Ethical approval was obtained from the Joint Research Imperial College London Research Ethics Committee (ICREC) before the beginning of the study (ICREC reference 19IC5158). This research did not receive a grant or funding from any agencies that are public, commercial, or within the not-for-profit sector. Quit Genius and Kwit provided free access to the study participants but had no other financial or material input.

Authors' Contributions

The 3 authors (NBR, NM, and FTF) jointly developed and designed the study. NBR handled day-to-day activities to manage the study and therefore conducted data collection. The statistical analyses presented in this paper were performed by NBR with assistance and guidance from NM and FTF. All authors reviewed the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study participant flowchart.
[DOCX File , 46 KB - games_v9i2e27290_app1.docx]

Multimedia Appendix 2
Screenshots of Quit Genius and Kwit.
[DOCX File , 646 KB - games_v9i2e27290_app2.docx]

Multimedia Appendix 3
Gamification features and tactics by Cugelman embedded in the Kwit and Quit Genius apps.
[DOCX File , 13 KB - games_v9i2e27290_app3.docx]

Multimedia Appendix 4
The t tests statistically examining the mean differences in self-efficacy and motivation to quit scores between study time points (n=116).
[DOCX File , 13 KB - games_v9i2e27290_app4.docx]

References

1. GBD 2013 Risk Factors Collaborators, Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, et al. . [doi: 10.1016/S0140-6736(15)00128-2] [Medline: 26364544]
2. Tobacco. World Health Organisation. 2020. URL: https://www.who.int/en/news-room/fact-sheets/detail/tobacco [accessed 2021-01-05]
3. Statistics on Smoking, England - 2019. National Health Service. 2019. URL: https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-smoking/statistics-on-smoking-england-2019/part-1-smoking-related-ill-health-and-mortality [accessed 2020-11-05]
4. Garnett C, Tombor I, Beard E, Jackson SE, West R, Brown J. Changes in smoker characteristics in England between 2008 and 2017. Addiction 2020 Apr;115(4):748-756 [FREE Full text] [doi: 10.1111/add.14882] [Medline: 31914486]
5. Health matters: stopping smoking – what works? Public Health England. URL: https://www.gov.uk/government/publications/health-matters-stopping-smoking-what-works/health-matters-stopping-smoking-what-works [accessed 2019-11-08]
6. Smit ES, Hoving C, Schelleman-Offermans K, West R, de Vries H. Predictors of successful and unsuccessful quit attempts among smokers motivated to quit. Addict Behav 2014 Sep;39(9):1318-1324. [doi: 10.1016/j.addbeh.2014.04.017] [Medline: 24837754]
7. Ockene J, Emmons K, Mermelstein R, Perkins KA, Bonollo DS, Voorhees CC, et al. Relapse and maintenance issues for smoking cessation. Health Psychol 2000 Jan;19(1S):17-31 [FREE Full text] [doi: 10.1037/0278-6133.19.suppl1.17] [Medline: 10709945]
8. Gwaltney CJ, Metrik J, Kahler CW, Shiffman S. Self-efficacy and smoking cessation: a meta-analysis. Psychol Addict Behav 2009 Mar;23(1):56-66 [FREE Full text] [doi: 10.1037/a0013529] [Medline: 19290690]
9. Piñeiro B, López-Durán A, Del Río EF, Martínez Ú, Brandon TH, Becoña E. Motivation to quit as a predictor of smoking cessation and abstinence maintenance among treated Spanish smokers. Addict Behav 2016 Feb;53:40-45. [doi: 10.1016/j.addbeh.2015.09.017] [Medline: 26441045]
10. Motivation to Stop Smoking. NHS Center for Smoking Cessation and Training. URL: https://www.ncsct.co.uk/usr/pub/Motivation%20to%20stop%20smoking.pdf [accessed 2018-01-07]
11. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review 1977;84(2):191-215 [FREE Full text] [doi: 10.1037/0033-295x.84.2.191]
12. Etter J, Bergman M, Humair J, Perneger TV. Development and validation of a scale measuring self-efficacy of current and former smokers. Addiction 2000 Jun;95(6):901-913. [doi: 10.1046/j.1360-0443.2000.9569017.x] [Medline: 10946439]
13. Lindson N, Thompson T, Ferrey A, Lambert JD, Aveyard P. Motivational interviewing for smoking cessation. Cochrane Database Syst Rev 2019 Jul 31;7:CD006936 [FREE Full text] [doi: 10.1002/14651858.CD006936.pub4] [Medline: 31425622]

14. Elshatarat RA, Yacoub MI, Khraim FM, Saleh ZT, Afaneh TR. Self-efficacy in treating tobacco use: A review article. Proceedings of Singapore Healthcare 2016 Sep 16;25(4):243-248. [doi: 10.1177/2010105816667137]
15. Use of appropriate digital technologies for public health. World Health Organisation. URL: http://www.apps.who.int/gb/ebwha/pdf_files/WHA71/A71_20-en.pdf?ua=1 [accessed 2019-11-05]
16. Deterding S, Dixon D, Khaled R. From game design elements to gamefulness: defining "gamification". 2181040: ACM; 2011 Presented at: Academic MindTrek Conference: Envisioning Future Media Environments; 28th-30th September; New York p. 9-15. [doi: 10.1145/2181037.2181040]
17. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
18. Harold DJ. A Qualitative Study on the Effects of Gamification on Student Self - Efficacy. In: Gamification and Self-Efficacy. 2015 Presented at: International Society for Technology in Education Conference; 2015; Philadelphia.
19. Thorsteinsen K, Vittersø J, Svendsen G. Increasing physical activity efficiently: an experimental pilot study of a website and mobile phone intervention. Int J Telemed Appl 2014;2014:746232 [FREE Full text] [doi: 10.1155/2014/746232] [Medline: 24963290]
20. 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 [FREE Full text] [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
21. Nacke, Deterding S. The maturing of gamification research. Computers in Human Behavior 2017 Jun;71:450-454 [FREE Full text] [doi: 10.1016/j.chb.2016.11.062]
22. Landers R, Bauer K, Callan R. Gamification of task performance with leaderboards: A goal setting experiment. Computers in Human Behavior 2017 Jun;71:508-515 [FREE Full text] [doi: 10.1016/j.chb.2015.08.008]
23. Shameli A, Althoff T, Saberi A, Leskovec J. How Gamification Affects Physical Activity: Large-scale Analysis of Walking Challenges in a Mobile Application. Proc Int World Wide Web Conf 2017 Apr;2017:455-463 [FREE Full text] [doi: 10.1145/3041021.3054172] [Medline: 28990011]
24. Corepal R, Best P, O'Neill R, Tully MA, Edwards M, Jago R, et al. Exploring the use of a gamified intervention for encouraging physical activity in adolescents: a qualitative longitudinal study in Northern Ireland. BMJ Open 2018 Apr 20;8(4):e019663 [FREE Full text] [doi: 10.1136/bmjopen-2017-019663] [Medline: 29678971]
25. Kappen D, Mirza-Babaei P, Nacke L. Gamification of Older Adults' Physical Activity: An Eight-Week Study. 2018 Presented at: 51st Hawaii International Conference on Systems Sciences; 2018; Hawaii p. 1207-1216 URL: http://hdl.handle.net/10125/50036 [doi: 10.24251/hicss.2018.149]
26. Pløhn T, Aalberg T. Using Gamification to Motivate Smoking Cessation. 2015 Presented at: European Conference on Games Based Learning; 2015; Norway. [doi: 10.1111/j.1475-6765.2011.02036.x]
27. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game On? Smoking Cessation Through the Gamification of mHealth: A Longitudinal Qualitative Study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]
28. Lin Y, Tudor-Sfetea C, Siddiqui S, Sherwani Y, Ahmed M, Eisingerich AB. Effective Behavioral Changes through a Digital mHealth App: Exploring the Impact of Hedonic Well-Being, Psychological Empowerment and Inspiration. JMIR Mhealth Uhealth 2018 Jun 15;6(6):e10024 [FREE Full text] [doi: 10.2196/10024] [Medline: 29907557]
29. Hiscock R, Dobbie F, Bauld L. Smoking Cessation and Socioeconomic Status: An Update of Existing Evidence from a National Evaluation of English Stop Smoking Services. Biomed Res Int 2015;2015:274056 [FREE Full text] [doi: 10.1155/2015/274056] [Medline: 26273602]
30. Webb J, Peerbux S, Smittenaar P, Siddiqui S, Sherwani Y, Ahmed M, et al. Preliminary Outcomes of a Digital Therapeutic Intervention for Smoking Cessation in Adult Smokers: Randomized Controlled Trial. JMIR Ment Health 2020 Oct 06;7(10):e22833 [FREE Full text] [doi: 10.2196/22833] [Medline: 33021488]
31. Lokker C, Jezrawi R, Gabizon I, Varughese J, Brown M, Trottier D, et al. Feasibility of a Web-Based Platform (Trial My App) to Efficiently Conduct Randomized Controlled Trials of mHealth Apps For Patients With Cardiovascular Risk Factors: Protocol For Evaluating an mHealth App for Hypertension. JMIR Res Protoc 2021 Feb 01;10(2):e26155 [FREE Full text] [doi: 10.2196/26155] [Medline: 33522978]
32. Huang F, Wu X, Xie Y, Liu F, Li J, Li X, et al. An automated structured education intervention based on a smartphone app in Chinese patients with type 1 diabetes: a protocol for a single-blinded randomized controlled trial. Trials 2020 Nov 23;21(1):944 [FREE Full text] [doi: 10.1186/s13063-020-04835-9] [Medline: 33225982]
33. Rajani NB, Weth D, Mastellos N, Filippidis FT. Use of gamification strategies and tactics in mobile applications for smoking cessation: a review of the UK mobile app market. BMJ Open 2019 Jun 17;9(6):e027883 [FREE Full text] [doi: 10.1136/bmjopen-2018-027883] [Medline: 31213452]
34. Rajani NB, Weth D, Mastellos N, Filippidis FT. Adherence of popular smoking cessation mobile applications to evidence-based guidelines. BMC Public Health 2019 Jun 13;19(1):743 [FREE Full text] [doi: 10.1186/s12889-019-7084-7] [Medline: 31196062]
35. Kretz G. Kwit. Strasbourg; 2020. URL: https://kwit.app/en [accessed 2018-02-11]
36. Sherwani Y. Quit Genius. London; 2020. URL: https://www.quitgenius.com/ [accessed 2018-01-07]

37. International Standard Classification of Education. International Labour Organisation. URL: https://ilostat.ilo.org/resources/methods/classification-education/ [accessed 2019-11-05]
38. Health statistics and information systems: Definition of regional groupings. World Health Organisation. URL: https://www.who.int/healthinfo/global_burden_disease/definition_regions/en/ [accessed 2019-11-06]
39. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
40. van Amsterdam J, Vorspan F, Snijder M, van den Brink W, Schene AH, Stronks K, et al. Use of the Fagerström test to assess differences in the degree of nicotine dependence in smokers from five ethnic groups: The HELIUS study. Drug Alcohol Depend 2019 Jan 01;194:197-204 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.10.011] [Medline: 30447512]
41. West R, Willis N. Double-blind placebo controlled trial of dextrose tablets and nicotine patch in smoking cessation. Psychopharmacology (Berl) 1998 Mar 11;136(2):201-204. [doi: 10.1007/s002130050557] [Medline: 9551778]
42. Ussher M, Kakar G, Hajek P, West R. Dependence and motivation to stop smoking as predictors of success of a quit attempt among smokers seeking help to quit. Addict Behav 2016 Feb;53:175-180 [FREE Full text] [doi: 10.1016/j.addbeh.2015.10.020] [Medline: 26547043]
43. Davis FD, Bagozzi RP, Warshaw PR. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
44. Uechi H, Tan N, Honda Y. Effects of gamification-based intervention for promoting health behaviors. Journal of Physical Fitness and Sports Medicine 2018;7(3):185-192. [doi: 10.7600/jpfsm.7.185]
45. Othman M, Ugap N, Rahman N. Play4fit: Enhancing Users' Engagement With Smartphone Health And Fitness Application Using Gamification Concept. International Journal of Scientific and Technology Research 2019;8(12):676-681 [FREE Full text]
46. Looyestyn J, Kernot J, Boshoff K, Ryan J, Edney S, Maher C. Does gamification increase engagement with online programs? A systematic review. PLoS One 2017;12(3):e0173403 [FREE Full text] [doi: 10.1371/journal.pone.0173403] [Medline: 28362821]
47. Cobb NK, Graham AL, Bock BC, Papandonatos G, Abrams DB. Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob Res 2005 Apr;7(2):207-216 [FREE Full text] [doi: 10.1080/14622200500055319] [Medline: 16036277]
48. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. JAMA 2001 Mar 07;285(9):1172-1177. [doi: 10.1001/jama.285.9.1172] [Medline: 11231746]
49. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. In: Flow. New York: Harper Perennial; 1990:1-303.
50. Rachels J. The effect of gamification on elementary students? Spanish language achievementacademic self-efficacy. Liberty University. 2016. URL: https://core.ac.uk/download/pdf/58827222.pdf [accessed 2021-04-21]
51. Richter G, Raban D, Rafaeli S. Studying Gamfiication: The Effect of Rewards and Incentives on Motivation. In: Gamification in Education and Business. Switzerland: Springer International Publishing; 2015:21-46.
52. Rajani NB, Weth D, Mastellos N, Filippidis FT. Use of gamification strategies and tactics in mobile applications for smoking cessation: a review of the UK mobile app market. BMJ Open 2019 Jun 17;9(6):e027883 [FREE Full text] [doi: 10.1136/bmjopen-2018-027883] [Medline: 31213452]
53. Scheiding R. The Relationship Between Smoking Cessation and Self-Efficacy. Marietta College. Ohio: Marietta College; 2009. URL: https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=marietta1260369637&disposition=inline [accessed 2021-04-21]
54. Schnoll RA, Martinez E, Tatum KL, Glass M, Bernath A, Ferris D, et al. Increased self-efficacy to quit and perceived control over withdrawal symptoms predict smoking cessation following nicotine dependence treatment. Addict Behav 2011 Jan;36(1-2):144-147 [FREE Full text] [doi: 10.1016/j.addbeh.2010.08.024] [Medline: 20869812]

Abbreviations

CBT: cognitive behavioral therapy
ICREC: Imperial College London Research Ethics Committee
mHealth: mobile health
PID: participant identification number

Edited by N Zary; submitted 20.01.21; peer-reviewed by J Moon, M Schmidt-Kraepelin; comments to author 12.02.21; revised version received 24.02.21; accepted 02.04.21; published 27.04.21.

Please cite as:
Rajani NB, Mastellos N, Filippidis FT
Impact of Gamification on the Self-Efficacy and Motivation to Quit of Smokers: Observational Study of Two Gamified Smoking Cessation Mobile Apps
JMIR Serious Games 2021;9(2):e27290
URL: https://games.jmir.org/2021/2/e27290
doi:10.2196/27290
PMID:33904824

Original Paper

Integration of Motor Learning Principles Into Virtual Reality Interventions for Individuals With Cerebral Palsy: Systematic Review

Marika Demers[1], PhD; Karen Fung[2], MSc; Sandeep K Subramanian[3], PhD; Martin Lemay[4,5], PhD; Maxime T Robert[2], PhD

[1]Division of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles, CA, United States

[2]Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Department of Rehabilitation, Faculty of Medicine, Université Laval, Quebec, QC, Canada

[3]Department of Physical Therapy, School of Health Professions, UT Health San Antonio, San Antonio, TX, United States

[4]Centre de Réadaptation Marie Enfant, Centre de Recherche du CHU Sainte Justine, Montreal, QC, Canada

[5]Département des Sciences de l'Activité Physique, Université du Québec à Montréal, Montreal, QC, Canada

Corresponding Author:
Maxime T Robert, PhD
Centre for Interdisciplinary Research in Rehabilitation and Social Integration
Department of Rehabilitation, Faculty of Medicine
Université Laval
525 boul Hamel
Quebec, QC, G1M 2S8
Canada
Phone: 1 418 529 9141 ext 6121
Email: maxime.robert@fmed.ulaval.ca

Abstract

Background: Increasing evidence supports the use of virtual reality systems to improve upper limb motor functions in individuals with cerebral palsy. While virtual reality offers the possibility to include key components to promote motor learning, it remains unclear if and how motor learning principles are incorporated into the development of rehabilitation interventions using virtual reality.

Objective: The objective of this study was to determine the extent to which motor learning principles are integrated into virtual reality interventions targeting upper limb function in individuals with cerebral palsy.

Methods: A systematic review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The search was performed in 10 databases using a combination of keywords related to cerebral palsy, virtual reality, video games, and rehabilitation. Studies were divided into 2 categories: commercial video game platforms and devices and custom virtual reality systems. Study quality was assessed using the modified Downs and Black checklist.

Results: The initial search yielded 1497 publications. A total of 26 studies from 30 publications were included, with most studies classified as "fair" according to the modified Downs and Black checklist. The majority of studies provided enhanced feedback and variable practice and used functionally relevant and motivating virtual tasks. The dosage varied greatly (total training time ranged from 300 to 3360 minutes), with only 6 studies reporting the number of movement repetitions per session. The difficulty progression and the assessment of skills retention and transfer were poorly incorporated, especially for the commercial video games.

Conclusions: Motor learning principles should be better integrated into the development of future virtual reality systems for optimal upper limb motor recovery in individuals with cerebral palsy.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020151982; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020151982

(JMIR Serious Games 2021;9(2):e23822) doi:10.2196/23822

KEYWORDS
virtual rehabilitation; upper limb; brain damage; feedback; active video games; learning

Introduction

Cerebral palsy (CP) is the most common neuromotor disorder in children, with a prevalence ranging from 1.5 to 2.5 per 1000 births [1-3], that continues throughout adulthood. It is defined as "a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain" [4]. Due to structural brain abnormalities, individuals with CP have a wide range of sensorimotor impairments, including muscle tone disorders [5], sensory deficits [6-8], and deficits in interjoint coordination [9], motor execution, and planning [10,11]. These impairments ultimately lead to altered upper limb function. Due to the important contribution of the arms and hands in daily activities, deficits in upper limb functions ultimately may result in a poorer quality of life [12].

To improve both motor function and quality of life in individuals with CP, moderate to strong evidence supports motor learning-based approaches in rehabilitation [13,14]. Motor learning is defined as a set of processes based on principles of neuroplasticity associated with practice or experience that lead to relatively permanent motor changes [15,16]. The brain's inherent ability to organize itself and to form new connections between neurons can be exploited with therapeutic rehabilitation approaches based on motor learning principles [17,18]. Key components of motor learning–based approaches include but are not limited to (1) intensive rehabilitation interventions involving a high number of task repetitions, (2) progressive incremental increases in task difficulty, and (3) salient interventions to enhance motivation and engagement in therapy [15,19]. The provision of extrinsic feedback on either the movement quality or the motor performance can also promote motor learning [20]. Extrinsic feedback can compensate for reduced availability and/or processing ability of intrinsic feedback in individuals with CP. However, provision of extrinsic feedback is often not individualized to the motor abilities of each individual nor standardized [20]. Additional challenges to incorporating motor learning principles in clinical practice include (1) accountability for the heterogeneity and severity of the sensorimotor impairments observed in CP, (2) personalization of interventions based on the individual's needs or goals, and (3) delivery of exercises that are both challenging and enjoyable [21].

Use of technology has helped to address some of the aforementioned challenges with rehabilitation interventions. Technology-based interventions including virtual reality and active video games have gained popularity in rehabilitation, with many systems designed to encourage upper limb function [22-24]. Virtual rehabilitation refers to interventions that are built on virtual reality platforms to meet rehabilitation goals. It encompasses a continuum of technologies of different types and technical complexities ranging from fully immersive 3D virtual reality viewed using commercially available head-mounted displays (eg, Oculus Rift; Facebook Technologies, LLC) to active video games or exergames (eg, commercial video games used for rehabilitation purposes or active video games primarily used for physical activity) [25]. Rehabilitation using virtual reality–based platforms offers the possibility to deliver high-intensity training in a multimodal training environment [26]. Virtual reality interventions also provide a unique opportunity to customize and standardize the levels of task difficulty by modifying the spatial and temporal constraints and the cognitive challenge. Feedback provision on the individual abilities and delivery modes can be easily manipulated. The task outcome and quality can be automatically recorded, which is useful to both clinicians and researchers [27]. Virtual reality has been shown to be safe and ecologically valid for the rehabilitation of individuals with CP [28-30]. The novelty of virtual reality technology and the interactive and engaging gaming characteristics are key components that provide a joyful training environment to sustain and enhance motivation to treatment [31,32]. Therefore, the attributes of virtual environments such as motivation, repetitive practice, and enhanced feedback make them an ideal modality to facilitate the incorporation of motor learning principles into the treatment of individuals with CP.

Several systematic reviews and one meta-analysis investigated the impact of virtual reality interventions on upper limb motor recovery in children and adolescents with CP [13,22,27,33-35]. Their results suggest, to an extent, that virtual reality can be effective and motivating for children with CP. Another literature review specifically looked at the effectiveness of virtual reality on motor learning in children with CP [36]. The results support virtual reality interventions to improve motor learning and encourage skill transfer to real-life situations. A current knowledge gap in the literature relates to which exact motor learning principles are incorporated into virtual reality–based platforms used for upper limb rehabilitation in individuals with CP (children, adolescents, and adults). The limited incorporation of key motor learning principles in therapy (eg, treatment intensity and specificity, feedback provision and delivery, and difficulty progression) could explain differences observed between studies and also limit the potential for motor learning in individuals with CP. This systematic review aims to identify the extent to which motor learning principles are integrated into virtual reality interventions targeting upper limb function in individuals with CP. The incorporation of motor learning principles will be identified for commercial video game platforms and devices and custom virtual reality systems for rehabilitation to help guide clinical decision making.

Methods

Protocol and Registration

This systematic review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, and results are reported using the PRISMA checklist [37]. The protocol for this systematic review was registered on PROSPERO (CRD42020151982).

Eligibility Criteria

We included publications related to the effectiveness of virtual reality interventions on upper limb sensorimotor function at the levels of body function/structure or activity limitations according to the International Classification of Functioning, Disability and Health in individuals with CP. We only included peer-reviewed publications in English or French in which the majority of the sample included individuals with CP (children or adults). We excluded publications that (1) used solely qualitative methodologies (eg, focus groups and interviews), (2) focused on the development of the virtual reality intervention, (3) used virtual reality as an assessment tool, or (4) were reviews, meta-analyses, or commentaries.

Information Sources

The following electronic databases were searched on August 22, 2019, and updated on July 5, 2020, using a combination of keywords related to CP, rehabilitation, video games, and virtual reality: MEDLINE, Embase, CINAHL, Web of Science, Google Scholar, OTseeker, Physiotherapy Evidence Database, IEEE Xplore, Scopus, and Cochrane Library, and the Cochrane Central Register of Controlled Trials. The reference lists of articles of interest were also searched for additional references. The search strategy was developed for the MEDLINE database and adapted for other databases. Various combinations of keywords and Medical Subject Headings (MeSH) or Embase subject headings (Emtree) related to virtual reality or video games, rehabilitation, and CP were used (see Multimedia Appendix 1 for the detailed search strategy). No time limit regarding the date of publication was applied to the search strategy.

Search and Study Selection

The search strategy was executed by one researcher (MD), and all publications from each database were extracted using citation management software (Endnote X9; Clarivate). Any duplications were removed. Titles and abstracts were screened independently by two researchers (MD and KF) based on the inclusion and exclusion criteria. For all potential eligible studies, full texts were retrieved, and eligibility was assessed by the same two researchers. Any conflict was resolved by discussion.

Data Collection Process and Data Items

For all publications meeting the inclusion criteria, one researcher (MD) extracted the following information into an Excel (Microsoft Inc) spreadsheet: author/date, study design, participants, sample size, virtual reality system, delivery method, number of movement repetitions, intensity, task specificity, difficulty progression, type of practice, type of feedback, feedback delivery schedule, motivation, motor recovery outcome measures, changes in upper limb motor function (body function/structure or activity limitation levels), assessment of skills retention, and assessment of transfer of skills. Another researcher (KF) validated the data extraction by reading all included publications and confirming that the data extracted were accurate and complete. Publications presenting the results from the same group of participants were considered a single study and the results were extracted together. Since there is a lack of clear definition of motor learning principles in CP, we used the definition and key descriptors of Maier et al [38] for dosage, type of practice, feedback provision, and task specificity. For treatment intensity and duration, the number of movement repetitions was extracted and the number of minutes of treatment was computed using the treatment duration, frequency, and number of weeks of the intervention. Data were analyzed separately for commercial video game platforms and devices and custom virtual reality systems for rehabilitation. Commercial video game platforms and devices included salon game consoles and commercially available applications on a tablet. Custom virtual reality systems for rehabilitation included commercially available or custom virtual reality software programmed for rehabilitation purposes. Studies using special hardware (eg, instrumented gloves or robotic devices) or repurposed commercial gaming hardware (eg, the Microsoft Kinect camera) were included in the custom virtual reality software category.

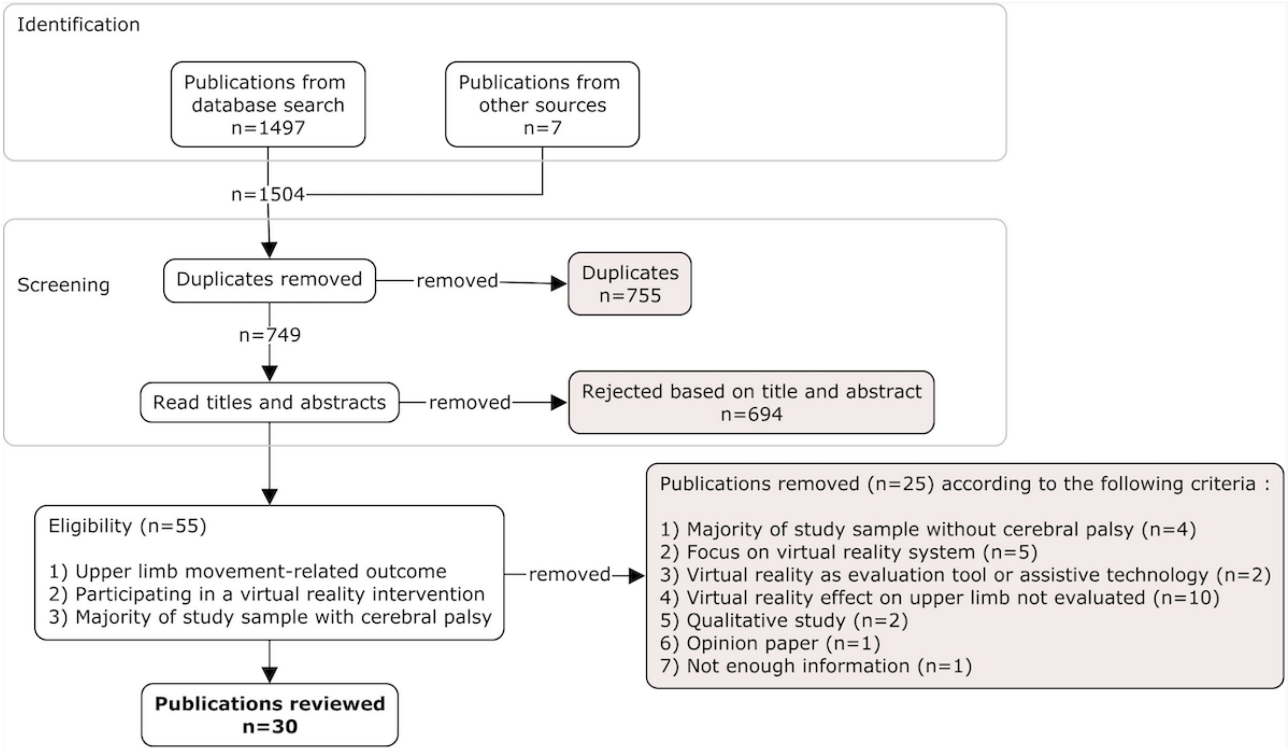
Risk of Bias in Individual Studies

Two authors (SKS and MTR) analyzed the quality of each study, and conflicts were resolved by discussion. The 27-item modified version [39] of the original Downs and Black checklist [40] helped assess the quality of the included randomized and nonrandomized studies. The overall quality of the research was scored out of 28 based on the following criteria: reporting, internal validity, power, and external validity. Scores on the modified Downs and Black checklist were classified as "excellent" (scores of 24-28), "good" (scores of 19-23), "fair" (scores of 14-18), or "poor" (scores ≤13) [41]. The Downs and Black checklist was chosen over other measures, such as the Effective Public Health Practice Project [42], because it considers sample size in the total score calculation and enables studies to be quantitatively classified into different categories based on the total score. The Downs and Black checklist has previously been used in studies involving virtual reality interventions for upper limb rehabilitation [43].

Results

The database search yielded 1497 publications, and 749 publications were screened for eligibility after duplicates were removed. After full-text review, 26 studies from 30 publications were included (see Figure 1 for flow diagram and reasons for exclusion).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart of study selection.



Quality Assessment

The quality ratings of the publications based on the Downs and Black checklist are shown in Table 1. Only 17 studies were rated because the remaining studies were either case studies or short papers. Of the 17 studies, 2 were rated as good, 14 as fair, and 1 as poor.

Table 1. Quality assessment of the reviewed studies based on the 27-item modified version [39] of the Downs and Black checklist [40].

First author and publication year	Downs and Black score/28	Quality of study
Avcil et al, 2020 [44]	21	Good
Bedair et al, 2016 [45]	16	Fair
Chen et al, 2007 [46]	15	Fair
Sharan et al, 2012 [47]	13	Fair
El-Shamy, 2018 [48]	16	Fair
Fluet et al, 2009 [49]	16	Fair
Fluet et al, 2010 [50]	15	Fair
Hernández et al, 2018 [51]	17	Fair
Jannink et al, 2008 [52]	14	Fair
Kassee et al, 2017 [53]	17	Fair
Odle et al, 2009 [54]	14	Fair
Rostami et al, 2012 [55]	18	Fair
Sahin et al, 2020 [56]	19	Good
Sandlund et al, 2014 [57]	13	Fair
Turconi et al, 2016 [58]	18	Fair
Weightman et al, 2011 [59]	12	Poor
Winkels et al, 2013 [60]	13	Fair

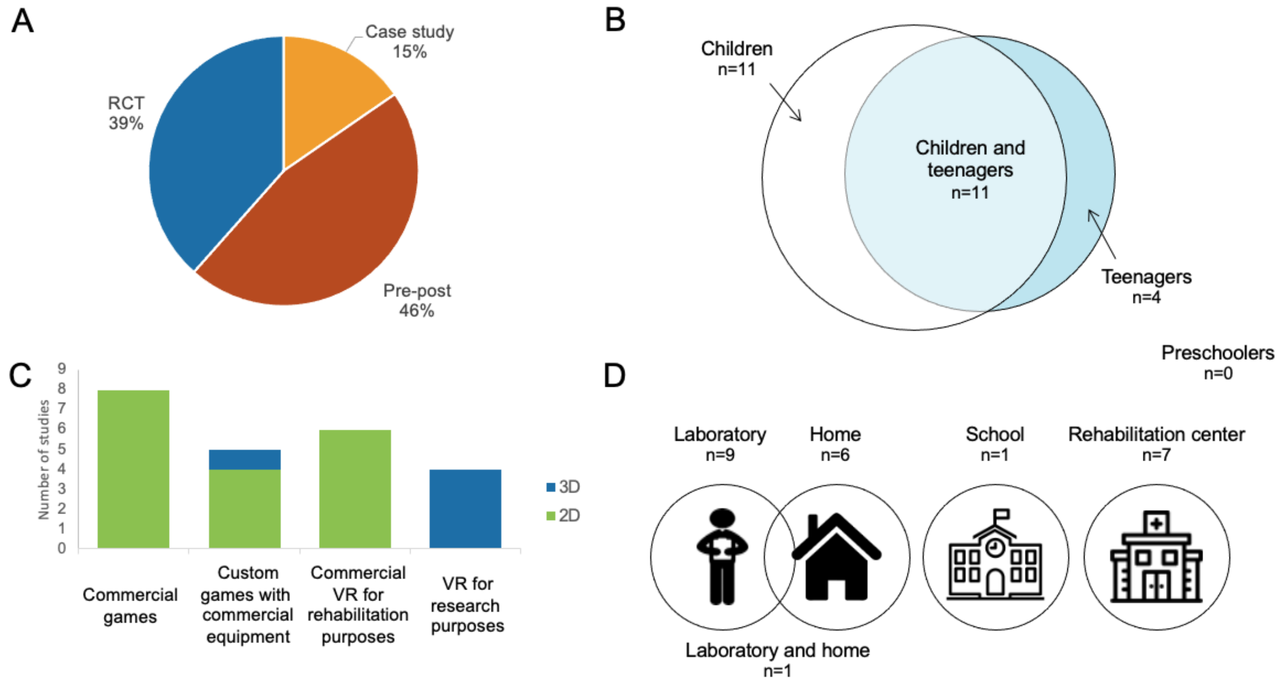
Characteristics of the Virtual Reality Systems

Of the 26 studies, 9 (35%) used commercial video game platforms and devices, such as the Nintendo Wii, the Sony PlayStation 2 or 3, Xbox consoles (Microsoft Corporation), or commercial applications [44,45,47,52,53,58,60-62]. Custom virtual reality systems for rehabilitation were used in 17 studies (65% [46,48,50,51,54-56,59,63-71]). Commercially available virtual reality systems designed for rehabilitation purposes were used in 5 studies (19%): ArmeoSpring (Hocoma) combined with virtual games [48]; E-Link Evaluation and Exercise System (Biometrics Ltd) [55]; IREX (GestureTek Health) [66]; Timocco [64]; and YouGrabber (YouRehab) [67]. Seven studies (27%) used custom-based games combined with commercially available accessories (ie, Microsoft Kinect camera [56,68,72], webcam [54], force-feedback motion controller [51], or instrumented gloves [69,70]). Custom virtual reality systems designed for rehabilitation research purposes were used in 4 studies (15%) [49,50,59,65,71]. Display media used to view the virtual environments included stereoscopic glasses to enable 3D media display view [49,50,71] and television or computer monitors to view 2D virtual environments with various 3D rendering features (ie, shadow, drop lines, etc). Virtual environments ranged from a simple display of reaching targets in a 2D plane to a detailed replication of real-life environments, such as a tennis court or a kitchen. None of the studies used immersive virtual reality through a head-mounted display.

Study Characteristics

The study settings and targeted participants varied greatly. Sample sizes ranged from 1 to 30 participants in the intervention group, with 12 studies (46%) having a sample size of 5 participants or fewer. While our search strategy did not exclude studies conducted in adults with CP, none of the studies targeted participants over 18 years of age (see Figure 2 for detailed study characteristics). All studies targeted school-aged children (over 4 years of age) or adolescents. In 17 studies (65%), treatment was delivered in a laboratory setting [46,48,50,55,58,61,65-67,71] or a rehabilitation center [45,47,51,52,54,60,68]. In 6 studies (23%), the virtual reality–based intervention was delivered at home using either telerehabilitation technologies (ie, videoconferencing and remote monitoring [64,69,70]) or a prescribed exercise program [53,57,59].

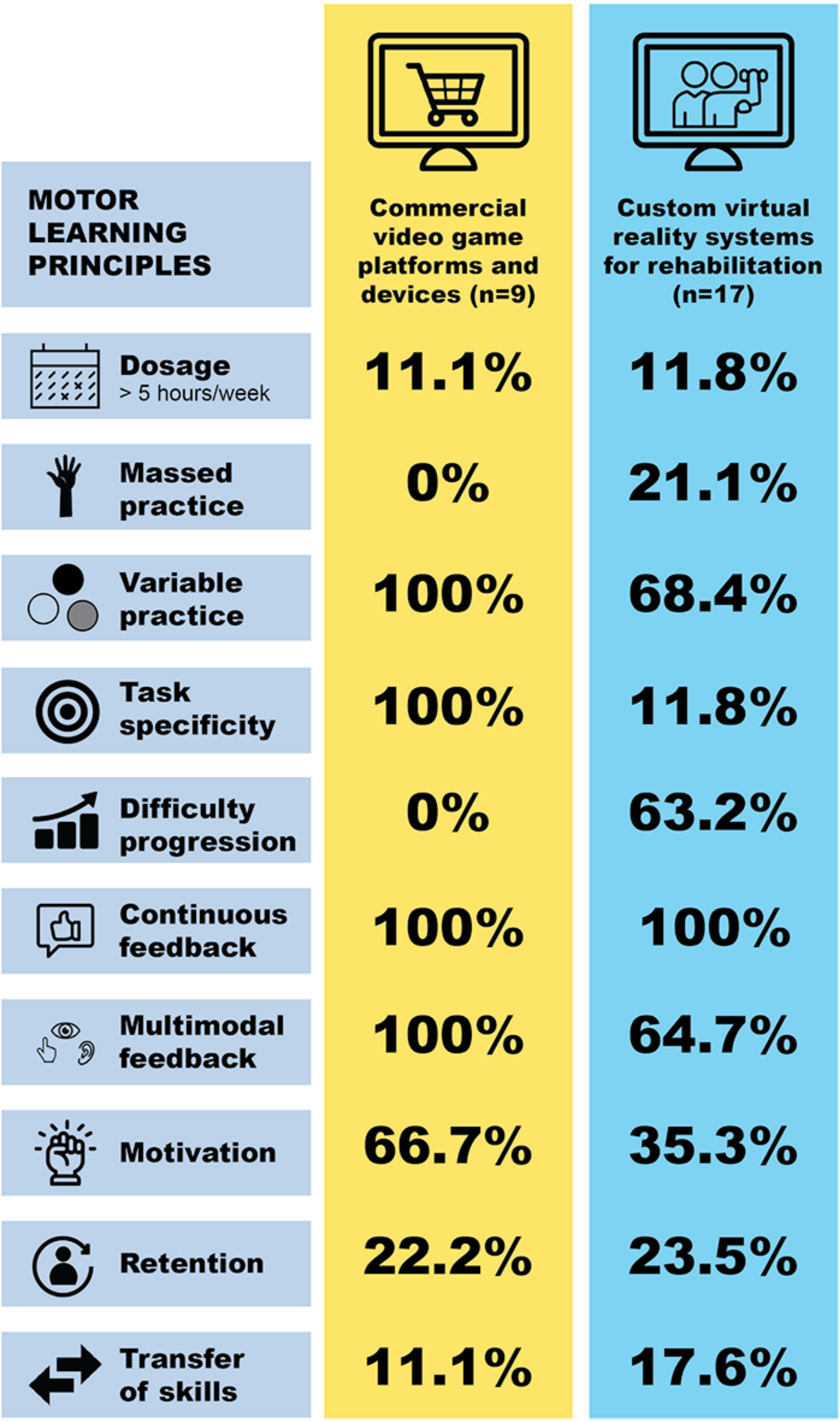
Figure 2. Characteristics of the reviewed studies according to (A) study design, (B) participants' age groups (preschoolers: aged 0-4 years; children: aged 4-12 years; teenagers: aged 13-18 years), (C) virtual reality (VR) system type, and (D) delivery environment. RCT: randomized controlled trial.



Motor Learning Principles in Studies Using Commercial Video Game Platforms and Devices

Among the 9 studies (35%) using commercial video game platforms and devices, the length of sessions varied from 20 to 60 minutes (Figure 3 and Multimedia Appendix 1). Frequency ranged from 2 to 5 sessions per week for 3 to 12 weeks, with a total treatment time ranging from 360 to 1440 minutes. Only Kassee et al [53] discussed the number of movement repetitions, suggesting that the number of movement repetitions in the virtual reality group was comparable with that in the resistance training group (144 repetitions per session). Jannink et al [52] indicated that the intensity level was moderate but did not define how intensity was measured in their study. Task-specific training, in which movements were goal-oriented or relevant for activities of daily living, was used in all 9 studies and the majority simulated competitive sports. Examples of task-specific training included playing various sports (eg, tennis, bowling, sword fighting), controlling a moving car, piloting a spaceship, and catching falling balls. Variable practice was delivered in all studies using commercial video game platforms and devices by using frequent changes of tasks. However, none of the studies clearly specified the type of practice offered with the intervention. None of the publications reported whether the games were adapted to match the participant's capacity and, if so, how they were adapted and how the difficulty levels were changed.

Figure 3. Percentage of commercial video game platforms and devices and custom virtual reality systems for rehabilitation integrating the principles of motor learning.



Visual and auditory feedback were provided in all studies as a display of total score and/or reward sounds. Additional haptic feedback (vibration) was offered by a motion controller held in the most-affected hand in 2 studies [44,53]. All studies delivered continuous feedback. For virtual interventions, an example of continuous feedback is the provision of knowledge of results after each trial, such as the rate of success or total score. Motivation with the intervention was assessed in 6 studies (23%) using questionnaires (eg, a visual analog scale for motivation and a participant and parent feedback questionnaire), therapists' observations, or semistructured interviews. High motivation to practice in a virtual environment was reported [47,52,53,58,60,62]. However, Sandlund et al [57] noted that children's interest in gaming faded somewhat over time. Winkels et al [60] observed that motivation varied between games and children. In 2 studies, greater motivation was observed in the Wii training group than in a control group [47,53]. Retention of motor skills was assessed in only 2 studies [53,61]; the results of both studies suggested that motor skills were retained 4 weeks after the intervention. Transfer of skills was only assessed in 1 study, with the skills shown to be transferred to a similar reaching task [57].

Motor Learning Principles in Studies Using Custom Virtual Reality Systems for Rehabilitation

In all 17 studies (65%) using custom virtual reality systems for rehabilitation, the treatment frequency and intervention dosage were clearly reported, with the total training time ranging from 300 to 3360 minutes. The number of movement repetitions was reported in only 5 studies (range of 45 to 550 repetitions per session) [46,50,51,59,72]. Nine studies (35%) delivered task-specific training with functionally relevant tasks [46,48-50,56,64,66-68,71]. Thirteen studies (50%) delivered variable practice, with 4 studies (15%) also delivering massed practice (eg, minimal time between sessions and large number of movement repetitions to a single target) [46,49-51,59] and 1 study (4%) using random practice (ie, randomized tasks to maintain patients' interest) [70]. Two studies (8%) delivered constant practice [64,65], while the remaining studies did not specify the type of practice used. Task difficulty was increasingly progressed in 12 studies [46,48,50,51,55,59,65,67-71] by different methods, including a modification of the ranges of motion required to accomplish the task, an increase or decrease in the amount of assistive and/or resistive force, and a change in the speed, accuracy, or target characteristics. Difficulty levels were progressed by the system using an algorithm, based on the judgment of a therapist according to task success, according to difficulties reported by the participants, or based on preset difficulty levels (eg, easy, medium, hard).

Multisensory feedback that combined visual and auditory and/or haptic feedback was offered in 11 studies (42%) [46,48-51,55,59,63,64,70,71,73]. Six studies (23%) provided only visual feedback [54,56,64-66,68]. Feedback was delivered continuously in all studies, but 1 study did not report on feedback frequency [51]. Motivation was assessed in 6 studies (23%) using subjective assessment or semistructured interviews [46,50,51,64,69,71]. The results suggested that the virtual environments were motivating, but Chen et al [46] and Hernández et al [51] reported that motivation levels were highly variable from one child to another, ranging from low to high. Retention of motor skills was assessed in only 4 studies [46,55,65,66]; in 3 of the 4 studies, the motor skills gained by the virtual reality intervention were maintained or improved at 1 to 3 months [46,55,66]. Three studies assessed the transfer of skills [51,59,66]. Hernández et al [51] reported that all children made significant progress on their self-selected goals, which largely targeted activities of daily living and leisure. In another study [66], 66% of the children showed a transfer of skills to a similar reach-to-grasp task for all kinematic variables. Weightman et al [59] noticed improvements in activities of daily living not directly targeted by the intervention.

Discussion

Principal Findings

The objective of this systematic review was to examine the extent to which motor learning principles are integrated into virtual reality interventions in individuals with CP. A total of 26 studies met the inclusion criteria, of which 9 used commercial video game platforms and devices and 17 developed custom virtual reality systems for rehabilitation. Overall, the studies were considered fair based on the Downs and Black checklist, given that the majority of them were small pilot or proof-of-concept studies. Nonetheless, the novelty of this review is that virtual reality is well-suited to incorporate motor learning principles into rehabilitation interventions targeted at children and adolescents with CP. Proper integration of motor learning principles is important, as demonstrated by the fact that the most effective therapies (ie, constraint-induced movement therapy and bimanual therapy) for improving upper limb motor functions are themselves based on motor learning principles and principles of neuroplasticity [13,74,75]. Enhanced feedback provision, variable practice, task specificity, and motivation were the motor learning principles most frequently adopted in virtual reality interventions for children and adolescents with CP. Dosage varied greatly from one study to another with only a few studies reporting the number of movement repetitions per session. The application of some principles (ie, difficulty progression and assessment of skills retention and transfer) were poorly integrated, especially in commercial video game platforms and devices.

It is widely accepted that rehabilitation interventions should be delivered at a high intensity (dose, frequency, and duration of training) to engage neuroplastic mechanisms [15]. Unfortunately, our results showed that many of the reviewed studies did not provide sufficient information on the number of repetitions performed per session. Although intensive and repetitive practice is important, repetitive motor activity alone is insufficient to induce experience-dependent plasticity [76]. Virtual reality enables the possibility to deliver high-intensity practice of engaging and meaningful tasks along with relevant feedback. This promotes a problem-solving approach known to be useful for enhanced rehabilitation outcomes. Additionally, a close collaboration between game developers, academia, and clinicians in the development of both commercial and custom virtual reality systems would help identify the clinical needs and optimize virtual reality interventions for individuals with CP [77].

Extrinsic Feedback

In both the commercial and custom virtual reality systems, feedback provision was well integrated. Feedback plays a crucial role to enhance motor learning and motivation level [78,79]. All studies reported providing feedback at a continuous frequency, yet provision of continuous feedback is often not optimal for motor learning. Continuous feedback limits the opportunity for learning to occur through exploration and increases the dependence of the user on the feedback to improve specific goals, thereby affecting the ability of learning to detect and correct errors [80]. While continuous feedback may increase skills acquisition, retention of improvements is rarely sustained over time beyond the cue that prompted them (ie, feedback). Alternatively, provision of faded feedback (ie, gradual decrease in feedback provision as the learner improves) or self-controlled feedback may encourage the participant to explore and internalize new movement patterns, thus increasing the retention of these newly acquired skills [81,82]. Based on the limited evidence on the feedback modality that should be prioritized, a combination of multimodal feedback (ie, visual, auditory, and haptic) is proposed to be more effective for improving motor performance [80,83]. All studies using commercial video game platforms and devices implemented multimodal feedback. In contrast, approximately one-third of the studies using custom virtual reality systems delivered solely visual feedback. This could be because of the technical complexity required to program multimodal feedback and the lack of knowledge about which feedback modality should be provided in rehabilitation settings. Feedback should be implemented in a structured manner considering the individual capacities and errors made and thus allow the progression of difficulty throughout the intervention period. The use of theoretical frameworks, such as the framework developed by Schüler et al [84], can help designers and researchers to identify the beneficial components of virtual reality systems for specific treatment goals.

Progression of Difficulty

Constant progression of difficulty promotes motor learning because an individual's abilities are considered within the conditions of a specific learning experience. According to the Challenge Point Framework [85], learning occurs through active problem solving. Errors committed during motor learning are necessary to both improve movement behavior and provoke neuroplasticity [86]. The majority of custom virtual reality systems for rehabilitation successfully implemented this concept. Unfortunately, all 9 studies using commercial video game platforms and devices did not report on the progression of difficulty, which does not enable assessment of whether the participants in these studies were appropriately challenged. This is somewhat surprising, since game developers generally integrate difficulty progression to maintain players' enjoyment. Commercial video game platforms and devices can have limited therapeutic value for individuals with physical disabilities because they are designed for nondisabled populations [87]. Depending on the severity of the sensorimotor impairments, individuals with CP may not achieve the minimum threshold to progress through the difficulty levels in a given game. The concept of progression of difficulty is nonnegligible, as it may have a negative impact if a task is deemed too difficult. Thus, the strong association between challenge and motivation highlights the importance of delivering interventions at an appropriate difficulty level.

Motivation

Motivation is a critical element of rehabilitation, especially in the pediatric population [88,89]. A lack of motivation both increases activity limitations and decreases the child's participation, thus hindering adherence to treatment [88]. Higher levels of motivation help lead an individual toward satisfying their specific needs and achieving goals in a persistent manner [90]. Given that most commercial games have high production value and include gamification elements to promote motivation and volition [91], most of the studies using commercial video game platforms and devices reported a high level of motivation. However, it is also possible that while commercial video games may be perceived as motivating and fun, they might still be too difficult for participants with CP. Driving elements of motivation include appropriately challenging tasks, game variability, setting realistic goals, and aspects of competition such as a reward system [92]. Unfortunately, these elements were often not incorporated into custom virtual reality systems for rehabilitation. Reasons contributing to these findings could include a lack of financial resources available in the development of a game and/or the lack of collaboration between game developers, clinicians, and end users. Future studies must comprehensively assess and report motivation levels and whether or not they were sustained in the long term. Moreover, strategies used to drive motivation should also be reported to completely understand the utility of such strategies.

Retention and Transfer of Skills

Two of the most important principles of motor learning pertain to how much the improvements are retained over an extended period and how much of the newly acquired skill can be transferred to performance of a similar task [93,94]. The majority of the studies reviewed, regardless of the type of virtual reality used, did not examine these principles in depth. A few studies, however, referenced the potential to retain the newly acquired upper limb skills [46,55,61,66], as well as the possibility to transfer motor skills to real-life activities [51,57,66,73]. Virtual reality is well suited to address important factors that potentially encourage retention and transfer of skills, such as high enjoyment level, physical fidelity of the practiced movement, and high repetitions. Thus, retention is an important factor that must be addressed in future studies.

Type of Practice

Relatively few details were provided regarding the practice conditions used for studies with the commercial video game platforms and devices. For the custom virtual reality systems, 4 studies reported using massed practice. The beneficial impact of massed or distributed practice on learning is not clear and is likely related to contextual factors such as the nature (discreet vs continuous) and the difficulty of the task or the expertise of the participants. A pilot study conducted in healthy participants that compared massed practice with distributed practice in a virtual reality setting did not report any significant difference between these 2 types of practice [95]. Thirteen studies reported using variable practice, while only 1 study reported using random practice. Both variable and random practice tend to negatively affect short-term performance but often have a positive long-term impact on skill retention and transfer [96]. Variable and random practice are notably more cognitively engaging [97], more challenging, and improve generalization and adaptability, as observed when performing a novel variation of a task [85]. Our results show that many interventions do not take full advantage of increasing variability within trials to optimize skill retention and transfer. The concept of variability was first introduced by Bernstein [98], who emphasized that the success of practice relies on the process of solving a problem repetitively. To this day, this concept remains relevant and should be considered when developing virtual reality interventions. Nonetheless, it should be noted that in some contexts, blocked and constant practice might be more suitable, notably in younger children or in difficult tasks [99-101]. Therefore, virtual reality software should be flexible enough to allow the clinician to adjust the variability within a block of trials in order to maximize motor learning.

Limitations

Overall, our interpretation of the results of these studies was limited by the available information provided in the publications. For example, inconsistent reporting of the type of feedback and delivery schedule hindered our ability to conclude whether feedback was provided in the form of knowledge of performance or knowledge of results [102]. In other words, our review was constrained by the level of detail in the studies' methodology sections, which often mirrored the quality of the studies. Thus, no conclusion was drawn on the effectiveness of virtual reality in rehabilitation, as it was beyond the scope of this review.

Conclusions

This review demonstrates the current integration of select principles of motor learning into commercial video game platforms and devices and custom virtual reality systems designed for upper limb motor recovery. Overall, motor learning principles are not yet being fully integrated into virtual reality systems, especially into commercial video game platforms and devices, because the target audience is not individuals with disabilities. Custom virtual reality systems are better tailored to the needs of individuals with CP and provide an experience better adapted to the capacity of individuals in term of difficulty. However, the custom virtual reality systems used in this review were not as engaging as commercial video game platforms and devices nor did they provide multimodal feedback. Nonetheless, designing an intervention using multimodal feedback may be feasible with the proper resources. The integration of motor learning principles into such a system would help maximize its efficiency and offer a cost-effective intervention to supplement standard treatments in the clinical setting. Future research should provide detailed methodology on the extent to which motor learning principles are integrated to help evaluate the efficacy of video game platforms and devices and virtual reality systems in improving upper limb function.

Acknowledgments

MD is supported by a postdoctoral fellowship by the Fonds de la Recherche du Québec en Santé. The funder did not play a role in the publication of these results.

Authors' Contributions

All authors contributed to the writing of the manuscript and its critical revision. MD performed the search strategy, study selection, and data extraction and analysis, and drafted the manuscript. KF performed the study selection and verified the data extraction. SKS and MTF assessed the quality of the reviewed studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed search strategy.

[PDF File (Adobe PDF File), 184 KB - games_v9i2e23822_app1.pdf]

References

1. Bishop DVM. Which neurodevelopmental disorders get researched and why? PLoS One 2010 Nov 30;5(11):e15112 [FREE Full text] [doi: 10.1371/journal.pone.0015112] [Medline: 21152085]
2. Paneth N, Hong T, Korzeniewski S. The descriptive epidemiology of cerebral palsy. Clin Perinatol 2006 Jun;33(2):251-267. [doi: 10.1016/j.clp.2006.03.011] [Medline: 16765723]
3. Oskoui M, Coutinho F, Dykeman J, Jetté N, Pringsheim T. An update on the prevalence of cerebral palsy: a systematic review and meta-analysis. Dev Med Child Neurol 2013 Jun;55(6):509-519 [FREE Full text] [doi: 10.1111/dmcn.12080] [Medline: 23346889]
4. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl 2007 Feb;109:8-14. [Medline: 17370477]
5. Shortland A. Muscle deficits in cerebral palsy and early loss of mobility: can we learn something from our elders? Dev Med Child Neurol 2009 Oct;51 Suppl 4:59-63 [FREE Full text] [doi: 10.1111/j.1469-8749.2009.03434.x] [Medline: 19740211]

6. Robert MT, Guberek R, Sveistrup H, Levin MF. Motor learning in children with hemiplegic cerebral palsy and the role of sensation in short-term motor training of goal-directed reaching. Dev Med Child Neurol 2013 Dec;55(12):1121-1128 [FREE Full text] [doi: 10.1111/dmcn.12219] [Medline: 23899048]
7. Auld ML, Boyd RN, Moseley GL, Ware RS, Johnston LM. Impact of Tactile Dysfunction on Upper-Limb Motor Performance in Children With Unilateral Cerebral Palsy. Arch Phys Med Rehabil 2012 Apr;93(4):696-702. [doi: 10.1016/j.apmr.2011.10.025]
8. Gupta D, Barachant A, Gordon AM, Ferre C, Kuo H, Carmel JB, et al. Effect of sensory and motor connectivity on hand function in pediatric hemiplegia. Ann Neurol 2017 Nov;82(5):766-780 [FREE Full text] [doi: 10.1002/ana.25080] [Medline: 29034483]
9. Eliasson A, Gordon AM. Impaired force coordination during object release in children with hemiplegic cerebral palsy. Dev Med Child Neurol 2000 Apr;42(4):228-234 [FREE Full text] [doi: 10.1017/s0012162200000396] [Medline: 10795560]
10. Kurz MJ, Becker KM, Heinrichs-Graham E, Wilson TW. Neurophysiological abnormalities in the sensorimotor cortices during the motor planning and movement execution stages of children with cerebral palsy. Dev Med Child Neurol 2014 Nov;56(11):1072-1077 [FREE Full text] [doi: 10.1111/dmcn.12513] [Medline: 24931008]
11. Steenbergen B, Gordon AM. Activity limitation in hemiplegic cerebral palsy: evidence for disorders in motor planning. Dev Med Child Neurol 2006 Aug 14;48(09):780. [doi: 10.1017/s0012162206001666]
12. Makris T, Dorstyn D, Crettenden A. Quality of life in children and adolescents with cerebral palsy: a systematic review with meta-analysis. Disabil Rehabil 2021 Feb;43(3):299-308. [doi: 10.1080/09638288.2019.1623852] [Medline: 31180733]
13. Novak I, Morgan C, Fahey M, Finch-Edmondson M, Galea C, Hines A, et al. State of the Evidence Traffic Lights 2019: Systematic Review of Interventions for Preventing and Treating Children with Cerebral Palsy. Curr Neurol Neurosci Rep 2020 Feb 21;20(2):3 [FREE Full text] [doi: 10.1007/s11910-020-1022-z] [Medline: 32086598]
14. Sakzewski L, Ziviani J, Boyd RN. Efficacy of upper limb therapies for unilateral cerebral palsy: a meta-analysis. Pediatrics 2014 Jan;133(1):e175-e204. [doi: 10.1542/peds.2013-0675] [Medline: 24366991]
15. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008 Feb;51(1):S225-S239. [doi: 10.1044/1092-4388(2008/018)] [Medline: 18230848]
16. Schmidt R, Lee T, Winstein C, Wulf G, Zelaznik. Motor Control and Learning: A Behavioural Emphasis Sixth Edition. Champaign, IL: Human Kinetics; Feb 19, 2018:1-552.
17. Winstein C, Kay D. Translating the science into practice: shaping rehabilitation practice to enhance recovery after brain damage. Prog Brain Res 2015;218:331-360. [doi: 10.1016/bs.pbr.2015.01.004] [Medline: 25890145]
18. Krakauer J. The applicability of motor learning to neurorehabilitation. In: Dietz V, Ward NS, editors. Oxford Textbook of Neurorehabilitation. 2nd edition. Oxford, United Kingdom: Oxford University Press; 2015:55-63.
19. Muratori LM, Lamberg EM, Quinn L, Duff SV. Applying principles of motor learning and control to upper extremity rehabilitation. J Hand Ther 2013;26(2):94-102; quiz 103 [FREE Full text] [doi: 10.1016/j.jht.2012.12.007] [Medline: 23598082]
20. Robert MT, Sambasivan K, Levin MF. Extrinsic feedback and upper limb motor skill learning in typically-developing children and children with cerebral palsy: Review. Restor Neurol Neurosci 2017;35(2):171-184. [doi: 10.3233/RNN-160688] [Medline: 28282845]
21. Onla-or S, Winstein CJ. Determining the optimal challenge point for motor skill learning in adults with moderately severe Parkinson's disease. Neurorehabil Neural Repair 2008;22(4):385-395. [doi: 10.1177/1545968307313508] [Medline: 18326891]
22. Johansen T, Strøm V, Simic J, Rike P. Effectiveness of training with motion-controlled commercial video games on hand and arm function in young people with cerebral palsy: A systematic review and meta-analysis. J Rehabil Med 2019 Dec 03;52(1):jrm00012 [FREE Full text] [doi: 10.2340/16501977-2633] [Medline: 31794044]
23. Leal AF, da Silva TD, Lopes PB, Bahadori S, de Araújo LV, da Costa MVB, et al. The use of a task through virtual reality in cerebral palsy using two different interaction devices (concrete and abstract) - a cross-sectional randomized study. J Neuroeng Rehabil 2020 Apr 29;17(1):59 [FREE Full text] [doi: 10.1186/s12984-020-00689-z] [Medline: 32349752]
24. Metin Ökmen B, Doğan Aslan M, Nakipoğlu Yüzer GF, Özgirgin N. Effect of virtual reality therapy on functional development in children with cerebral palsy: A single-blind, prospective, randomized-controlled study. Turk J Phys Med Rehabil 2019 Jun;65(4):371-378 [FREE Full text] [doi: 10.5606/tftrd.2019.2388] [Medline: 31893274]
25. Weiss PL, Tirosh E, Fehlings D. Role of virtual reality for cerebral palsy management. J Child Neurol 2014 Aug;29(8):1119-1124. [doi: 10.1177/0883073814533007] [Medline: 24799367]
26. Burke JW, McNeill MDJ, Charles DK, Morrow PJ, Crosbie JH, McDonough SM. Optimising engagement for stroke rehabilitation using serious games. Vis Comput 2009 Aug 27;25(12):1085-1099. [doi: 10.1007/s00371-009-0387-4]
27. Snider L, Majnemer A, Darsaklis V. Virtual reality as a therapeutic modality for children with cerebral palsy. Dev Neurorehabil 2010;13(2):120-128. [doi: 10.3109/17518420903357753] [Medline: 20222773]
28. Gordon AM, Okita SY. Augmenting pediatric constraint-induced movement therapy and bimanual training with video gaming technology. Technol Disabil 2010 Dec 01;22(4):179-191. [doi: 10.3233/TAD-2010-0302] [Medline: 32892679]

29. Rizzo A, Cohen I, Weiss P, Kim JG, Yeh SC, Zali B, et al. Design and development of virtual reality based perceptual-motor rehabilitation scenarios. Conf Proc IEEE Eng Med Biol Soc 2004;2004:4852-4855. [doi: 10.1109/IEMBS.2004.1404342] [Medline: 17271398]
30. Weiss P, Klinger E. Moving beyond single user, local virtual environments for rehabilitation. Stud Health Technol Inform 2009;145:263-276. [Medline: 19592799]
31. Bryanton C, Bossé J, Brien M, McLean J, McCormick A, Sveistrup H. Feasibility, motivation, and selective motor control: virtual reality compared to conventional home exercise in children with cerebral palsy. Cyberpsychol Behav 2006 Apr;9(2):123-128. [doi: 10.1089/cpb.2006.9.123] [Medline: 16640463]
32. Levac D, Sveistrup H. Motor Learning in Virtual Reality. In: Weiss PL, Keshner EA, Levin MF, editors. Virtual Reality for Physical and Motor Rehabilitation. New York (NY): Springer; 2014:25-43.
33. Rathinam C, Mohan V, Peirson J, Skinner J, Nethaji KS, Kuhn I. Effectiveness of virtual reality in the treatment of hand function in children with cerebral palsy: A systematic review. J Hand Ther 2019;32(4):426-434.e1. [doi: 10.1016/j.jht.2018.01.006] [Medline: 30017414]
34. Chen Y, Lee S, Howard AM. Effect of virtual reality on upper extremity function in children with cerebral palsy: a meta-analysis. Pediatr Phys Ther 2014;26(3):289-300. [doi: 10.1097/PEP.0000000000000046] [Medline: 24819682]
35. Galvin J, McDonald R, Catroppa C, Anderson V. Does intervention using virtual reality improve upper limb function in children with neurological impairment: a systematic review of the evidence. Brain Inj 2011;25(5):435-442. [doi: 10.3109/02699052.2011.558047] [Medline: 21401370]
36. Massetti T, Silva TD, Ribeiro D. Motor learning through virtual reality in cerebral palsy - a literature review. Med Express 2014;1(6):302-306. [doi: 10.5935/medicalexpress.2014.06.04]
37. 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]
38. 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]
39. Morton S, Barton CJ, Rice S, Morrissey D. Risk factors and successful interventions for cricket-related low back pain: a systematic review. Br J Sports Med 2014 Apr;48(8):685-691. [doi: 10.1136/bjsports-2012-091782] [Medline: 23666021]
40. 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]
41. O'Connor SR, Tully MA, Ryan B, Bradley JM, Baxter GD, McDonough SM. Failure of a numerical quality assessment scale to identify potential risk of bias in a systematic review: a comparison study. BMC Res Notes 2015 Jun 06;8:224 [FREE Full text] [doi: 10.1186/s13104-015-1181-1] [Medline: 26048813]
42. Effective Public Health Practice Project. Quality assessment tool for quantitative studies. National Collaborating Centre for Methods and Tools, McMaster University, Hamilton, Ontario. 1998. URL: https://merst.ca/wp-content/uploads/2018/02/quality-assessment-tool_2010.pdf [accessed 2021-03-24]
43. Subramanian S, Caramba S, Hernandez O, Morgan Q, Cross M, Hirschhauser C. Is the Downs and Black scale a better tool to appraise the quality of the studies using virtual rehabilitation for post-stroke upper limb rehabilitation? InInternational Conference on Virtual Rehabilitation (ICVR). : IEEE; 2019 Presented at: International Conference on Virtual Rehabilitation; 2019; Tel Aviv, Israel. [doi: 10.1109/icvr46560.2019.8994724]
44. Avcil E, Tarakci D, Arman N, Tarakci E. Upper extremity rehabilitation using video games in cerebral palsy: a randomized clinical trial. Acta Neurol Belg 2020 Jun 11. [doi: 10.1007/s13760-020-01400-8]
45. Bedair R, Al-Talawy H, Shoukry K, Abdul-Raouf E. Impact of virtual reality games as an adjunct treatment tool on upper extremity function of spastic hemiplegic children. Int J PharmTech Res 2016;9(6):1-8 [FREE Full text]
46. Chen Y, Kang L, Chuang T, Doong J, Lee S, Tsai MW, et al. Use of virtual reality to improve upper-extremity control in children with cerebral palsy: a single-subject design. Phys Ther 2007 Nov;87(11):1441-1457. [doi: 10.2522/ptj.20060062] [Medline: 17895352]
47. Sharan D, Ajeesh PS, Rameshkumar R, Mathankumar M, Paulina RJ, Manjula M. Virtual reality based therapy for post operative rehabilitation of children with cerebral palsy. Work 2012;41 Suppl 1:3612-3615. [doi: 10.3233/WOR-2012-0667-3612] [Medline: 22317271]
48. El-Shamy SM. Efficacy of Armeo® Robotic Therapy Versus Conventional Therapy on Upper Limb Function in Children With Hemiplegic Cerebral Palsy. Am J Phys Med Rehabil 2018 Mar;97(3):164-169. [doi: 10.1097/PHM.0000000000000852] [Medline: 29059068]
49. Fluet G, Qiu Q, Saleh S. Robot-assisted virtual rehabilitation (NJIT-RAVR) system for children with upper extremity hemiplegia. 2009 Presented at: International Conference on Virtual Rehabilitation; 9 June-2 July 2009; Haifa, Israel p. 189-192. [doi: 10.1109/icvr.2009.5174230]

50. Fluet GG, Qiu Q, Kelly D, Parikh HD, Ramirez D, Saleh S, et al. Interfacing a haptic robotic system with complex virtual environments to treat impaired upper extremity motor function in children with cerebral palsy. Dev Neurorehabil 2010;13(5):335-345 [FREE Full text] [doi: 10.3109/17518423.2010.501362] [Medline: 20828330]
51. Hernández HA, Khan A, Fay L, Roy J, Biddiss E. Force Resistance Training in Hand Grasp and Arm Therapy: Feasibility of a Low-Cost Videogame Controller. Games Health J 2018 Aug;7(4):277-287. [doi: 10.1089/g4h.2017.0193] [Medline: 30106641]
52. Jannink MJA, van der Wilden GJ, Navis DW, Visser G, Gussinklo J, Ijzerman M. A low-cost video game applied for training of upper extremity function in children with cerebral palsy: a pilot study. Cyberpsychol Behav 2008 Feb;11(1):27-32. [doi: 10.1089/cpb.2007.0014] [Medline: 18275309]
53. Kassee C, Hunt C, Holmes MW, Lloyd M. Home-based Nintendo Wii training to improve upper-limb function in children ages 7 to 12 with spastic hemiplegic cerebral palsy. J Pediatr Rehabil Med 2017 May 17;10(2):145-154. [doi: 10.3233/PRM-170439] [Medline: 28582885]
54. Odle B, Irving A, Foulds R. Usability of an adaptable video game platform for children with cerebral palsy. 2009 Presented at: Bioeng Proc Northeast Conf; 3-5 April 2009; Cambridge, MA, USA p. 1-2. [doi: 10.1109/nebc.2009.4967748]
55. Rostami HR, Arastoo AA, Nejad SJ, Mahany MK, Malamiri RA, Goharpey S. Effects of modified constraint-induced movement therapy in virtual environment on upper-limb function in children with spastic hemiparetic cerebral palsy: a randomised controlled trial. NeuroRehabilitation 2012;31(4):357-365. [doi: 10.3233/NRE-2012-00804] [Medline: 23232158]
56. Şahin S, Köse B, Aran OT, Bahadır Ağce Z, Kayıhan H. The Effects of Virtual Reality on Motor Functions and Daily Life Activities in Unilateral Spastic Cerebral Palsy: A Single-Blind Randomized Controlled Trial. Games Health J 2020 Feb;9(1):45-52. [doi: 10.1089/g4h.2019.0020] [Medline: 31335174]
57. Sandlund M, Domellöf E, Grip H, Rönnqvist L, Häger CK. Training of goal directed arm movements with motion interactive video games in children with cerebral palsy - a kinematic evaluation. Dev Neurorehabil 2014 Oct;17(5):318-326. [doi: 10.3109/17518423.2013.776124] [Medline: 23863100]
58. Turconi AC, Biffi E, Maghini C, Peri E, Servodio Iammarone F, Gagliardi C. Can new technologies improve upper limb performance in grown-up diplegic children? Eur J Phys Rehabil Med 2016 Oct;52(5):672-681 [FREE Full text] [Medline: 26554345]
59. Weightman A, Preston N, Levesley M, Holt R, Mon-Williams M, Clarke M, et al. Home based computer-assisted upper limb exercise for young children with cerebral palsy: a feasibility study investigating impact on motor control and functional outcome. J Rehabil Med 2011 Mar;43(4):359-363 [FREE Full text] [doi: 10.2340/16501977-0679] [Medline: 21347508]
60. Winkels DGM, Kottink AIR, Temmink RAJ, Nijlant JMM, Buurke JH. Wii™-habilitation of upper extremity function in children with cerebral palsy. An explorative study. Dev Neurorehabil 2013;16(1):44-51. [doi: 10.3109/17518423.2012.713401] [Medline: 23030054]
61. Do J, Yoo E, Jung M, Park HY. The effects of virtual reality-based bilateral arm training on hemiplegic children's upper limb motor skills. NeuroRehabilitation 2016;38(2):115-127. [doi: 10.3233/NRE-161302] [Medline: 26923353]
62. Sandlund M, Waterworth EL, Häger C. Using motion interactive games to promote physical activity and enhance motor performance in children with cerebral palsy. Dev Neurorehabil 2011;14(1):15-21. [doi: 10.3109/17518423.2010.533329] [Medline: 21241174]
63. Chang Y, Chen S, Huang J. A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. Res Dev Disabil 2011;32(6):2566-2570. [doi: 10.1016/j.ridd.2011.07.002] [Medline: 21784612]
64. Reifenberg G, Gabrosek G, Tanner K, Harpster K, Proffitt R, Persch A. Feasibility of Pediatric Game-Based Neurorehabilitation Using Telehealth Technologies: A Case Report. Am J Occup Ther 2017;71(3):7103190040p1-7103190040p8. [doi: 10.5014/ajot.2017.024976] [Medline: 28422630]
65. Rios DC, Gilbertson T, McCoy SW, Price R, Gutman K, Miller KEF, et al. NeuroGame Therapy to improve wrist control in children with cerebral palsy: a case series. Dev Neurorehabil 2013 Dec;16(6):398-409. [doi: 10.3109/17518423.2013.766818] [Medline: 23617243]
66. Robert M, Guberek R, Levin M, Sveistrup H. Motor learning of the upper limb in children with cerebral palsy after virtual and physical training intervention. 2013 Presented at: International Conference on Virtual Rehabilitation ICVR; 2013; Valencia, Spain p. 194-195. [doi: 10.1109/icvr.2013.6662125]
67. van Hedel H, Wick K, Meyer-Heim A, Eng K. Improving dexterity in children with cerebral palsy. : IEEE; 2011 Presented at: International Conference on Virtual Rehabilitation ICVR; 2011; Zurich, Switzerland p. 1-2. [doi: 10.1109/icvr.2011.5971872]
68. Dinomais M, Veaux F, Yamaguchi T, Richard P, Richard I, Nguyen S. A new virtual reality tool for unilateral cerebral palsy rehabilitation: two single-case studies. Dev Neurorehabil 2013 Dec;16(6):418-422. [doi: 10.3109/17518423.2013.778347] [Medline: 23845037]
69. Golomb MR, Warden SJ, Fess E, Rabin B, Yonkman J, Shirley B, et al. Maintained hand function and forearm bone health 14 months after an in-home virtual-reality videogame hand telerehabilitation intervention in an adolescent with hemiplegic cerebral palsy. J Child Neurol 2011 Mar;26(3):389-393 [FREE Full text] [doi: 10.1177/0883073810394847] [Medline: 21383228]

70. Huber M, Rabin B, Docan C. PlayStation 3-based tele-rehabilitation for children with hemiplegia. 2008 Presented at: International Conference on Virtual Rehabilitation ICVR; Aug 25-28, 2008; Vancouver, Canada p. 105-112. [doi: 10.1109/icvr.2008.4625145]
71. Qiu Q, Ramirez DA, Saleh S, Fluet GG, Parikh HD, Kelly D, et al. The New Jersey Institute of Technology Robot-Assisted Virtual Rehabilitation (NJIT-RAVR) system for children with cerebral palsy: a feasibility study. J Neuroeng Rehabil 2009 Nov 16;6:40 [FREE Full text] [doi: 10.1186/1743-0003-6-40] [Medline: 19917124]
72. Chang Y, Chen S, Huang J. A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. Res Dev Disabil 2011;32(6):2566-2570. [doi: 10.1016/j.ridd.2011.07.002] [Medline: 21784612]
73. Golomb MR, McDonald BC, Warden SJ, Yonkman J, Saykin AJ, Shirley B, et al. In-home virtual reality videogame telerehabilitation in adolescents with hemiplegic cerebral palsy. Arch Phys Med Rehabil 2010 Jan;91(1):1-8.e1. [doi: 10.1016/j.apmr.2009.08.153] [Medline: 20103390]
74. Charles J, Gordon AM. Development of hand-arm bimanual intensive training (HABIT) for improving bimanual coordination in children with hemiplegic cerebral palsy. Dev Med Child Neurol 2006 Nov;48(11):931-936 [FREE Full text] [doi: 10.1017/S0012162206002039] [Medline: 17044964]
75. Gordon AM, Charles J, Wolf SL. Methods of constraint-induced movement therapy for children with hemiplegic cerebral palsy: development of a child-friendly intervention for improving upper-extremity function. Arch Phys Med Rehabil 2005 Apr;86(4):837-844. [doi: 10.1016/j.apmr.2004.10.008] [Medline: 15827942]
76. Nudo R. Adaptive plasticity in motor cortex: implications for rehabilitation after brain injury. J Rehabil Med 2003 May(41 Suppl):7-10 [FREE Full text] [doi: 10.1080/16501960310010070] [Medline: 12817650]
77. Demers M, Mbiya N, Levin M. Industry and academia collaboration in the design of virtual reality applications for rehabilitation. : IEEE; 2017 Presented at: International Conference on Virtual Rehabilitation; June 19-22, 2017; Montreal, Canada p. 1-2. [doi: 10.1109/ICVR.2017.8007545]
78. Winstein CJ, Schmidt RA. Reduced frequency of knowledge of results enhances motor skill learning. J Exp Psychol Learn Mem Cogn 1990;16(4):677-691. [doi: 10.1037/0278-7393.16.4.677]
79. Laufer Y, Weiss PL. Virtual reality in the assessment and treatment of children with motor impairment: A systematic review. J Phys Ther Educ 2011;25(1):59-71. [doi: 10.1097/00001416-201110000-00011]
80. 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]
81. Chiviacowsky S, Wulf G, de Medeiros FL, Kaefer A, Wally R. Self-controlled feedback in 10-year-old children: higher feedback frequencies enhance learning. Res Q Exerc Sport 2008 Mar;79(1):122-127. [doi: 10.1080/02701367.2008.10599467] [Medline: 18431958]
82. Hemayattalab R, Arabameri E, Pourazar M, Ardakani MD, Kashefi M. Effects of self-controlled feedback on learning of a throwing task in children with spastic hemiplegic cerebral palsy. Res Dev Disabil 2013 Sep;34(9):2884-2889. [doi: 10.1016/j.ridd.2013.05.008] [Medline: 23810928]
83. Subramanian SK, Massie CL, Malcolm MP, Levin MF. Does provision of extrinsic feedback result in improved motor learning in the upper limb poststroke? A systematic review of the evidence. Neurorehabil Neural Repair 2010 Feb;24(2):113-124. [doi: 10.1177/1545968309349941] [Medline: 19861591]
84. Schüler T, Ferreira dos Santos L, Hoermann S. Designing virtual environments for motor rehabilitation: Towards a framework for the integration of best-practice information. 2015 Presented at: International Conference on Virtual Rehabilitation; 2015; Valencia, Spain p. 145-146. [doi: 10.1109/ICVR.2015.7358620]
85. Guadagnoli MA, Lee TD. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. J Mot Behav 2004 Jun;36(2):212-224. [doi: 10.3200/JMBR.36.2.212-224] [Medline: 15130871]
86. Metcalfe J. Learning from Errors. Annu Rev Psychol 2017 Jan 03;68:465-489. [doi: 10.1146/annurev-psych-010416-044022] [Medline: 27648988]
87. Levac DE, Miller PA. Integrating virtual reality video games into practice: clinicians' experiences. Physiother Theory Pract 2013 Oct;29(7):504-512. [doi: 10.3109/09593985.2012.762078] [Medline: 23362843]
88. Majnemer A. Importance of motivation to children's participation: a motivation to change. Phys Occup Ther Pediatr 2011 Feb;31(1):1-3. [doi: 10.3109/01942638.2011.541747] [Medline: 21174548]
89. Meyns P, Roman de Mettelinge T, van der Spank J, Coussens M, Van Waelvelde H. Motivation in pediatric motor rehabilitation: A systematic search of the literature using the self-determination theory as a conceptual framework. Dev Neurorehabil 2018 Aug;21(6):371-390. [doi: 10.1080/17518423.2017.1295286] [Medline: 28277817]
90. Majnemer A, Shikako-Thomas K, Lach L, Shevell M, Law M, Schmitz N, QUALA group. Mastery motivation in adolescents with cerebral palsy. Res Dev Disabil 2013 Oct;34(10):3384-3392. [doi: 10.1016/j.ridd.2013.07.002] [Medline: 23911644]
91. Tatla SK, Sauve K, Jarus T, Virji-Babul N, Holsti L. The effects of motivating interventions on rehabilitation outcomes in children and youth with acquired brain injuries: a systematic review. Brain Inj 2014;28(8):1022-1035. [doi: 10.3109/02699052.2014.890747] [Medline: 24661000]
92. Harris K, Reid D. The influence of virtual reality play on children's motivation. Can J Occup Ther 2005 Feb;72(1):21-29. [doi: 10.1177/000841740507200107] [Medline: 15727045]

93. Kitago T, Krakauer J. Motor learning principles for neurorehabilitation. In: Barnes MP, Good DC, editors. Handbook of Clinical Neurology. Vol 110. Amsterdam, Netherlands: Elsevier; 2013:93-103.
94. Levac DE, Huber ME, Sternad D. Learning and transfer of complex motor skills in virtual reality: a perspective review. J Neuroeng Rehabil 2019 Oct 18;16(1):121 [FREE Full text] [doi: 10.1186/s12984-019-0587-8] [Medline: 31627755]
95. Yanovich E, Ronen O. The use of virtual reality in motor learning: A multiple pilot study review. APE 2015;05(03):188-193. [doi: 10.4236/ape.2015.53023]
96. Breslin G, Hodges NJ, Steenson A, Williams AM. Constant or variable practice: recreating the especial skill effect. Acta Psychol (Amst) 2012 Jun;140(2):154-157. [doi: 10.1016/j.actpsy.2012.04.002] [Medline: 22627159]
97. Lage GM, Ugrinowitsch H, Apolinário-Souza T, Vieira MM, Albuquerque MR, Benda RN. Repetition and variation in motor practice: A review of neural correlates. Neurosci Biobehav Rev 2015 Oct;57:132-141. [doi: 10.1016/j.neubiorev.2015.08.012] [Medline: 26299808]
98. Bernstein N. The Co-ordination and Regulation of Movements. Oxford, United Kingdom: Pergamon Press; Mar 1968.
99. Guadagnoli MA, Holcomb WR, Weber TJ. The relationship between contextual interference effects and performer expertise on the learning of a putting task. J Hum Mov Stud 1999;37(1):19-36.
100. Zipp GP, Gentile AM. Practice Schedule and the Learning of Motor Skills in Children and Adults: Teaching Implications. Journal of College Teaching & Learning 2010 Nov 03;7(2). [doi: 10.19030/tlc.v7i2.87]
101. Pinto-Zipp G, Gentile A. Practice schedule in motor learning: children versus adults. 1995 Presented at: Soc of Neurosci Abst; Nov 11-16, 1995; San Diego, CA.
102. Levin M, Sveistrup H, Subramanian S. Feedback and virtual environments for motor learning and rehabilitation. Schedae 2010;1:19-36 [FREE Full text]

Abbreviations

CP: cerebral palsy
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by N Zary; submitted 25.08.20; peer-reviewed by J Li, S Hoermann, N Mohammad Gholi Mezerji; comments to author 08.10.20; revised version received 06.11.20; accepted 13.03.21; published 07.04.21.

Please cite as:
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;9(2):e23822
URL: https://games.jmir.org/2021/2/e23822
doi:10.2196/23822
PMID:33825690

Original Paper

A Gamification Framework for Cognitive Assessment and Cognitive Training: Qualitative Study

Ali Khaleghi[1], PhD; Zahra Aghaei[1], MSc; Mohammad Amin Mahdavi[1], PhD

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

Corresponding Author:
Ali Khaleghi, PhD
Department of Computer Engineering
Imam Khomeini International University
Norouzian Boulevard
Qazvin
Iran
Phone: 98 9121003006
Email: ali.khaleghi.ir@gmail.com

Abstract

Background: Cognitive tasks designed to measure or train cognition are often repetitive and presented in a monotonous manner, features that lead to participant boredom and disengagement. In this situation, participants do not put forth their best effort to do these tasks well. As a result, neuropsychologists cannot draw accurate conclusions about the data collected, and intervention effects are reduced. It is assumed that greater engagement and motivation will manifest as improved data quality. Gamification, the use of game elements in nongame settings, has been heralded as a potential mechanism for increasing participant engagement in cognitive tasks. Some studies have reported a positive effect of gamification on participant performance, although most studies have shown mixed results. One reason for these contrasting findings is that most studies have applied poor and heterogeneous design techniques to gamify cognitive tasks. Therefore, an appropriate gamification design framework is needed in these tasks.

Objective: This study aimed to propose a framework to guide the design of gamification in cognitive tasks.

Methods: We employed a design science research (DSR) approach to provide a framework for gamifying cognitive assessments and training by synthesizing current gamification design frameworks and gamification works in cognitive assessment and training, as well as incorporating field experiences. The prototypes of the framework were iteratively evaluated with 17 relevant experts.

Results: We proposed a framework consisting of 7 phases: (1) preparation; (2) knowing users; (3) exploring existing tools for assessing or training a targeted cognitive context and determining the suitability of game-up and mapping techniques; (4) ideation; (5) prototyping using the Objects, Mechanics, Dynamics, Emotions (OMDE) design guideline; (6) development; and (7) disseminating and monitoring.

Conclusions: We found that (1) an intermediate design framework is needed to gamify cognitive tasks, which means that game elements should be selected by considering current cognitive assessment or training context characteristics since game elements may impose an irrelevant cognitive load that, in turn, can jeopardize data quality; (2) in addition to developing a new gamified cognitive task from scratch, 2 gamification techniques are widely used (first, adding game elements to an existing cognitive task and second, mapping an existing game to a cognitive function or impairment to assess or train it); and (3) further research is required to investigate the interplay of cognitive processes and game mechanics.

(JMIR Serious Games 2021;9(2):e21900) doi:10.2196/21900

KEYWORDS

cognitive tasks; boredom; motivation; gamification; game elements; framework; process; gamification design; cognitive training; cognitive assessment

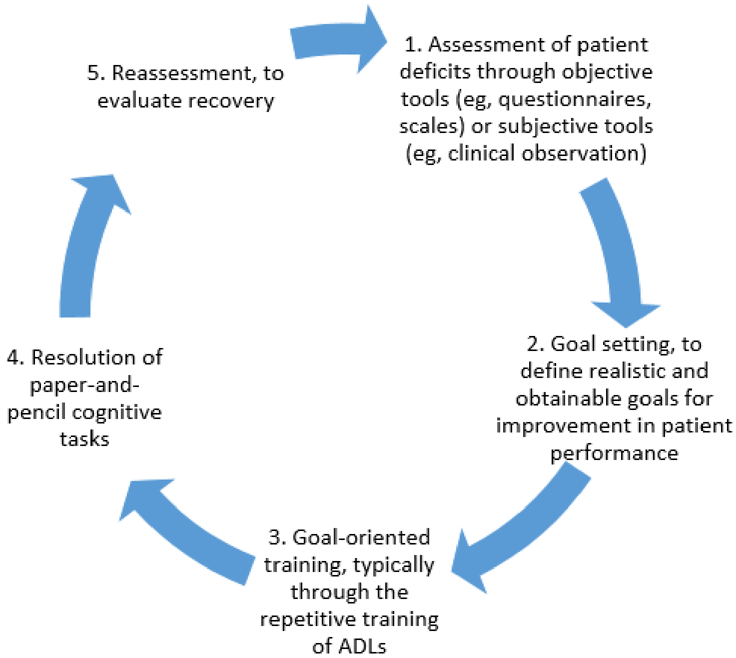
Introduction

Background

Statistics show that the cognitive assessment and training market will achieve a growth rate of 32.39% from 2018 to 2022, and gamification will be one of the key vendors operating in this market [1]. These statistics confirm another estimation that shows 1 out of 5 people will be over 60 years old in the next 40 years [2]. Minor and major neurocognitive disorders, which have prevalences of approximately 10%-20% and 5%-7%, respectively, are global health issues due to the aging population; according to the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders, edition 5, minor and major neurocognitive disorders encompass mild cognitive impairment and dementia, respectively [3]. In minor neurocognitive disorders, cognitive abilities decline, but changes are not severe enough to significantly affect individuals' activities of daily living (ADLs); 35% of individuals with a minor neurocognitive disorder progress to a major neurocognitive disorder within a 3-year period [4]. In this stage, individuals lose independence in their ADLs and require care and support from others. It is not only aging that causes cognitive impairments; they may also exist since childhood (such as attention deficit hyperactivity disorder [ADHD] and autism spectrum disorder) or because of different factors like alcohol or drug abuse [5-7].

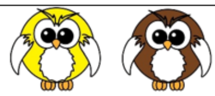
Cognitive assessment and training play an essential role in preventing loss of autonomy and independence in ADLs [8]. Cognitive assessment is associated with evaluating individuals' cognitive abilities (eg, working memory, attention, executive functions) [9]. Cognitive training refers to using cognitive tasks to maintain or improve a particular aspect of cognitive functioning [9]. There are cognitive tasks to assess or train cognitive functions. Cognitive tasks such as the Mini-Mental State Examination (MMSE) [10], Confusion Assessment Method (CAM) [11], and Montreal Cognitive Assessment (MoCA) [12] are widely used for cognitive assessment. There are also cognitive training companies that have developed cognitive training tasks such as Cogmed [13], Nintendo Brain Age [14], Lumosity [15], and Posit Science BrainHQ [16]. The prevailing approach in these companies for designing a training task is to convert lab-based and individualistic cognitive assessment tasks into a training task [17] (eg, Nintendo Brain Age includes several cognitive assessment tasks like the Stroop task [18]). Cognitive tasks are a vital tool for the assessment and training of cognitive impairments. However, participants often view them as monotonous and boring since they have a repetitive nature and are rigidly presented [8,9,19-23]. These features increase the frequency of insufficient efforts to perform these tasks, and consequently, the reliability of data collected decreases [9,19,23-26]. The typical process taken by cognitive specialists to perform a cognitive training program is shown in Figure 1.

Figure 1. The process of traditional cognitive training programs (the figure was drawn based on descriptions provided by Vourvopoulos et al [8]). ADLs: activities of daily living.



It is assumed that greater engagement and motivation will manifest as improved data quality in cognitive tasks [9,19-27]. Among the existing solutions, gamification, which is the process of adding game elements (eg, scoring system, leaderboard, badge) to nongame contexts (eg, education context, business context, cognitive tasks) [28], stands as one of the most influential and promising solutions to improve motivation in monotonous tasks [9]. A greater understanding of human motivation helps maintain users' encouragement to participate in cognitive tasks over time [17]. Motivation is multidimensional and falls on a continuum from intrinsic motivation to extrinsic motivation to amotivation (little to no motivation exists) [29]. Intrinsic motivation is regulated internally and refers to performing activities for their inherent satisfaction. Intrinsic motivation is required for long-term engagement and long-term changes. In contrast, extrinsic motivation (doing activities solely for their outcomes and rewards) is useful for short-term engagement and short-term changes and is also regulated externally [29,30]. Self-determination theory (SDT) [31] and Flow [32] theory are widely used to improve users' participation and motivation. According to SDT [31], intrinsic motivation can be sustained by satisfying 3 psychological needs of relatedness (is experienced when individuals feel connected to others), autonomy (the need for freedom to make choices based on one's volition during an activity), and competence (the need for challenge and feelings of self-efficacy). Flow refers to "the state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it" [32]. Gamification can combine intrinsic motivation with extrinsic motivation to raise motivation and engagement [30]. Game elements such as badges, points, game levels, a leaderboard, and avatars lead to extrinsic motivation and are useful for capturing early user motivation [17]. Based on SDT [31] and Flow [32] theory, gamification can also improve the intrinsic motivation of participants through game elements such as optimal challenges and positive feedback (these elements satisfy human needs of competence). Figure 2 presents some examples of gamified cognitive tasks from Craven and Groom [33], Boendermaker et al [34], Lumsden et al [9], and Van de Weijer-Bergsma et al [35].

Figure 2. Examples of gamified cognitive tasks. A task was categorized as an assessment or training according to the self-prescription provided by the authors of the publication. The owl and apple pictures were used or adapted by [33] from openclipart [36] and 4vector [37] (online media repository of free graphics). ADHD: attention deficit hyperactivity disorder.

Author, year	Cognitive task	Targeted cognitive function(s) or disorder(s)	Assessment or training	Short description	Corresponding pictures
Craven and Groom, 2015 [33]	Go/No-Go	ADHD	Both	Focuses on gamifying the Go/No-Go task, which is a cognitive task used to assess the ability to inhibit a response. Participants must respond or withhold a reply depending on whether a Go or No-Go stimulus is presented. In the gamified version (awkward owls), a narrative is given for participants. They play a conservation volunteer's who collects rare yellow owls (Go stimulus) for a wildlife survey and ignore brown owls (No-Go stimulus). After each response, depending on the displayed stimulus, the message "Hoot!" or "Boo!!!" is shown.	
Craven and Groom, 2015 [33]	Stop Signal Task (SST)	ADHD	Both	Here, SST was gamified, called wormy fruit. A sequence of fruits is presented. Fruits and wormy fruits are the Go and Stop signals, respectively. Eaten fruit and wormy fruit result from a response to the stimulus and missing the Stop, including clicking early before the Stop when the fruit is wormy.	
Boendermaker et al, 2015 [34]	Go/No-Go	Working memory	Training	To train alcohol bias, Go/No-Go was gamified in a game called cheese ninja. The game consists of a mouse ninja that walks through a hallway and passes posters of the same beverages used in the regular task. A good (cheese) or bad (cat) cue is shown in front of the posters. The player should collect as many useful objects as possible while ignoring the bad ones.	Not provided
Lumsden et al, 2016 [9]	Go/No-Go	Response inhibition	Assessment	Two gamified versions of the Go/ No-Go task were developed to assess response inhibition ability. The standard task contains 20 everyday objects: 15 green (Go stimulus) and 5 red (No-Go stimulus) objects. The first gamified variant was identical to the standard task except that a point system based on that of Miranda and Palmer [27] was incorporated. In the second gamified version, the original task was themed as a cowboy shootout game in which participants are the sheriff of a town and a group of criminals hole up in a saloon and take hostages. During the game, a cartoon saloon appears, and stimuli are shown in the doorway. The participants have to shoot cowboys as Go targets and hold their response when they see innocent civilians as No-Go targets. By responding to the displayed No-Go stimulus, a blood splat overlays on it for the remainder of the game time.	 
Van de Weijer-Bergsma et al, 2014 [35]	Working memory task	Visual-spatial working memory, updating, shifting, and inhibition skills	Assessment	Focuses on gamifying a working memory task for children, named the lion game. A 4x4 matrix containing 16 bushes is presented. In different locations for 2000 ms, 8 lions of different colors are continuously shown. Participants should remember the last place where a lion of a specific color (e.g., blue) appears using the mouse button to click on that location after the sequence ends. The game consists of 5 levels, in which working memory capacity is manipulated by changing lions' colors and locations.	

Challenges With Designing Gamification for Cognitive Assessment or Training

Despite the growing trend towards using gamification in cognitive tasks, its impacts on participant engagement and data quality are not stable. Studies have [20-22,27,38-40] stated that gamification has a positive effect on users' engagement as well as data quality. However, other studies [19,24,26,34] have reported that no effects were observed on users' performance by adding game elements, but they perceived the gamified task as funnier and more challenging than the nongamified version. In addition, other studies [19,27] showed that gamification worsened data quality but had positive effects on engagement. The gamification applied by Birk et al [23] not only did not have a positive impact on data quality and engagement but also worsened them [19]. These mixed findings are potentially due to 4 main reasons: (1) Most gamified cognitive tasks have been developed by cognitive psychologists, not professional gamification designers, and for scientists, the clinical effectiveness of a gamified task is important, with less focus on employing effective and creative gamification designs [41]; (2) a variety of gamification techniques have been applied to cognitive tasks [9,24,38]; (3) gamification techniques have been applied to different cognitive tasks [24]; and (4) the results obtained from gamified cognitive tasks are often preliminary and limited by small sample sizes. Also, the considered duration for evaluating the efficacy of gamified tasks is relatively short [9,23-25,42,43].

There are recommendations and design guidelines to integrate game elements into cognitive tasks (eg, [9,30]). However, to our knowledge, they did not propose a detailed and step-by-step framework that clearly shows what factors are essential to designing gamification in these tasks from early stages (eg, planning and preparation phases) to develop, evaluate, and disseminate gamified tasks, followed by monitoring the efficacy of such tasks in the long term. In general, several gamification frameworks have been developed by experts, not designed for cognitive tasks (eg, [44-53]). However, they suffer from 3 main limitations: (1) Most of them have been designed for enterprise and business contexts, with less focus on health contexts [54]; (2) cognitive tasks need to engage participants for the long term, but existing frameworks have not been designed for this purpose [54,55]; and (3) they have not specified how game elements should be added to a particular context. Incorporating game elements in cognitive tasks may jeopardize data quality by imposing an additional cognitive load to these tasks [56]. For example, Katz et al [25] gamified the N-Back task by including a real-time scoring system while completing the task. The gamified task, in comparison to the actual task, negatively impacted data quality. One possible explanation is that the game features imposed irrelevant cognitive demands by distracting the players' attention.

Objectives

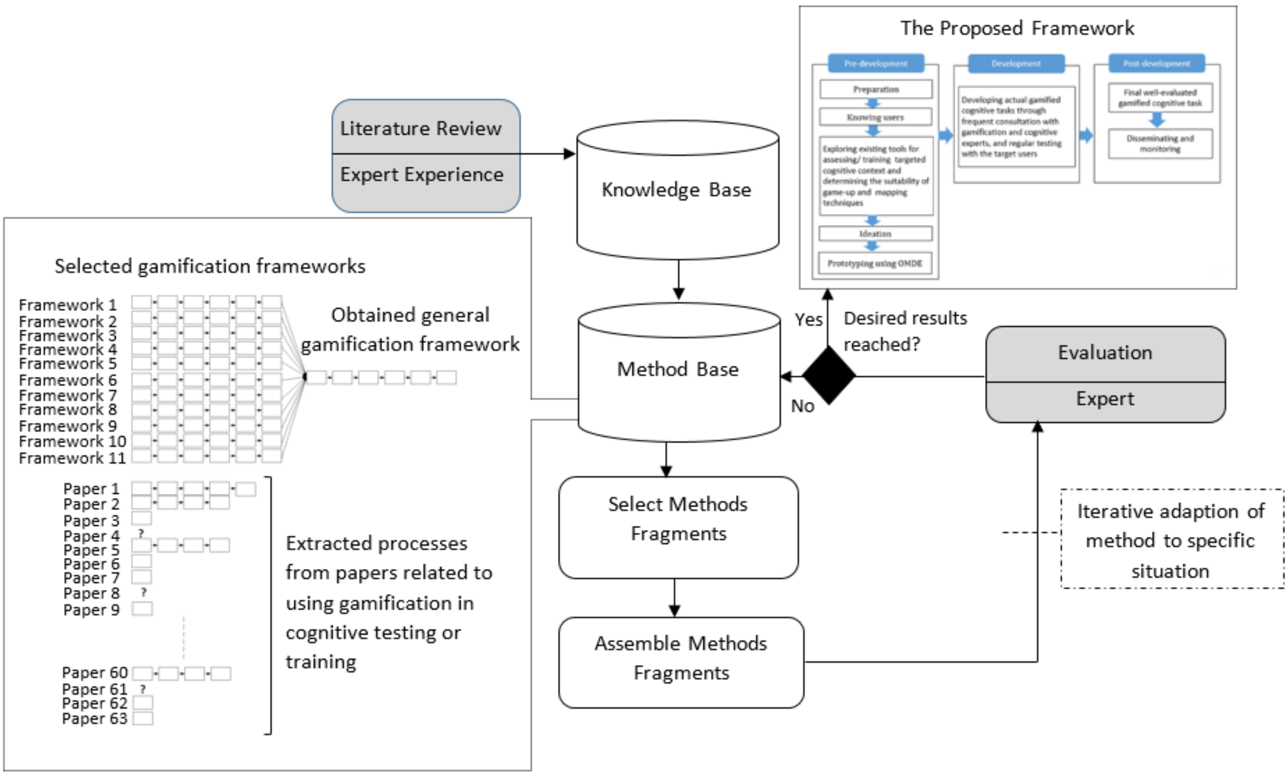
Despite the papers that have shown mixed findings of using game elements in cognitive tasks, we assume that gamification can positively influence data quality and user engagement. Therefore, we are proposing a framework to guide the process of incorporating game elements in cognitive tasks by synthesizing (1) existing gamification design frameworks, (2) gamification efforts in cognitive assessment and training, and (3) field experiences.

Methods

Overview

We approached the research problem through the design science research (DSR) methodology [57]. Design science is an accepted research methodology in information systems. It emphasizes that research should be firmly grounded in existing knowledge and target the context in which the developed artifact must be used, to create scientifically sound artifacts (eg, theories, models, and methods) [58]. The 2 main steps for conducting DSR include (1) developing artifacts and (2) evaluation of the developed artifacts [53]. In our DSR approach, similar to Morschheuser et al [53], we employed an assembly-based situational method engineering methodology proposed by Brinkkemper [59]. Method engineering is an approach in information systems to develop new methods from recognized fragments of existing methods knowledge to propose a situational method tuned to the situation of the project at hand. According to method engineering, 3 phases are needed to develop a new method [59]: (1) building a method database, which includes all the resources required for the development of a new situational method; (2) constructing the situational method through assembling of the methods fragments from the method database; and (3) evaluation of the developed method. Figure 3 provides an overview of our assembly-based situational method engineering.

Figure 3. Situational method engineering approach followed for developing the framework (adapted from [53,59]).



Knowledge Base

Defining Sources of Evidence

We selected the following 3 resources to gather the knowledge required for proposing the intended framework: (1) studies that proposed a gamification design framework, since by synthesizing them, we could extract the general framework for the design of gamification; (2) projects that used gamification in cognitive assessment or training to extract key factors and considerations for gamifying these contexts; (3) the experiences of relevant experts to integrate the evidence from live environments.

Exploring Relevant Papers

We used systematic literature review strategies in the Google search engine, Google Scholar, PubMed, and Research Gate to use a wide variety of relevant papers. The search was started in May 2017 and lasted until the submission date. Two search strings were developed based on keywords, their synonyms, and related terms: (1) ("gamification" AND ("framework" OR "platform" OR "process" OR "method")) and (2) ("gamification" OR "serious games" OR "game up" OR "video game") AND (("cognitive" AND ("training "OR "assessment")) OR "Go /No-Go" OR "N-Back" OR "MMSE" OR "MoCA" OR "ADHD" OR "stop-signal task" OR "dyslexia"). Furthermore, search strategies such as checking the reference lists of included studies and cited reference searching were applied.

Inclusion Criteria for Selecting Gamification Frameworks

To select the most highly regarded gamification frameworks, the following 4 metrics were used: (1) framework was not focused on the parts or steps of the gamification design process, but covering the maximum number of steps; (2) framework was determined to be worthy in terms of efficacy by calculating the number of its citations; (3) framework was developed by gamification experts (we considered an individual an expert based on whether she or he published at least 10 scientific articles concerning gamification issues); (4) framework was developed using a robust methodology.

Inclusion Criteria for Selecting Gamification Projects in Cognitive Assessment and Training

First, we included projects that published reports about the impacts of game elements on data quality and user engagement in the form of a scientific paper, to obtain sufficient author-presented analytical expressions on how game elements should be incorporated into cognitive tasks. Second, similar to Lumsden et al [9], we did not select projects based on whether they included the word "gamification"; instead, we selected projects if our search strategies found them. We intentionally did not define gamification as defined by Deterding et al [28] ("the use of game design elements in non-gaming contexts"), since precisely defining the elements that make a game is challenging and limiting [9]. Therefore, we decided that a cognitive task was gamified if its purpose was to increase participants' commitment and motivation. However, it used other game-inspired designs such as serious games, video games, games with a purpose, and game-like interventions. We erred on the side of caution to minimize the potential loss of relevant sources.

Expert Evaluation

From the early stages, the framework was screened and judged by a homogeneous group of 17 experts from relevant disciplines, including information technology, game, gamification, and cognitive psychology. The average years of experience are presented in Table 1.

Table 1. Experts' background and average years of experience.

Background	Average experience (years)
Information technology (IT)	20
Game	4/5
Gamification	3
Cognitive psychology	11 years of academic experience and 6 years of clinical experience

After extracting each piece of evidence and then applying it to the under-development framework, the whole of the framework was visualized for expert evaluation. Then, the framework was refined based on the feedback collected.

Method Base

Extracting the General Gamification Framework

After comparing the selected frameworks in terms of main characteristics, merits, and demerits, 2 general gamification design frameworks were elaborated. The first framework was based on analyses presented in Multimedia Appendix 1. The second framework was based on the Mechanics, Dynamics, Aesthetics (MDA) process, a formal framework for designing and analyzing games [60], and 2 adapted versions of MDA presented in [44,52], now known as Objects, Mechanics, Dynamics, Emotions (OMDE). The phases and activities of these frameworks will be described in detail in the subsequent sections while describing the proposed framework.

Customizing the General Gamification Frameworks for Cognitive Assessment and Training

An exhaustive number (n=63) of empirical project reports or theoretical works that applied gamification into cognitive tasks were gathered to customize the obtained general frameworks for cognitive assessment and training contexts. We tried to extract an abstract process for each paper by observing the papers' different sections. As abstracted in Figure 3, most articles did not use a specific or formal framework for gamifying cognitive tasks. Therefore, either we could not extract a process, or the process obtained consisted of only one chunk. Finally, the isolated processes, fragments, and general gamification frameworks were converged based on their commonalities and unique features to assemble the intended framework.

Results

Search Results

The Explored Gamification Design Frameworks

We identified a total of 35 gamification design frameworks (these frameworks are listed in Multimedia Appendix 1). Of these, 11 frameworks were selected for more in-depth analysis (these frameworks are highlighted in bold in Multimedia Appendix 1).

The Explored Gamification Projects in Cognitive Assessment and Training

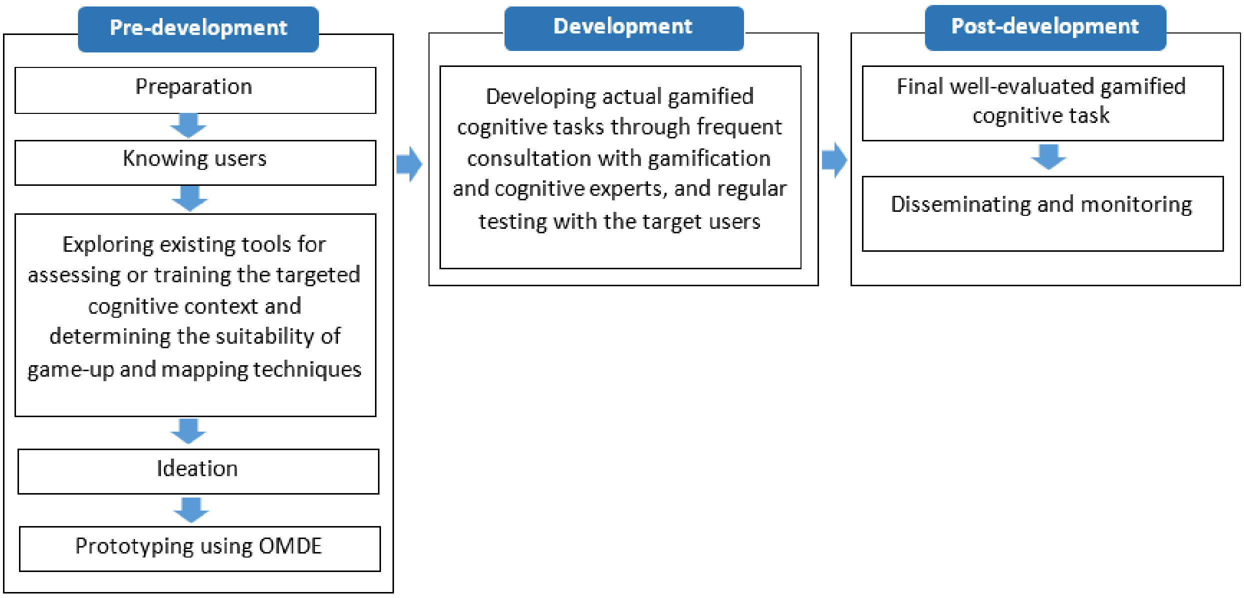
A total of 63 gamification projects in cognitive assessment or training was selected. Of these, 41 (41/63, 64%) were empirical project reports; 2 (2/63, 3%) used empirical and theoretical methods, and 20 (20/63, 32%) were theoretical. As for the purpose of the papers, 22 (22/63, 35%) were for assessment, 32 (32/63, 51%) for training, and 9 (9/63, 14%) for both assessment and training. In addition, 19 (19/63, 38%) were for children aged 3-14 years; 16 (16/63, 34%) for adolescents, youth, and adults aged 15-55 years; and 13 (13/63, 28%) for older adults aged 56-94 years. For more details and raw data about the works included, such as their targeted cognitive functions or impairments, please see both tables in Multimedia Appendix 2.

The Proposed Framework

Overview

To address the need for a framework that more effectively integrates game elements into cognitive assessment and training, we introduce a framework consisting of 7 phases: (1) preparation; (2) knowing users; (3) exploring existing tools for assessing or training the targeted cognitive context and determining the suitability of game-up and mapping techniques; (4) ideation; (5) prototyping using OMDE; (6) development; and (7) disseminating and monitoring. These phases are grouped into 3 overarching categories: predevelopment, development, and postdevelopment (Figure 4). Although the framework phases are presented sequentially, they are not necessarily to be conducted linearly since different ideas and directions may be explored while integrating gamification into targeted cognitive tasks. Therefore, projects are encouraged to loop back through the phases continuously [61,62].

Figure 4. The proposed framework. OMDE: Objects, Mechanics, Dynamics, Emotions.



The framework aims to facilitate creating more effective gamified cognitive tasks by using an interdisciplinary team of gamification designers, cognitive experts, and target users. Only a truly multidisciplinary team has the knowledge and expertise to address the complex factors involved in the design of gamification into cognitive tasks [61,63,64]. Gamification designers are not familiar enough with the process and execution of targeted cognitive tasks. As a result, they may incorporate game elements inappropriately [19,23-25,33,65]. Therefore, gamification designer and cognitive expert involvement is needed throughout the design process, and target users should be involved throughout phases 2, 4, 5, and 6. The involvement of target users in these phases places their needs and motivations at the center of attention. The first 3 framework phases are primarily about information gathering to develop a more well-accepted and scientific gamified task in later stages. Steps 4, 5, and 6 (ideation, prototyping using OMDE, and development, respectively) follow 2 main objectives: (1) generating gamification design ideas around targeted cognitive tasks (stages 4 and 5) and (2) developing actual gamified cognitive tasks through frequent consultation with gamification and cognitive experts and regular testing with target users (stage 6). Finally, once the efficacy of a gamified cognitive task has been demonstrated in phase 6, the task is disseminated to its target audience and then monitored periodically to maintain its effectiveness over the long term (stage 7).

Phase 1: Preparation

The primary purpose of the preparation stage is to have a good understanding of a gamification project's objectives [53,66]. Defining objectives has been recommended in most reviewed gamification frameworks (9 of 11) and will support a later stage to figure out if the desired goals have been achieved [66]. The interdisciplinary team should list all potential objectives and then rank and justify the list in terms of importance since trade-offs of less important goals for more important ones might be needed [49,53]. Finally, as the team goes through gamification design and development, it can go back to the list to focus on what is really important [49]. Therefore, the defined objectives should be achievable, specific, relevant, measurable, and time-bound [66]. For instance, an initially broad goal of "increasing participants' motivation to complete their cognitive training exercises" may be refined to "conduct 10 minutes of training each day."

According to frameworks such as [50,53,66], gamification's suitability as a possible way to intervene should be examined before starting any gamification design process. It can be carried out by detecting the problem that gamification should solve by gathering and analyzing quantitative and qualitative information. After determining the problem, the root reason that caused the problem must be motivational. Otherwise, gamification is not suitable [66]. The root reason can be identified by the "Five Whys" technique that determines the root cause of a problem by repeating the question "Why?" [67].

It is also essential to identify the standard project requirements and constraints such as scope (time, personnel, budget) and legal and ethical constraints since they can affect a gamification project's success [50,53].

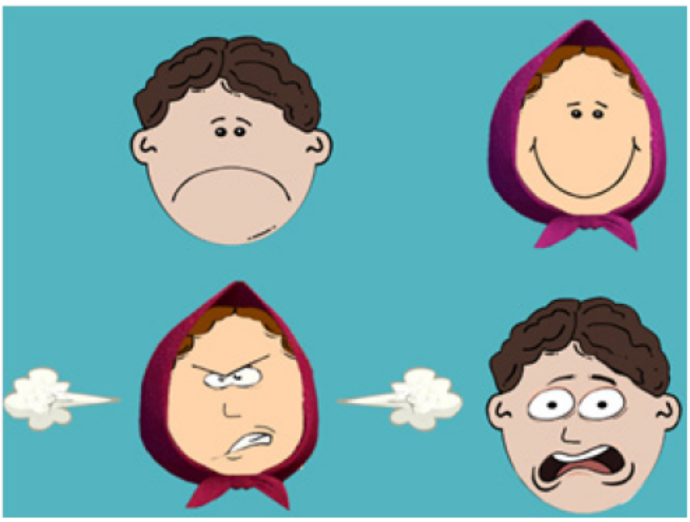
Phase 2: Knowing Users

During this stage, the interdisciplinary team must select one or more of a variety of methods to collect information about target users' motivations and needs (eg, interviews, observation of target users' behaviors, surveys, focus groups, questionnaires) [50,53]. After collecting and analyzing users' data, users with similar characteristics should be segmented into groups to create user personas. The segmentation helps the team choose a more acceptable design to gamify targeted cognitive tasks.

Typically, gamification through motivational affordances enriches information systems [53]. Therefore, it is essential to conduct this phase (9 of 11 frameworks have had one step for understanding users). People are motivated by different motivational affordances based on characteristics such as their age, gender, and culture. The Octalysis gamification framework is widely used to segment users based on their motivations [68]. Octalysis was developed by Chou [68] as an octagon with 8 core drivers of individuals on each side: (1) epic meaning and calling, (2) development and accomplishment, (3) empowerment of creativity and feedback, (4) ownership and possession, (5) social influence and relatedness, (6) scarcity and impatience, (7) unpredictability and curiosity and, (8) loss and avoidance. The game strategies or elements that are associated with each driver have been grouped next to it.

In cognitive contexts, in addition to considering the users' motivations, their needs should be identified since they may suffer from mild to severe cognitive dysfunction, which may sometimes be accompanied by physical disabilities. Afrasiabi Navan and Khaleghi [7] developed the game "Smile 1" to help Iranian autistic children recognize emotional states such as happiness, sadness, anger, and fear in the cartoon faces of girls and boys that appear in the game. The girls have a scarf (Figure 5) since these children only identify women and girls who wear a scarf (in Iranian culture, girls and women wear a scarf). To identify users' needs, most gamified assessments and training have tried to implement a gamified experience that allows users natural and straightforward interactions, often using touch-based technologies such as smartphones and tablets (22 of 63). For more information about designing a user-friendly interface for people facing cognitive and physical disabilities, please see [2,69-73].

Figure 5. Screenshot of the game "Smile 1," which was developed to help autistic children recognize different emotions [7].



Phase 3: Exploring Existing Tools for Assessing or Training the Targeted Cognitive Context and Determining the Suitability of Game-Up and Mapping Techniques

At this stage, the interdisciplinary team should thoroughly acquaint itself with existing tools for assessing or training the targeted cognitive functions or impairments through methods such as consulting with cognitive experts and gathering quantitative and qualitative information. This work helps the team incorporate game elements into these tools without changing their process and execution and find integration points for adding game elements [74]. Describing the tools at a granular level is required. Otherwise, it is not guaranteed that the next framework steps will lead to the desired outcomes [53,74]. Three main tools that can be explored for cognitive assessment and training purposes are standard computerized or not computerized cognitive tasks, existing cognitive games, and existing video games.

For standard computerized or not computerized cognitive tasks, there may be more than one cognitive task designed to assess or train a cognitive function (eg, continuous performance test, Go/No-Go test, stop signal task for assessing attention, inhibitory, and motor skills). Selecting appropriate cognitive tasks is very important. Some tasks may have better performance than others [24,75]. Valladares-Rodriguez et al [75] chose the California Learning Verbal Test II to assess episodic memory. The task, in comparison to other tasks such as the Children's Memory Scale, Rey Auditory Verbal Learning Test, and Wechsler Memory Scale, has a large number of variables and produces more qualitative information. Computer versions have now been made for many standard cognitive tasks, which are cheaper, more repeatable, and easier to administer and distribute [76]. Many of these tasks can be found in [77-79].

Regarding existing cognitive games, many games have been developed based on standard cognitive tasks [80]. It is beneficial to find these games since they can be reused for current cognitive assessment and training purposes, or the initial inspiration for gamifying current tasks can be obtained by reviewing the style of these games for integrating game elements into a cognitive task [25,33,43,75]. Brain games can be found from platforms such as Cogmed [13], Nintendo Brain Age [14], Lumosity [15], and Posit Science BrainHQ [16].

Regarding existing video games, it has also been demonstrated that classical video games of different genres that have not been inherently designed to assess or train cognitive functions can be reused as a standard cognitive task. Video game challenges come in various forms, and players have to use their underlying neural systems and cognitive abilities to win these games [81-90]. Each cognitive function is typically characterized by a set of parameters estimated from a gameplay to reuse for assessing and training. In other words, the team must identify which cognitive skills are central to each gameplay [91]. For winning games like Tetris and Candy Crush, mental rotation and spatial reasoning skills are required [91]. Card games like Solitaire and FreeCell have a reasonable correlation with classical measurements of executive functions and planning abilities [92,93]. The team can explore existing games from platforms such as the App Store and Google Play. They rank games based on their rate of downloads and players' comments, which help select games according to the target user's preferences. According to Green and Bavelier [83], Doherty et al [91], and expert experiences, it is unnecessary to find appropriate games through earlier methods for categorizing games such as genre-based methods since they are no longer effective. Games that have never overlapped in terms of content and mechanics now have many points of overlap [83].

After collecting the tools, the interdisciplinary team should determine whether game-up and mapping techniques can be used instead of designing a new gamified cognitive task from scratch. Game-up refers to adding game elements such as colors, animations, sound effects, and a backstory into standard cognitive tasks without changing their fundamental properties such as stimuli, design, and procedure [5,26,94] (21 of 63 studies used the game-up technique). The developed gamified cognitive tasks based on game-up are often presented in the form of a battery of mini-games. Each mini-game focuses on a specific cognitive function (eg, [64,76,95]). For example, Zeng et al [76] gamified a computerized cognitive test battery to detect impairments in 5 cognitive functions involved in developing a major neurocognitive disorder. For each test, some mini-games were designed in the context of ADLs such as cooking, cleaning, and shopping. The main feature of game-up is simplicity in terms of its mechanics and design [7,65,96,97]. This feature is useful for individuals who suffer from cognitive impairments like children with learning disabilities and ADHD who have weak working memory capacity [65,96]. Therefore, gamified cognitive tasks for these children should be broken into short and discrete tasks [65,96]. Mapping refers to reusing an existing game (cognitive and classical games) as a cognitive task and can save considerable time and effort that have been applied in the design of explored games [98,99]. The mapping technique was used by 21 of the 63 studies (eg, [72,92,93,100-104]). Explored games should be adjusted appropriately since they usually do not provide cognitive psychologists with sufficient quantitative data about the participant's performance and progression on one hand. On the other hand, the used game elements and storylines may not align with participants' preferences [103,105,106]. Moreover, they may impose an additional cognitive load. Therefore, the exact cognitive demands of selected games should be identified by analyzing their structural characteristics [83]. Each game's structural characteristics should be examined individually since different games, even those that fall into one category such as action or first-shooting person, may require greatly different cognitive demands [81,85,107,108]. In mapping, it is also possible to mash up various games for cognitive assessment and training purposes [81]. For example, dyslexia is associated with a variety of underlying deficits in phonological, auditory, motor, memory, and visual attentional processes. According to previous findings showing the core deficit in dyslexia is related to attentional problems, Franceschini et al [81] explored 10 action games to train dyslexic children. Action games can enhance a wide variety of visual attentional abilities, such as segmenting items both in time and across space.

Phase 4: Ideation

The interdisciplinary team is involved in a highly iterative design process through the next 3 phases (ideation, prototyping using OMDE, and development). Iterative processes enable the team to obtain more creative and effective gamification designs. Of the 11 frameworks, 5 have one or more steps that should be iterated until the desired designs and outcomes are reached (ie, [46,48-50,53]). The steps that are often performed iteratively are ideation, prototyping, and development [53].

At this stage, the team combines the analyses and materials obtained in new ways to produce apt and innovative ideas to engage target users. It is necessary to involve a cross-functional group of people from cognitive experts, gamification designers, and target users to start this stage [61,62]. This work helps the interdisciplinary team to collect a greater number of more varied and creative ideas [61,62]. The participants should be encouraged to use different ways to be creative [61,62]. Brainstorming, co-creation workshops, and mind mapping are some methods [53,61,62]. The important question at this stage is how to help participants find the ideas. One solution is to explore existing games, gamification designs, and examples that may be a perfect fit for the current project [62]. For instance, in the game "Whack a Mole," moles hide quickly, and the player trying to hit them with a hammer has to be faster. Various types of moles exist in different game versions, such as ninja, pirate, samurai, and batman moles [109]. Based on studies reusing the Whack a Mole to measure attention, inhibitory control, and executive functions [72,102,103], one idea is to use ninja and samurai moles as metaphors for Go and No/Go stimuli, respectively. As a result, a gamified cognitive task that mimics the Go/No-Go design can be created using the Whack-a-Mole style. Exploring many games and gamification examples and then mashing them up to fit the current problem is another right approach for generating ideas to gamify the current task [62].

After preparing ideas, similar ideas should be clustered using affinity diagrams; then, the clusters should be prioritized using methods such as dot voting. This work helps the team to focus on important ideas in the next 2 phases [50,53].

Phase 5: Prototyping Using OMDE

After collecting the right ideas, the interdisciplinary team needs to start prototyping. Prototyping is the stage in which the team implements the ideas into tangible forms to see how they actually work. In this stage, low-fidelity (ie, "quick and dirty") prototypes are developed rapidly to gather feedback from relevant experts and target users early and often [61]. Prototyping saves time and resources by helping the team to identify refinements required before solidifying a design [61,62].

During each iteration of the prototyping, the team can use the OMDE design guideline to (1) check the motivational characteristics of prototypes (such as fun, flow, engagement, positive emotions) and (2) validate prototypes from cognitive psychology aspects. OMDE divides the components of a gamified cognitive task into 4 categories: objects, mechanics, dynamics, and emotions. Objects are a gamified cognitive task's assets, such as visual assets, images, audios, videos, and animations [60]. Mechanics refer to a gamified task's components at the level of game rules, algorithms, and data representation [60]. Dynamics are run-time users' behaviors that emerge as users partake in the gamified task such as competition and cooperation [60]. Emotions refer to whatever emotions users experience while interacting with the gamified task [44]. Participants may experience different emotions such as fear, happiness, anger, sadness, and pride while interacting with the gamified task [44]. Dynamics and emotions emerge from the selected objects and mechanics [44,52,60]. For instance, a leaderboard mechanic leads to dynamics such as competition and comparison and emotions such as fear and happiness. Some participants may be afraid of being judged by others, and the use of the leaderboard may demotivate them from continuing the gamified task. Or, many participants may enjoy these dynamics, and the leaderboard can motivate them. Therefore, displaying participants' statuses in the leaderboard must be an optional feature in a gamified task.

Good dynamics and emotions are vital to ensuring a strong user commitment to participation [44,52,60]. To check gamified tasks' motivational features with OMDE, the interdisciplinary team must first define the desired dynamics and emotional responses that the designed task should evoke among users. Then, in each iteration, the team must list what dynamics and emotions emerge from the gamified task in practice and then compare the responses with the desired ones to determine if the desired responses have been reached. The team cannot accurately predict what dynamics and emotions will emerge from a gamified task. Therefore, it is necessary to use OMDE iteratively [44,52].

In gamified cognitive tasks, it is also essential to validate the components of OMDE from cognitive aspects because they may impose an additional cognitive load. Objects and mechanics can cause difficulty in categorizing cognitive tasks' stimuli for participants or can evoke emotions such as anxiety and stress that may distract participants' attention from completing gamified tasks [19,23-25,33,65]. In this circumstance, participant errors increase, and the reliability of the data obtained decreases. For instance, in the study by Birk et al [23], the gamified Go/No-Go task decreased users' performance. In the standard task, a sequence of stimuli is presented for 500 ms. Participants should respond to circles but not to squares. In the gamified version, participants should shoot blond zombies (Go stimulus) but not yellow hat moles (No/ Go stimulus). In the standard task, a circle is very different from a square. In contrast, in the gamified task, the colors of yellow hat moles and blond zombies are close to each other and can cause difficulty while participants are trying to identify the gamified task's stimuli. In order to gamify the Go/No-Go task, Lumsden et al [19] suggested that red and green colors be used instead of cartoon characters because participants are more familiar with colors. The components of OMDE can be validated by discussions with cognitive experts and answering questions such as: (1) Is it possible to gamify the cognitive tasks' stimuli? If yes, how can we do so? (2) Does the team have the freedom to choose objects for gamifying cognitive tasks, or should they be selected among those listed in a specified set or ones that participants are more familiar with, such as everyday objects? (3) What degree of structural similarities (such as shape, size, and color) between objects and mechanics should be adjusted? (4) Is it possible to gamify the surrounding environment of the selected cognitive tasks' stimuli? If yes, how should the degree of separation between cognitive and gamified sections be adjusted? (5) Does the designed gamified cognitive task lead to negative emotions like anxiety and stress?

Phase 6: Development

During this stage, actual gamified cognitive tasks are developed through frequent consultation with relevant experts (gamification and cognitive experts) and regular testing with target users (Figure 6). Based on the examined gamification efforts in cognitive tasks, to test the efficacy of gamified tasks, rigorous evaluations are required in terms of user engagement and data quality (eg, [9,19-21,23-27,34,38-40]).

Figure 6. Overall structure of the development phase.

Two methods are widely used to evaluate how gamified cognitive tasks influence users' engagement and motivation [24,27,38,39]: (1) subjective measures of engagement, in which the motivation level of a gamified task is measured through self-report questionnaires based on SDT [31] and Flow [32] theory (eg, Flow State Scale [110]) and (2) objective measures of engagement such as the number of times participants used a gamified task or the number of optional cognitive assessment or training sessions performed by participants, methods that might be more preferred by the interdisciplinary team [24,38,39,63,111]. A combination of both methods is often used to measure gamified tasks' motivation levels [24].

To assess how gamified cognitive tasks impact the quality of data and to indicate the maturity of these tasks to be used as a valid clinical tool, they must be evaluated for 2 essential properties [43]: (1) reliability, which refers to the extent to which a task's results are consistent and repeatable, and there are 4 types of reliability (test-retest reliability, parallel forms reliability, internal consistency reliability, and interrater reliability) and (2) validity, which refers to how well a task measures what it claims and includes criterion validity, content validity, construct validity, face validity, external validity, and ecological validity.

In cognitive training, it may also be necessary to measure to what extent gamified tasks can impact transfer effects [38,41,83]. The term "transfer" is frequently used in clinical practice and refers to the extent to which considered cognitive training tasks can improve untrained cognitive abilities. New tasks and situations are included to measure transfer effects. Transfer effects are divided into near and far transfer effects. Cognitive training has near transfer effects if it improves cognitive skills that are highly similar to trained cognitive skills. Far transfer effects refer to improvements in cognitive skills that are less similar to trained skills.

There are 2 other essential factors for conducting rigorous evolutions: (1) selecting sufficient sample sizes and (2) selecting an appropriate duration for evaluation. Most gamification efforts in cognitive contexts have used small sample sizes to evaluate the efficacy of gamified cognitive tasks [9]. Also, little consideration has been given to using statistical analyses such as power analysis for a sample size calculation [9,112]. For more information about how to calculate sufficient sample sizes, please see [113,114]. Insufficient sample sizes limit the reliability and generalizability of the results [9,115,116]. Moreover, only a few studies have used randomized controlled trials (RCTs) to evaluate gamified cognitive tasks [30]. In clinical research, RCTs are considered the most robust study design for evaluating the effectiveness of a new tool due to the ability of RCTs to minimize several forms of bias [61]. RCTs randomly assign participants to an experimental group and a control group. The use of an RCT design comparing gamified (experimental group) and nongamified (control group) versions of the same cognitive task has been highly recommended to evaluate the potential efficacy of gamified tasks [9,30]. Regarding selecting the appropriate duration for evaluation, participants are not involved in the gamified task over the long term but instead participate for a short time. In turn, it remains unclear after how long participants feel boredom and how the quality of data will alter in these circumstances [19,25].

Moreover, a short duration can cause errors due to participants' unfamiliarity with the gamified task. In this regard, using short tutorials and warm-up sessions before actual evaluation sessions has been recommended [65,76].

Phase 7: Disseminating and Monitoring

Once the gamified cognitive task's efficacy has been demonstrated in the previous step, the task is finally disseminated to its target audience. There have long been calls for disseminating and sharing well-evaluated digital health interventions due to the abundance of low-quality interventions currently available to the public [17,61]. Disseminating gamified tasks provides access to the broader population that may benefit the most from these tasks and helps the industry invest in these interventions more quickly [61,73]. Disseminating can occur via the App Store or Google Play. Also, industry partnerships can support a more effective and sustainable dissemination of gamified cognitive tasks [61,117]. It is also highly recommended that projects disseminate their findings, experiences, and methods for developing gamified cognitive tasks to scientific journals, conferences, researchers, and digital mental health intervention developers. It can advance future gamified tasks and improve their effectiveness [61,117]. Disseminating can also include open sharing of gamified cognitive task codes via GitHub [118] or allowing free access to a mobile health platform such as Mobile Sensor Data-to-Knowledge (MD2K) [119].

For disseminating gamified tasks, 2 other important factors should be considered by the team. First, according to expert experiences and [91] in collaboration with cognitive experts, appropriate guidelines and prescriptions should be prepared for using gamified cognitive tasks by clinics and target users (eg, determining the minimum effort and time that target users should spend to improve their cognitive skills). These instructions help mental experts and target users to use gamified tasks effectively. Second, ethical concerns are another primary concern that should be considered in this stage since performance variables are embedded in gamified tasks to track improvements in users' cognitive functions. Therefore, gathering target users' information should be undertaken with their explicit consent [23,47,50,53,73].

After disseminating, gamified tasks' performance should also be accompanied by postlaunch monitoring and evaluation, in which usage and applicability of the gamified task through methods introduced in the previous phases are observed in regular intervals to compile a list of possible improvements [47,53]. For instance, different levels of cognitive impairments may exist among users, and after a cognitive training program through the gamified task, users' cognitive status may improve. Therefore, according to the users' new levels, some changes may need to be applied in the gamified task [69,73]. Also, gamified tasks' motivation levels should be continuously monitored to maintain intrinsic motivation for the long term. The initial effects of game elements can diminish over time [47,53].

Discussion

Overview

This paper proposes a 7-step framework to guide the design, development, and evaluation of gamified cognitive tasks designed to assess or train cognition. Within these steps, there are a series of key recommendations on how each step should be operationalized. Along with the framework, the article presents the OMDE guideline at stage 5 of the framework (prototyping using OMDE), which contains vital recommendations for advancing the understanding of design complexities when applying gamification in cognitive tasks.

The prototypes of the framework were designed and evaluated extensively by evidence from 3 sources: (1) existing gamification design frameworks, (2) project reports of applying game elements into cognitive assessment and training, and (3) expert experiences. To our knowledge, this is the first study of its kind that has converged these sources to propose a unified model to design gamification in cognitive tasks. The significant point about gamification efforts in cognitive tasks is that they do not use a specific design process to incorporate game elements. We used an alternative solution, such that by observing each work, we tried to extract clues or pieces of information to propose an abstract process for each of them, if possible. This work helped us to identify critical factors and considerations for gamifying cognitive tasks.

Comparison of the Proposed Framework With General Gamification Frameworks

Like general gamification frameworks, the proposed framework follows a user-centered design to improve users' participation but has many added features that make it appropriate to gamify cognitive tasks. One main finding of this study that gamification designers need to be aware of is that an intermediate design is required for gamifying cognitive tasks, meaning that game elements cannot be selected without considering the targeted cognitive context characteristics. Otherwise, an irrelevant cognitive load may jeopardize data quality. Therefore, collaboration between both gamification and mental experts is required to examine the interplay of game elements and cognitive processes [9]. The intermediate design is a crucial feature that differentiates the gamification design framework required for cognitive assessment and training from other contexts. Based on examined gamification efforts in cognitive tasks, we concluded that 2 techniques have been widely used to gamify cognitive tasks besides designing a new gamified task from scratch: (1) gaming-up an existing cognitive task and (2) mapping an existing game (cognitive and classical games) to a cognitive function or impairment to assess or train it. The required details to use game-up and mapping techniques were provided in step 3 of the framework as far as possible.

Limitations

The main limitation was that only a few studies have discussed how game elements impact participants when interacting with the gamified task and how they should be utilized in cognitive tasks to positively influence data quality and user engagement. Gathering evidence from the mentioned sources only gave an initial evaluation for the proposed framework because the number of experts and relevant studies was limited. Hence, a more robust evaluation is necessary. For this purpose, we listed most of the recognized experts in the area to evaluate and refine the framework in subsequent studies.

Possible Future Studies

By further developing the framework exploited in this work and utilizing machine learning and deep learning algorithms, it is possible to create a recommender system that can suggest the most appropriate game elements according to characteristics of the targeted cognitive context and users' preferences or requirements. In our work, only existing games that have been developed or examined in scientific papers were investigated. It is also possible to analyze a great number of current games, from brain games to classical games, for further development of the framework. Due to the time constraints, establishing the feasibility of proposing different gamification design frameworks for cognitive functions that inherently share similar cognitive processes like processing speed (Gs) and working memory [20] was not provided.

Conclusions

While more work is needed to further refine and evaluate the framework, we believe our framework has great potential to be used as a foundation for developing effective gamified cognitive tasks. Furthermore, ideas presented in the paper can be further developed and researched by many other researchers and practitioners.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Explored gamification design frameworks.
[DOC File , 100 KB - games_v9i2e21900_app1.doc]

Multimedia Appendix 2
Included papers applied gamification in cognitive training/ testing.
[DOC File , 231 KB - games_v9i2e21900_app2.doc]

References

1. Global cognitive assessment and training market 2018-2022. Business Wire. 2018 Feb 08. URL: https://www.businesswire.com/news/home/20180208005680/en/Global-Cognitive-Assessment-Training-Market-2018-2022-- [accessed 2019-04-05]
2. Gamberini L, Martino F, Seraglia B, Spagnolli A, Fabregat M, Ibanez F, et al. Eldergames project: An innovative mixed reality table-top solution to preserve cognitive functions in elderly people. 2009 Presented at: 2nd Conference on Human System Interactions; May 21-23, 2009; Catania, Italy. [doi: 10.1109/hsi.2009.5090973]
3. Chan JY, Kwong JS, Wong A, Kwok TC, Tsoi KK. Comparison of Computerized and Paper-and-Pencil Memory Tests in Detection of Mild Cognitive Impairment and Dementia: A Systematic Review and Meta-analysis of Diagnostic Studies. J Am Med Dir Assoc 2018 Sep;19(9):748-756.e5. [doi: 10.1016/j.jamda.2018.05.010] [Medline: 29921507]
4. Potter G, Steffens D. Contribution of depression to cognitive impairment and dementia in older adults. Neurologist 2007 May;13(3):105-117. [doi: 10.1097/01.nrl.0000252947.15389.a9] [Medline: 17495754]
5. Boendermaker WJ, Prins PJ, Wiers RW. Cognitive Bias Modification for adolescents with substance use problems--Can serious games help? J Behav Ther Exp Psychiatry 2015 Dec;49(Pt A):13-20 [FREE Full text] [doi: 10.1016/j.jbtep.2015.03.008] [Medline: 25843611]
6. Mayes R, Bagwell C, Erkulwater J. ADHD and the rise in stimulant use among children. Harv Rev Psychiatry 2008;16(3):151-166. [doi: 10.1080/10673220802167782] [Medline: 18569037]
7. Afrasiabi Navan A, Khaleghi A. Using Gamification to Improve the Quality of Educating Children with Autism. Rev. Cient 2020 Jan 01;37(1):90-106. [doi: 10.14483/23448350.15431]
8. Vourvopoulos A, Faria A, Ponnam K, Bermudez BS. RehabCity: designvalidation of a cognitive assessment and rehabilitation tool through gamified simulations of activities of daily living. 2014 Presented at: ACE '14: 11th Conference on Advances in Computer Entertainment Technology; November 2014; Funchal, Portugal. [doi: 10.1145/2663806.2663852]
9. 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]
10. 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]
11. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med 1990 Dec 15;113(12):941-948. [doi: 10.7326/0003-4819-113-12-941] [Medline: 2240918]

12. Nasreddine Z, Phillips N, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
13. CogMed. URL: https://www.cogmed.com/ [accessed 2020-09-21]
14. Nintendo Brain Age. URL: https://www.nintendo.com/games/detail/brain-age-train-your-brain-in-minutes-a-day-wii-u/ [accessed 2021-04-17]
15. Lumosity. URL: https://www.lumosity.com [accessed 2020-09-21]
16. BrainHQ. Posit Science. URL: https://www.brainhq.com/ [accessed 2020-09-21]
17. Gray SI, Robertson J, Manches A, Rajendran G. BrainQuest: The use of motivational design theories to create a cognitive training game supporting hot executive function. International Journal of Human-Computer Studies 2019 Jul;127:124-149. [doi: 10.1016/j.ijhcs.2018.08.004]
18. Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 1935;18(6):643-662. [doi: 10.1037/h0054651]
19. Lumsden J, Skinner A, Woods A, Lawrence N, Munafò M. The effects of gamelike features and test location on cognitive test performance and participant enjoyment. PeerJ 2016;4:e2184 [FREE Full text] [doi: 10.7717/peerj.2184] [Medline: 27441120]
20. McPherson J, Burns NR. Assessing the validity of computer-game-like tests of processing speed and working memory. Behav Res Methods 2008 Nov;40(4):969-981. [doi: 10.3758/BRM.40.4.969] [Medline: 19001388]
21. McPherson J, Burns NR. Gs invaders: assessing a computer game-like test of processing speed. Behav Res Methods 2007 Nov;39(4):876-883. [doi: 10.3758/bf03192982] [Medline: 18183904]
22. Ninaus M, Pereira G, Stefitz R, Prada R, Paiva A, Neuper C, et al. Game elements improve performance in a working memory training task. IJSG 2015 Feb 10;2(1):1. [doi: 10.17083/ijsg.v2i1.60]
23. Birk MV, Mandryk R, Bowey J, Buttlar B. The effects of adding premise and backstory to psychological tasks. 2015 Presented at: CHI'15 Workshop Researching Gamification: Strategies, Opportunities, Challenges, Ethics; April 19, 2015; Seoul, Korea.
24. 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]
25. 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]
26. Hawkins GE, Rae B, Nesbitt KV, Brown SD. Gamelike features might not improve data. Behav Res Methods 2013 Jun 9;45(2):301-318. [doi: 10.3758/s13428-012-0264-3] [Medline: 23055169]
27. Miranda AT, Palmer EM. Intrinsic motivation and attentional capture from gamelike features in a visual search task. Behav Res Methods 2014 Mar 9;46(1):159-172. [doi: 10.3758/s13428-013-0357-7] [Medline: 23835649]
28. Deterding S, Dixon D, Khaled R, Nacke L. From Game Design Elements to Gamefulness: Defining "Gamification". 2011 Presented at: MindTrek '11: 15th International Academic MindTrek Conference: Envisioning Future Media Environments; September 28-30, 2011; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]
29. Ryan RM, Deci EL. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]
30. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DML. 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]
31. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
32. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. New York, NY: Harper Collins; 1990.
33. Craven MP, Groom MJ. Computer Games for User Engagement in Attention Deficit Hyperactivity Disorder (ADHD) Monitoring and Therapy. 2015 Presented at: International Conference on Interactive Technologies and Games; October 22-23, 2015; Nottingham, UK p. 22-23. [doi: 10.1109/itag.2015.9]
34. Boendermaker WJ, Boffo M, Wiers RW. Exploring Elements of Fun to Motivate Youth to Do Cognitive Bias Modification. Games Health J 2015 Dec;4(6):434-443. [doi: 10.1089/g4h.2015.0053] [Medline: 26421349]
35. Van de Weijer-Bergsma E, Kroesbergen EH, Prast EJ, Van Luit JEH. Validity and reliability of an online visual-spatial working memory task for self-reliant administration in school-aged children. Behav Res Methods 2015 Sep 26;47(3):708-719. [doi: 10.3758/s13428-014-0469-8] [Medline: 24771322]
36. Openclipart. URL: https://openclipart.org/ [accessed 2020-09-25]
37. 4vector. URL: https://4vector.com/ [accessed 2020-09-25]
38. Dörrenbächer S, Müller PM, Tröger J, Kray J. Dissociable effects of game elements on motivation and cognition in a task-switching training in middle childhood. Front Psychol 2014 Nov 13;5:1275 [FREE Full text] [doi: 10.3389/fpsyg.2014.01275] [Medline: 25431564]

39. Prins PJ, Dovis S, Ponsioen A, ten Brink E, van der Oord S. Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? Cyberpsychol Behav Soc Netw 2011 Mar;14(3):115-122. [doi: 10.1089/cyber.2009.0206] [Medline: 20649448]
40. Prins PJ, Brink ET, Dovis S, Ponsioen A, Geurts HM, de Vries M, et al. "Braingame Brian": Toward an Executive Function Training Program with Game Elements for Children with ADHD and Cognitive Control Problems. Games Health J 2013 Feb;2(1):44-49. [doi: 10.1089/g4h.2013.0004] [Medline: 26196554]
41. Baniqued PL, Lee H, Voss MW, Basak C, Cosman JD, Desouza S, et al. Selling points: What cognitive abilities are tapped by casual video games? Acta Psychol (Amst) 2013 Jan;142(1):74-86 [FREE Full text] [doi: 10.1016/j.actpsy.2012.11.009] [Medline: 23246789]
42. Brown HR, Zeidman P, Smittenaar P, Adams RA, McNab F, Rutledge RB, et al. Crowdsourcing for cognitive science--the utility of smartphones. PLoS One 2014 Jul 15;9(7):e100662 [FREE Full text] [doi: 10.1371/journal.pone.0100662] [Medline: 25025865]
43. Valladares-Rodríguez S, Pérez-Rodríguez R, Anido-Rifón L, Fernández-Iglesias M. Trends on the application of serious games to neuropsychological evaluation: A scoping review. J Biomed Inform 2016 Dec;64:296-319 [FREE Full text] [doi: 10.1016/j.jbi.2016.10.019] [Medline: 27815228]
44. Robson K, Plangger K, Kietzmann JH, McCarthy I, Pitt L. Is it all a game? Understanding the principles of gamification. Business Horizons 2015 Jul;58(4):411-420. [doi: 10.1016/j.bushor.2015.03.006]
45. Aparicio AF, Vela FLG, Sánchez JLG, Montes JLI. Analysis and application of gamification. 2012 Presented at: 13th International Conference on Interacción Persona-Ordenador; October 2012; Elche, Spain p. 03-05. [doi: 10.1145/2379636.2379653]
46. Brito J, Vieira V, Duran A. Towards a Framework for Gamification Design on Crowdsourcing Systems: The G.A.M.E. Approach. 2015 Presented at: 12th International Conference on Information Technology - New Generations; April 13-15, 2015; Las Vegas, NV. [doi: 10.1109/itng.2015.78]
47. Kumar J. Gamification at work: Designing engaging business software. In: International Conference of Design, User Experience, and Usability. Berlin, Germany: Springer Publishing Company; 2012:528-537.
48. Marache-Francisco C, Brangier E. Process of Gamification From The Consideration of Gamification To Its Practical Implementation. 2013 Presented at: CENTRIC 2013: The Sixth International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services; October 2013; Venice, Italy p. 527-537.
49. Werbach K, Hunter D. For The Win: How Game Thinking Can Revolutionize Your Business. Pennsylvania, PA: Wharton Digital Press; 2012.
50. Deterding S. The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. Human–Computer Interaction 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471]
51. Li X. A Method to Support Gamification Design Practice with Motivation Analysis and Goal Modeling. 2018 Presented at: 2nd International GamiFIN Conference; 2018; Pori, Finland p. 151-158.
52. Liu D, Santhanam R, Webster J. Toward Meaningful Engagement: A Framework for Design and Research of Gamified Information Systems. MISQ 2017 Apr 4;41(4):1011-1034. [doi: 10.25300/misq/2017/41.4.01]
53. Morschheuser B, Hassan L, Werder K, Hamari J. How to design gamification? A method for engineering gamified software. Information and Software Technology 2018 Mar;95:219-237. [doi: 10.1016/j.infsof.2017.10.015]
54. Mora A, Riera D, González C, Arnedo-Moreno J. Gamification: a systematic review of design frameworks. J Comput High Educ 2017 May 23;29(3):516-548. [doi: 10.1007/s12528-017-9150-4]
55. Azouz O, Lefdaoui Y. Gamification design frameworks: a systematic mapping study. 2018 Presented at: 6th International Conference on Multimedia Computing and Systems (ICMCS); May 10-12, 2018; Rabat, Morocco. [doi: 10.1109/icmcs.2018.8525900]
56. Hansen OG. What gamification design do users want in a self-management application for chronic diseases? - The case of Cystic Fibrosis. Norwegian University of Science and Technology. 2017. URL: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2458481 [accessed 2021-04-17]
57. Johannesson P, Perjons E. An Introduction to Design Science. New York, NY: Springer Publishing Company; 2014.
58. Hevner, March, Park, Ram. Design Science in Information Systems Research. MIS Quarterly 2004;28(1):75. [doi: 10.2307/25148625]
59. Brinkkemper S. Method engineering: engineering of information systems development methods and tools. Information and Software Technology 1996 Jan;38(4):275-280. [doi: 10.1016/0950-5849(95)01059-9]
60. Hunicke R, LeBlanc M, Zubek R. MDA: A formal approach to game designgame research. 2004 Presented at: AAAI Workshop on Challenges in Game AI; January 2004; San Jose, CA.
61. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): A Framework and Toolkit of Strategies for the Development of More Effective Digital Interventions to Change Health Behavior. J Med Internet Res 2016 Dec 16;18(12):e317 [FREE Full text] [doi: 10.2196/jmir.5927] [Medline: 27986647]
62. Herger M. Enterprise Gamification: Engaging people by letting them have fun. Leipzig: CreateSpace Independent Publishing Platform; 2014.

63. Manera V, Petit P, Derreumaux A, Orvieto I, Romagnoli M, Lyttle G, et al. 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Front Aging Neurosci 2015 Mar 17;7:24 [FREE Full text] [doi: 10.3389/fnagi.2015.00024] [Medline: 25852542]
64. Delgado MT, Uribe PA, Alonso AA, Díaz RR. TENI: A comprehensive battery for cognitive assessment based on games and technology. Child Neuropsychol 2016 Nov 14;22(3):276-291. [doi: 10.1080/09297049.2014.977241] [Medline: 25396766]
65. Shaban A, Pearson E. A Learning Design Framework to Support Children with Learning Disabilities Incorporating Gamification Techniques. 2019 Presented at: CHI EA '19: 2019 CHI Conference on Human Factors in Computing Systems; May 2019; Glasgow, Scotland. [doi: 10.1145/3290607.3312806]
66. Carreño AM. A framework for agile design of personalized gamification services. Universitat Oberta de Catalunya. 2018 Jun. URL: https://www.tesisenred.net/bitstream/handle/10803/664190/Thesis_June.pdf [accessed 2021-04-17]
67. Serrat O. The five whys technique. In: Knowledge Solutions. Singapore: Springer Publishing Company; May 23, 2017:307-310.
68. Chou Y. Octalysis: complete gamification framework. Gamification & Behavioral Design. URL: http://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/ [accessed 2020-09-17]
69. Söderberg S. A Case Study on the Creative Process of Making a Game for Cognitively Impaired Children. Mälardalens University School of Innovation, Design and Engineering. 2011 Oct 2. URL: https://www.diva-portal.org/smash/get/diva2:445565/FULLTEXT01.pdf [accessed 2021-04-17]
70. Guidelines for the development of entertaining software for people with multiple learning disabilities. MediaLT. 2004. URL: http://www.medialt.no/rapport/entertainment_guidelines/ [accessed 2019-07-25]
71. Jaramillo-Alcázar A, Luján-Mora S, Salvador-Ullauri L. Inclusive Education: Mobile Serious Games for People with Cognitive Disabilities. Enfoque UTE 2018 Mar 30;9(1):53-66. [doi: 10.29019/enfoqueute.v9n1.236]
72. Tong T, Chignell M, Tierney MC, Lee J. A Serious Game for Clinical Assessment of Cognitive Status: Validation Study. JMIR Serious Games 2016 May 27;4(1):e7 [FREE Full text] [doi: 10.2196/games.5006] [Medline: 27234145]
73. Robert PH, König A, Amieva H, Andrieu S, Bremond F, Bullock R, et al. Recommendations for the use of Serious Games in people with Alzheimer's Disease, related disorders and frailty. Front Aging Neurosci 2014 Mar 24;6:54 [FREE Full text] [doi: 10.3389/fnagi.2014.00054] [Medline: 24715864]
74. Klevers M, Sailer M, Günthner W. Implementation model for the gamification of business processes: a study from the field of material handling. In: Kaneda T, Kanegae H, Toyoda Y, Rizzo P, editors. Simulation and gaming in the network society. New York, NY: Springer Publishing Company; 2016:173-184.
75. Valladares-Rodriguez S, Perez-Rodriguez R, Facal D, Fernandez-Iglesias M, Anido-Rifon L, Mouriño-Garcia M. Design process and preliminary psychometric study of a video game to detect cognitive impairment in senior adults. PeerJ 2017;5:e3508 [FREE Full text] [doi: 10.7717/peerj.3508] [Medline: 28674661]
76. Zeng Z, Fauvel S, Hsiang BTT, Wang D, Qiu Y, Khuan PCO, et al. Towards long-term tracking and detection of early dementia: a computerized cognitive test battery with gamification. 2018 Presented at: ICCSE'18: 3rd International Conference on Crowd Science and Engineering; July 2018; Singapore. [doi: 10.1145/3265689.3265719]
77. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. Am J Alzheimers Dis Other Demen 2015 Feb 13;30(1):13-28 [FREE Full text] [doi: 10.1177/1533317514522852] [Medline: 24526761]
78. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov 01;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
79. Tierney MC, Lermer MA. Computerized cognitive assessment in primary care to identify patients with suspected cognitive impairment. J Alzheimers Dis 2010;20(3):823-832. [doi: 10.3233/JAD-2010-091672] [Medline: 20413868]
80. Gray SI. Developing and evaluating the feasibility of an active training game for smartphones as a tool for promoting executive function in children. The University of Edinburgh. 2017 Jun 07. URL: https://era.ed.ac.uk/handle/1842/23414 [accessed 2021-04-17]
81. Franceschini S, Gori S, Ruffino M, Viola S, Molteni M, Facoetti A. Action video games make dyslexic children read better. Curr Biol 2013 Mar 18;23(6):462-466 [FREE Full text] [doi: 10.1016/j.cub.2013.01.044] [Medline: 23453956]
82. 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]
83. Green CS, Bavelier D. Action video game training for cognitive enhancement. Current Opinion in Behavioral Sciences 2015 Aug;4:103-108. [doi: 10.1016/j.cobeha.2015.04.012]
84. Dye MW, Green CS, Bavelier D. Increasing Speed of Processing With Action Video Games. Curr Dir Psychol Sci 2009 Dec;18(6):321-326 [FREE Full text] [doi: 10.1111/j.1467-8721.2009.01660.x] [Medline: 20485453]
85. Boot WR. Video games as tools to achieve insight into cognitive processes. Front Psychol 2015;6:3 [FREE Full text] [doi: 10.3389/fpsyg.2015.00003] [Medline: 25653634]
86. Bavelier D, Green CS. The Brain-Boosting Power of Video Games. Sci Am 2016 Jul 14;315(1):26-31. [doi: 10.1038/scientificamerican0716-26] [Medline: 27348376]
87. Anguera JA, Gazzaley A. Video games, cognitive exercises, and the enhancement of cognitive abilities. Current Opinion in Behavioral Sciences 2015 Aug;4:160-165. [doi: 10.1016/j.cobeha.2015.06.002]

88. Brilliant TD, Nouchi R, Kawashima R. Does Video Gaming Have Impacts on the Brain: Evidence from a Systematic Review. Brain Sci 2019 Sep 25;9(10):251 [FREE Full text] [doi: 10.3390/brainsci9100251] [Medline: 31557907]
89. Green C, Bavelier D. Learning, attentional control, and action video games. Curr Biol 2012 Mar 20;22(6):R197-R206 [FREE Full text] [doi: 10.1016/j.cub.2012.02.012] [Medline: 22440805]
90. Bavelier D, Green CS, Han DH, Renshaw PF, Merzenich MM, Gentile DA. Brains on video games. Nat Rev Neurosci 2011 Nov 18;12(12):763-768 [FREE Full text] [doi: 10.1038/nrn3135] [Medline: 22095065]
91. Doherty SM, Keebler JR, Davidson SS, Palmer EM, Frederick CM. Recategorization of Video Game Genres. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2018 Sep 27;62(1):2099-2103. [doi: 10.1177/1541931218621473]
92. Gielis K, Brito F, Tournoy J, Abeele VV. Can Card Games Be Used to Assess Mild Cognitive Impairment? A Study of Klondike Solitaire and Cognitive Functions. 2017 Presented at: CHI PLAY '17: Annual Symposium on Computer-Human Interaction in Play; October 2017; Amsterdam, Netherlands p. 269-276. [doi: 10.1145/3130859.3131328]
93. Jimison H, Pavel M, McKanna J, Pavel J. Unobtrusive monitoring of computer interactions to detect cognitive status in elders. IEEE Trans Inf Technol Biomed 2004 Sep;8(3):248-252. [doi: 10.1109/titb.2004.835539] [Medline: 15484429]
94. Groznik V, Sadikov A. Gamification in Cognitive Assessment and Cognitive Training for Mild Cognitive Impairment. In: Geroimenko V, editor. Augmented Reality Games II. New York, NY: Springer Publishing Company; 2019:179-204.
95. Valladares-Rodriguez S, Pérez-Rodriguez R, Fernandez-Iglesias JM, Anido-Rifón L, Facal D, Rivas-Costa C. Learning to Detect Cognitive Impairment through Digital Games and Machine Learning Techniques. Methods Inf Med 2018 Sep 24;57(04):197-207. [doi: 10.3414/me17-02-0011]
96. Peijnenborgh JC, Hurks PP, Aldenkamp AP, van der Spek ED, Rauterberg G, Vles JS, et al. A Study on the Validity of a Computer-Based Game to Assess Cognitive Processes, Reward Mechanisms, and Time Perception in Children Aged 4-8 Years. JMIR Serious Games 2016 Sep 22;4(2):e15 [FREE Full text] [doi: 10.2196/games.5997] [Medline: 27658428]
97. Cota TT, Ishitani L, Vieira N. Mobile game design for the elderly: A study with focus on the motivation to play. Computers in Human Behavior 2015 Oct;51:96-105. [doi: 10.1016/j.chb.2015.04.026]
98. Castronova E, Knowles I. Modding board games into serious games: The case of climate policy. International Journal of Serious Games 2015;2(3):41-62 [FREE Full text]
99. Vasiljevic GAM, de Miranda LC. Brain–Computer Interface Games Based on Consumer-Grade EEG Devices: A Systematic Literature Review. International Journal of Human–Computer Interaction 2019 Jun 03;36(2):105-142. [doi: 10.1080/10447318.2019.1612213]
100. Joselli M, Binder F, Clua E, Soluri E. Concept, development and evaluation of a mind action game with the electro encephalograms as an auxiliary input. JIS 2017 Sep 14;8(1):1. [doi: 10.5753/jis.2017.675]
101. Chesham A, Gerber SM, Schütz N, Saner H, Gutbrod K, Müri RM, et al. Search and Match Task: Development of a Taskified Match-3 Puzzle Game to Assess and Practice Visual Search. JMIR Serious Games 2019 May 09;7(2):e13620 [FREE Full text] [doi: 10.2196/13620] [Medline: 31094325]
102. Tong T, Chignell M, Sieminowski T. Case Study: A Serious Game for Neurorehabilitation Assessment. Procedia Computer Science 2015;69:125-131. [doi: 10.1016/j.procs.2015.10.013]
103. Tong T, Chignell M. Developing a Serious Game for Cognitive Assessment: Choosing Settings and Measuring Performance. 2014 Presented at: Chinese CHI '14: Second International Symposium of Chinese CHI; April 26, 2014; Toronto, Ontario, Canada p. 70-79. [doi: 10.1145/2592235.2592246]
104. Tong T, Chan JH, Chignell M. Serious Games for Dementia. 2017 Presented at: WWW '17 Companion: 26th International Conference on World Wide Web Companion; April 2017; Perth, Australia p. 1111-1115. [doi: 10.1145/3041021.3054930]
105. Khademi M, Hondori HM, McKenzie A, Dodakian L, Lopes CV, Cramer SC. Free-hand interaction with leap motion controller for stroke rehabilitation. 2014 Presented at: CHI EA '14: Human Factors in Computing Systems; April 2014; Toronto, Ontario, Canada p. 1663-1668. [doi: 10.1145/2559206.2581203]
106. Tan CW, Yu PD, Lin L. Teaching Computational Thinking Using Mathematics Gamification in Computer Science Game Tournaments. In: Kong SC, Abelson H, editors. Computational Thinking Education. Singapore: Springer Publishing Company; 2019:167-181.
107. Green CS, Seitz AR. The Impacts of Video Games on Cognition (and How the Government Can Guide the Industry). Policy Insights from the Behavioral and Brain Sciences 2015 Oct 01;2(1):101-110. [doi: 10.1177/2372732215601121]
108. Bavelier D, Green C, Seidenberg M. Cognitive development: gaming your way out of dyslexia? Curr Biol 2013 Apr 08;23(7):R282-R283 [FREE Full text] [doi: 10.1016/j.cub.2013.02.051] [Medline: 23578877]
109. Trajkoski M. Whack A Mole Game. App Store. URL: https://apps.apple.com/ms/app/whack-a-mole-game/id1233662398 [accessed 2019-12-26]
110. Jackson SA, Marsh HW. Development and validation of a scale to measure optimal experience: the flow state scale. Journal of Sport and Exercise Psychology 1996;18(1):17-35. [doi: 10.1123/jsep.18.1.17]
111. Dörrenbächer S, Kray J. The Impact of Game-Based Task-Shifting Training on Motivation and Executive Control in Children with ADHD. J Cogn Enhanc 2018 Jun 19;3(1):64-84. [doi: 10.1007/s41465-018-0083-2]

112. Friehs MA, Dechant M, Vedress S, Frings C, Mandryk RL. Effective Gamification of the Stop-Signal Task: Two Controlled Laboratory Experiments. JMIR Serious Games 2020 Sep 08;8(3):e17810 [FREE Full text] [doi: 10.2196/17810] [Medline: 32897233]
113. Lan L, Lian Z. Application of statistical power analysis – How to determine the right sample size in human health, comfort and productivity research. Building and Environment 2010 May;45(5):1202-1213. [doi: 10.1016/j.buildenv.2009.11.002]
114. Faul F, ErdFelder E, Lang AG, Buchner A. G*power 3.1 manual. Behav Res Methods 2007;39(2):175-191 [FREE Full text]
115. Faber J, Fonseca LM. How sample size influences research outcomes. Dental Press J Orthod 2014 Aug;19(4):27-29 [FREE Full text] [doi: 10.1590/2176-9451.19.4.027-029.ebo] [Medline: 25279518]
116. Aalbers T, Baars MAE, Olde Rikkert MGM, Kessels RPC. Puzzling with online games (BAM-COG): reliability, validity, and feasibility of an online self-monitor for cognitive performance in aging adults. J Med Internet Res 2013 Dec 03;15(12):e270 [FREE Full text] [doi: 10.2196/jmir.2860] [Medline: 24300212]
117. Fedele D, McConville A, Moon J, Thomas J. Topical Review: Design Considerations When Creating Pediatric Mobile Health Interventions: Applying the IDEAS Framework. J Pediatr Psychol 2019 Apr 01;44(3):343-348 [FREE Full text] [doi: 10.1093/jpepsy/jsy084] [Medline: 30346561]
118. GitHub. URL: https://github.com/ [accessed 2020-09-24]
119. MD2K. URL: https://md2k.org/ [accessed 2020-09-24]

Abbreviations

ADHD: attention deficit hyperactivity disorder
ADLs: activities of daily living
CAM: Confusion Assessment Method
DSR: Design Science Research
MD2K: Mobile Sensor Data-To-Knowledge
MDA: Mechanics, Dynamics and Aesthetics
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
OMDE: Objects, Mechanics, Dynamics, Emotions
RCT: randomized controlled trial
SDT: self-determination theory

Edited by N Zary; submitted 29.06.20; peer-reviewed by T Tong, P Robert, C Lara; comments to author 12.08.20; revised version received 05.10.20; accepted 08.02.21; published 18.05.21.

Please cite as:
Khaleghi A, Aghaei Z, Mahdavi MA
A Gamification Framework for Cognitive Assessment and Cognitive Training: Qualitative Study
JMIR Serious Games 2021;9(2):e21900
URL: https://games.jmir.org/2021/2/e21900
doi:10.2196/21900
PMID:33819164

Original Paper

Standardizing the Development of Serious Games for Physical Rehabilitation: Conceptual Framework Proposal

María Del Pilar Beristain-Colorado[1*], MSc; Jorge Fernando Ambros-Antemate[1*], MSc; Marciano Vargas-Treviño[2*], PhD; Jaime Gutiérrez-Gutiérrez[2*], PhD; Adriana Moreno-Rodriguez[3*], PhD; Pedro Antonio Hernández-Cruz[4*], PhD; Itandehui Belem Gallegos-Velasco[4*], PhD; Rafael Torres-Rosas[5*], PhD

[1]Doctorado en Biociencias, Facultad de Medicina y Cirugía, Universidad Autónoma “Benito Juárez” de Oaxaca, Oaxaca de Juárez, Mexico

[2]Escuela de Sistemas Biológicos e Innovación Tecnológica, Universidad Autónoma “Benito Juárez” de Oaxaca, Oaxaca de Juárez, Mexico

[3]Facultad de Ciencias Químicas, Universidad Autónoma “Benito Juárez” de Oaxaca, Oaxaca de Juárez, Mexico

[4]Laboratorio de genómica y proteómica, Centro de Investigación UNAM-UABJO, Facultad de Medicina y Cirugía UABJO, Oaxaca de Juárez, Mexico

[5]Laboratorio de Inmunología, Centro de Estudios en Ciencias de la Salud y la Enfermedad, Facultad de Odontología, Universidad Autónoma “Benito Juárez” de Oaxaca, Oaxaca de Juárez, Mexico

[*]all authors contributed equally

Corresponding Author:
María Del Pilar Beristain-Colorado, MSc
Doctorado en Biociencias
Facultad de Medicina y Cirugía
Universidad Autónoma “Benito Juárez” de Oaxaca
Ex Hacienda de Aguilera S/N, Calz. San Felipe del Agua.
Oaxaca de Juárez, 68120
Mexico
Phone: 52 9511332346
Email: pberistain.cat@uabjo.mx

Abstract

Background: Serious games have been used as supportive therapy for traditional rehabilitation. However, most are designed without a systematic process to guide their development from the phases of requirement identification, planning, design, construction, and evaluation, which reflect the lack of adaptation of rehabilitation requirements and thus the patient’s needs.

Objective: The aim of this study was to propose a conceptual framework with standardized elements for the development of information systems by using a flexible and an adaptable process centered on the patient’s needs and focused on the creation of serious games for physical rehabilitation.

Methods: The conceptual framework is based on 3 fundamental concepts: (1) user-centered design, which is an iterative design process focused on users and their needs at each phase of the process, (2) generic structural activities of software engineering, which guides the independent development process regardless of the complexity or size of the problem, and (3) gamification elements, which allow the transformation of obstacles into positive and fun reinforcements, thereby encouraging patients in their rehabilitation process.

Results: We propose a conceptual framework to guide the development of serious games through a systematic process by using an iterative and incremental process applying the phases of context identification, user requirements, planning, design, construction of the interaction devices and video game, and evaluation.

Conclusions: This proposed framework will provide developers of serious games a systematic process with standardized elements for the development of flexible and adaptable software with a high level of patient commitment, which will effectively contribute to their rehabilitation process.

(JMIR Serious Games 2021;9(2):e25854) doi:10.2196/25854

KEYWORDS

serious game; physical rehabilitation; framework; software engineering; gamification

Introduction

Background

Human motor skills can be affected by numerous adverse situations such as trauma, stroke, and degenerative diseases. Rehabilitation exercises play a fundamental role in reducing the degree of disability. The traditional assisted rehabilitation model consists of daily supervised exercise sessions with a therapist [1]. These exercises must maintain patient motivation through interactive and stimulating environments to be effective. The therapist must customize rehabilitation exercises according to the patient's needs. Many rehabilitation therapies are intense and involve numerous repetitions and exercises. Patients often experience frustration owing to mobility loss. It leads to states of depression, causing some patients to become discouraged and lose interest in therapeutic exercises [2-4]. To avoid this, some complementary techniques combined with traditional rehabilitation, such as serious games, allow enhanced recovery [5]. They offer a more attractive environment and maintain the interest in the process of motor rehabilitation, focusing on the game instead of its limitation [6]. Serious games are video games that are meant for education instead of entertainment [7]. Therefore, the game must impart additional experiences to the user in addition to knowledge or skills. Although several serious games for rehabilitation have been developed [5,8-10], there are still misunderstandings in spite of the systematic and standardized process for their creation. These misunderstandings cause the inadequate implementation of important elements such as specialist monitorization, motivation, game levels, and evaluation scales to objectively quantify the degree of the disability. This paper explains the creation of a conceptual framework with a systematic, standardized, flexible, and adaptable approach for the development of serious games in physical rehabilitation. A conceptual framework helps synthesize the knowledge of different areas to obtain a broad understanding of the topics [11]. This framework is based on the structural activities of software engineering applied in a user-centered design (UCD) approach with an iterative and incremental process that allows the visualization of prototypes from the beginning of development with gamification elements to increase commitment and motivation.

Related Works

Previous studies have developed serious game frameworks in various areas such as physical rehabilitation and education. The following papers were obtained when reviewing the existing literature.

Amengual et al [12] proposed a system based on a two-dimension (activities and incremental development) iterative process. It consists of 4 phases: project initiation, interaction elements, serious games, and evaluation. In project initiation, therapists identify the patients' needs. The interaction mechanism for the patient's movement is then selected. A serious game is created, and finally, the patient is clinically evaluated. The framework is based on (1) scrum to manage and control iterative work at a project level, (2) a web application development model because the authors consider its requirements to have a certain similarity with web development, and (3) a process that requires a set of sequential phases, where each phase attempts to meet or define some objective because as per Amengual et al [12], the development of a serious game for motor rehabilitation is similar to that of a clinical trial.

Ushaw et al [13] proposed a paradigm identifying a benefit delivery system for serious games. It is classified into 5 elements: repetition, exploration, strategy, progress, and social interaction. They proposed a triangle for resources (time), benefit (serious), and game (fun) based on the "iron triangle" of software development, which is focused on quality. The application development phase is carried out with design, implementation, testing, and assessment phases.

Olszewski et al [14] proposed a structured framework for game development in medical education. It is an iterative process comprising 3 phases of development (preparation, design, and development) and a formative evaluation process. In the preparation and design, a team of medical experts is created according to the serious game developed. They will state the necessary knowledge to the development team. A design script is created, visualizing the hospital room and the game organization through navigation elements. In the development phase, the illustration components of the game are made to improve visual communication interface learning. Prototypes are created for a team of experts to analyze and make adjustments. During evaluation, formative, design problems, functionality, and usability problems are identified in the game. Then, the finalized project is delivered.

Pirovano et al [15] proposed a four-step framework. The first step is exercise, which begins with the therapy goal through exercises and is classified as primary and secondary goals. The second step is virtualization. The primary goals are turned into virtual exercises. In the third step, the virtual exercise becomes a real serious game. Finally, the secondary goals are managed through a monitoring module to adjust the patient's progress.

Table 1 summarizes the characteristics of each framework with structural activities and identifies the gamification elements. None of the studies in the literature review proposed activities to build an interaction device.

The main differences between our framework and that mentioned in similar studies are as follows: (1) our framework contains 5 structural activities of software engineering applied to a UCD; (2) physical rehabilitation–oriented gamification elements were included and classified into 3 groups (flow enhancement, immersion, and progress), which are implemented in the design phase to motivate the patient and to generate an immersive environment, thereby preventing dropouts; and (3) we propose a phase to develop a data acquisition interface to process the patient's movements when commercial devices do not adapt to the rehabilitation process.

Table 1. Summary of the related studies.

Framework	Structural activity	Gamification elements	Information on interaction device
Amengual et al [12]	Project initiation Planning and control (communication and planning) Modeling Construction Evaluation (deploy)	Levels	No
Ushaw et al [13]	Serious goal and game-play mechanic (communication and planning) Design (modeling) Implementation (construction) Testing assessment (deploy)	Benefit delivery mechanic: repetition, exploration, strategy, reward, measurement	No
Olszewski et al [14]	Preparation (communication) Design (modeling) Development (construction) Formative evaluation (deploy)	—[a]	No
Pirovano et al [15]	Exercise definition (communication) Virtualization (modeling) Primary and secondary goals (modeling and construction) Game design (construction and deploy)	Feedback and motivational factors	No

[a]Not available.

Methods

Study Design

Our conceptual framework is based on 3 fundamental concepts: (1) software engineering because serious games are based on the principles of information systems; (2) UCD, which is an iterative design process that focuses on users and their needs in each phase of the project; and (3) gamification, which allows the transformation of obstacles into positive and fun reinforcements, thereby encouraging patients in their rehabilitation process. The concepts used in our framework proposal are shown in Figure 1.

Figure 1. Conceptual elements of the framework.

Framework in Software Engineering

The software engineering framework establishes a high level of abstraction for software development, applying concepts, models, and other elements. It provides solutions to a series of similar problems, generally describing the phases that must be followed to fix them without further detail of the activities in each phase [16]. The objective is that developers use the framework as a guide for the creation of software systems, applying its phases as "building blocks" depending on the problem.

Structural Activities in Software Engineering

The work associated with the development of information systems in software engineering is classified into generic structural activities [17,18], regardless of the field of application, project size, or complexity. The structural activities are communication, planning, modeling, construction, and deployments, which are defined below:

1. Communication: This activity focuses on identifying the context and key requirements of the system through collaboration between the client and the development team. This phase determines the information processed, developed interfaces, design restrictions, and validation criteria.
2. Planning: This activity identifies requirements and develops resource estimates. Development tasks are identified and a work plan is created. Then, techniques are applied to define a work path and the strategic goal of the project.
3. Modeling: With a multidisciplinary team, the models must understand the real entity and represent the characteristics that the users need in addition to the information obtained and transformed with the software. The models must meet these objectives at different abstraction levels, including the illustration of software from the user perspective and on a technical level for the development team.
4. Construction: In this activity, models are coded in a programming language, errors are detected through tests, and they are corrected, resulting in a smart operating software for the client or end user.
5. Deployment: The prototype is delivered to the end user. The customer must provide feedback on the project for improvements. The software development process is iterative and incremental, and as a result, several deployments are made until the software development is completed.

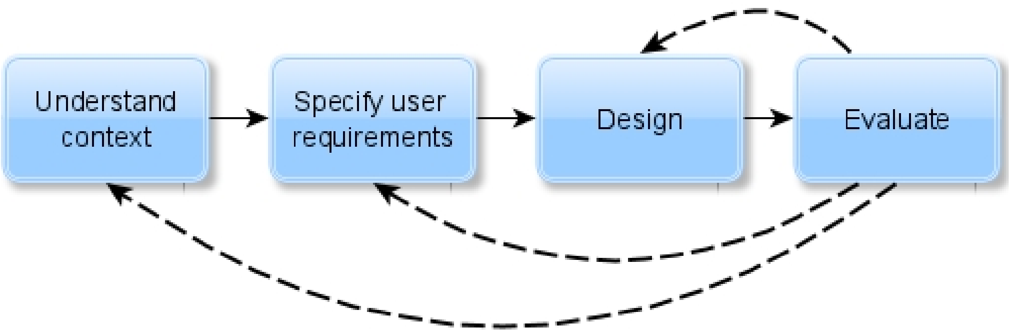
These 5 generic structural activities are used during software development. The process details will be different in each case, but the structural activities remain the same.

UCD

UCD is an iterative design process that focuses on users and their needs in each phase of the design process. Users are involved during the design process through research and design techniques to create highly usable and accessible products for them [19]. According to Karat and Karat [20], "UCD is characterized as a multi-phase problem-solving process that requires designers, analysis and foreseeing of the product or service employment, and verifying the validity of the behavioral assumptions in the real world."

UCD is an iterative process that includes the following key principles [21]: end users must be actively involved from the onset, and throughout the life cycle, the development comprises many iterative and incremental cycles to meet the end user requirements; design and prototypes must be created early and continuously to help visualize and evaluate ideas; and the development process must be performed by interdisciplinary teams. The principles described above facilitate the development, communication, and evaluation of the UCD to create interactive and useful systems, covering the design, evaluation, construction, and implementation phases of the product. Each UCD iteration involves 4 phases: contextual analysis, the definition of requirements, design, and evaluation. In the contextual analysis, the designer team must understand the context in which the system is used. When defining requirements, these are identified and specified according to the user. In the design phase, the team develops solutions such as simplified prototypes and designs on paper. In the last phase, results are evaluated from the assessment of the context and user requirements, verifying the design performance and satisfaction of relevant user needs. Depending on the results, the project team takes up phases again to optimize the product. These repetitions are performed until a satisfactory response is obtained from the users. Figure 2 shows the UCD process.

Figure 2. User-centered design process.



Need of a UCD for People With Disabilities

Helander and Landauer [22] established that people with disabilities have similarities with older adults since "they have multiple physical problems with a general reduction of their functionalities." People with special needs such as mobility limitations require an adaptable rehabilitation process for their needs. Thimbleby [23] stated that the purpose of UCD for a user with special needs is to increase their work productivity. Thus, the design must have end-user acceptance as they can feel more comfortable using the end product. To perform a physical rehabilitation process through serious games, the patient's movement to control the video game should be obtained. Therefore, the interaction device must be easy to wear and match the motor capacity of the patient. For example, if the therapy requires finger movement, it is difficult to put a haptic glove on the user (which could be a nuisance). Thus, the use of optical devices such as cameras and computer vision techniques are preferred.

Gamification

Gamification is a relatively new concept. Its objective is to apply game mechanics in different contexts to attract users to mundane but fun activities with motivational and cognitive benefits [24]. It allows the transformation of obstacles to positive and fun reinforcements, thereby encouraging users in making the right decisions for their health and well-being [25]. Several authors have proposed gamification elements for serious games, which are described below.

Cheek et al [26] identified design elements in a serious game called SPARX for adolescents with depression with a user-centered perspective. They identified 4 important areas and a series of associated elements: computer games (challenge, companionship, exploration, fantasy, and fidelity), accessibility (perceivable information, operable interface, and understandable, robust, and reliable information), working alliance (goal, task, and bond), and learning in immersion (situational learning, multiple perspective, real-life simulation, and immersive factors).

Zain et al [27] introduced a framework based on the flow theory of computer game usability and user experience. This framework consists of 8 elements: player skills, challenge, concentration, feedback, immersion, learning opportunities, accessibility, and adaptability.

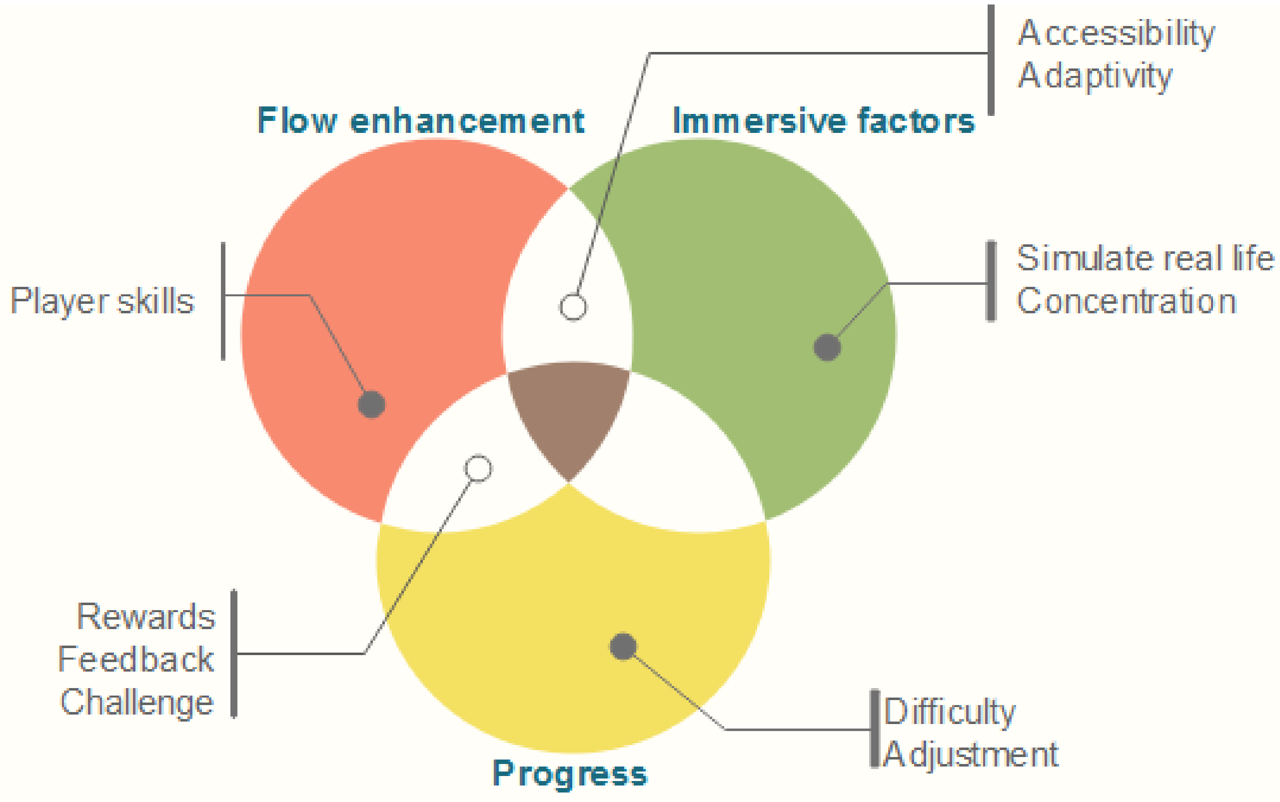
Specifically, Schulz et al [28] proposed a series of design element specifications based on functional and professional requirements: immersion, support for different roles, flow enhancement, visual enhancement, support for different learning phases and experience levels, design for interactivity, and progress.

Vermeir et al [29] presented a systemic review and meta-analysis of the gamification effects on computerized cognitive training. The elements identified were an avatar, challenge, competition, difficulty adjustment, feedback loops, levels, progress, rewards, social interaction, sound effects, and story/theme.

Bergeron [30] proposed a series of design elements: concept, game features, setting, story and backstory, effectors, game flow, screens and menus, control, options, sound and music, levels, score tracking, help, and localization.

From the gamification elements mentioned before, we chose those shared in every study and those that were convenient in a serious game for physical rehabilitation, and we classified the elements into 3 groups: flow enhancement, immersive factors, and progress. These concepts may appear in more than one group. For example, the element “rewards” is included in flow enhancement and progress. The shared gamification elements are shown in Figure 3.

Figure 3. Classification of gamification elements.



1. Challenge: According to Zain et al [27], the game must be challenging enough, and it has to match the player’s skill level. In physical rehabilitation, the game must adapt to the patient’s possibilities and be challenging enough to prevent boredom.
2. Accessibility: This element refers to the capacity to adapt to the patient’s disabilities. For example, when identifying hand movements in a patient who cannot hold an object, a camera can be used to track movements or a device that can be held by the patient.

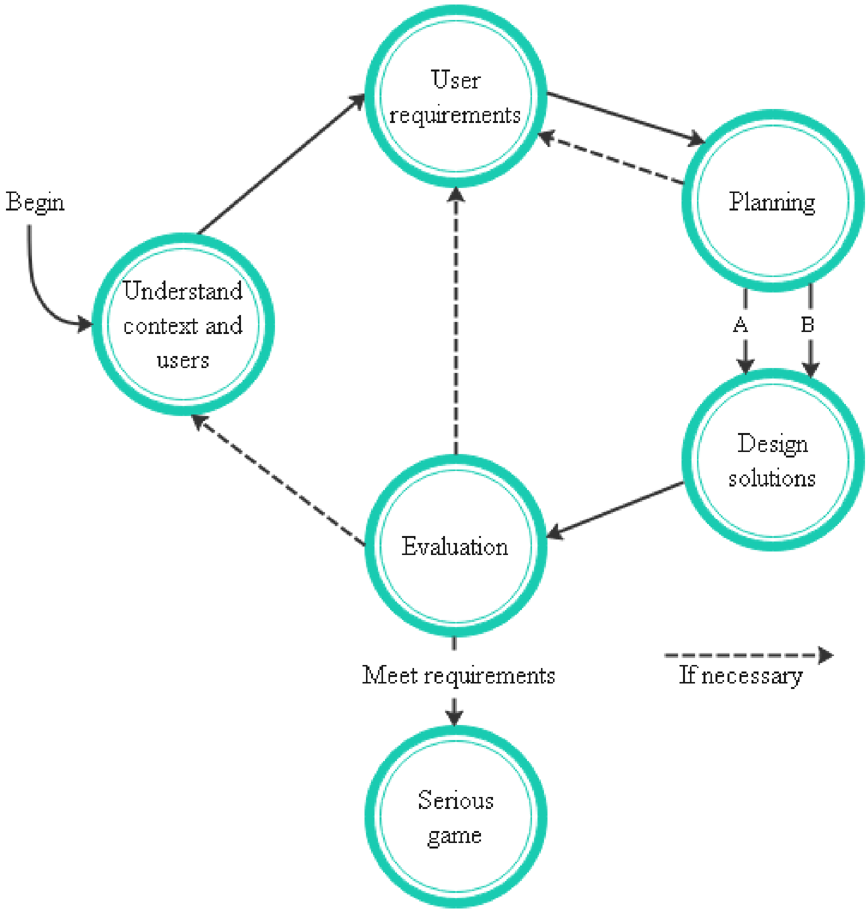
3. Adaptability: According to Zain et al [27], the user needs 3 factors: (1) user motivation, that is, why are you interested? (2) experiences and skills, that is, what skills are required to play? and (3) detection, that is, identify when a level change is necessary.
4. Player skill: The skill must be consistent with the serious game. As the game progresses, the patient develops more skills that motivate him/her to continue the rehabilitation process.
5. Rewards: Use indicators of the patient's progress such as points, virtual coins, badges, or any virtual object to motivate the user to continue with the rehabilitation process.
6. Real-life simulation: Games simulating real-life activities allow patients with physical disabilities to immerse themselves in the game.
7. Concentration: A serious game is motivating when the patients can fully focus on the game. Zain et al [27] indicate that serious games should attract the patient's attention at all times, avoiding distractions from the main task.
8. Feedback: There are different ways to provide feedback to the patient: (1) through progress, when the patient has correctly performed the exercise and must be motivated; (2) when indicating how to correctly perform an exercise; (3) through rewards with badges or virtual gifts when completing a challenge.
9. Difficulty adjustment: The serious game must be developed such that it allows the therapist to indicate the start level and make the necessary adjustments to the rehabilitation exercises.

Results

Proposal of a Conceptual Framework

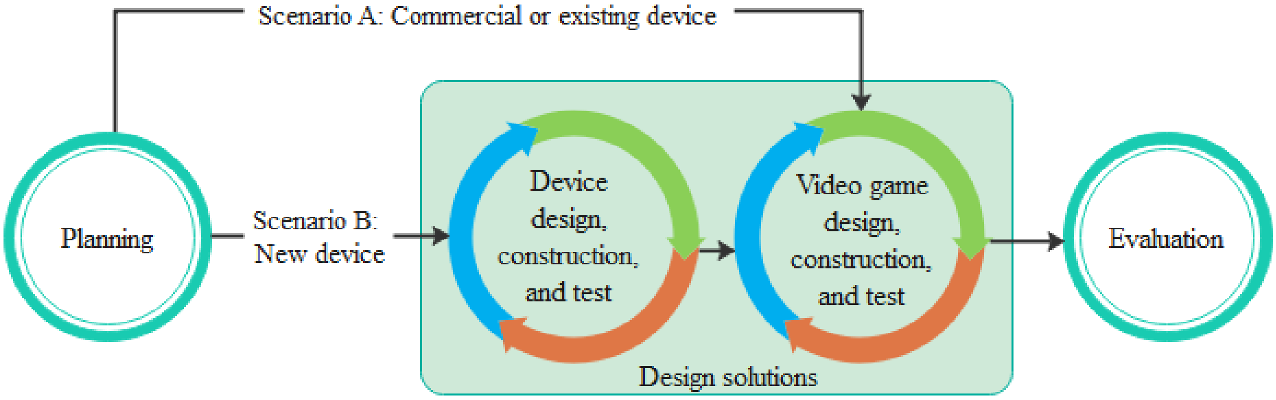
Few studies use a framework to develop serious games systematically. Therefore, our objective was to propose a conceptual framework based on UCD. Our framework consists of the adequate application of gamification elements and structural activities and guidance of meaningful, pleasant, relevant, and motivating serious games for physical rehabilitation. We use certain phases of the original UCD process, including a planning phase to establish estimates and priorities of the requirements and a modified designing phase to identify between creating an interaction device or using a commercial one. Figure 4 shows and describes the conceptual framework and the relation between the phases.

Figure 4. Proposed conceptual framework for the development of serious games for physical rehabilitation.



The framework begins in the phase of context and user understanding, and then the user requirements are identified. When new requirements are needed in the planning stage, it is necessary to return to the previous phase and include them in the user requirements. From planning to designing solutions, there are 2 possible scenarios shown in Figure 5: scenario A is applied when an existing or a commercial device is being used and scenario B is applied when creating a new device for a serious game. The evaluation has 3 options: (1) user requirements, in case of changes or improvements in the prototype; (2) context and user understanding, when clarifying the context for a requirement that needs to be adapted; and (3) serious game development when requirements are met, and the project satisfies the users' needs.

Figure 5. Contemplated scenarios.



Define Use of Context and Users

When developing a system or product, certain characteristics must be considered, such as context and user population with specific goals and tasks. Other conditions are technical, physical, social, or organizational that may affect its use. The quality of use of a system, user-friendliness, user health, and safety will depend on having an adequate understanding of the context. Identifying the correct context will help specify the user requirements and provide a solid foundation for subsequent evaluation activities. For well-known systems, the identification of stakeholders and context use review is sufficient. Further analysis of context and a study of existing users is required for more complex systems.

User Identification

The identification of direct and indirect users (people who influence or are affected by the system) ensures that every need is met and is tested as its construction progresses. User Mapping is a tool to identify users, as proposed by Taylor et al [31].

Analysis of the Context of Use

There are structured methods that obtain detailed information to understand the context of system use as a foundation for subsequent usability activities, particularly the specification and evaluation of user requirements. Some methods for context analysis have been proposed by Maguire [32], Taylor et al [31], and Thomas and Bevan [33]. For example, a guidebook for context analysis was developed by Thomas and Bevan [33], while Taylor et al [31] described the background and importance of understanding the context of use by developing a set of tools to identify the types of users, their needs, characteristics, and translation of this information into user requirements. This method is especially directed to nonexperts in the area of UCD and evaluation.

User Requirements

This stage identifies and documents the potential user requirements derived from the context information. Establishing and documenting user requirements will lead to the design process of a system [17]. User requirements include summarized descriptions of the system tasks and the features provided to support them. Therefore, user requirements describe the system characteristics to meet the context of use characteristics. Requirements engineering is needed to carry out this phase. It establishes a process of discovering, analyzing, documenting, and verifying requirements. Requirements engineering can be described in 5 distinct steps: requirement elicitation, requirement analysis and negotiation, requirement specification, system modeling, requirement validation, and requirement management [17]. Other techniques are proposed by Saiedan and Dale [34]. Once the requirements are obtained, they are analyzed with everyone involved. Then they must be documented with a user requirement(s) document or software requirements specification. An example of documenting requirements is the user stories used in the agile methodology XP [35].

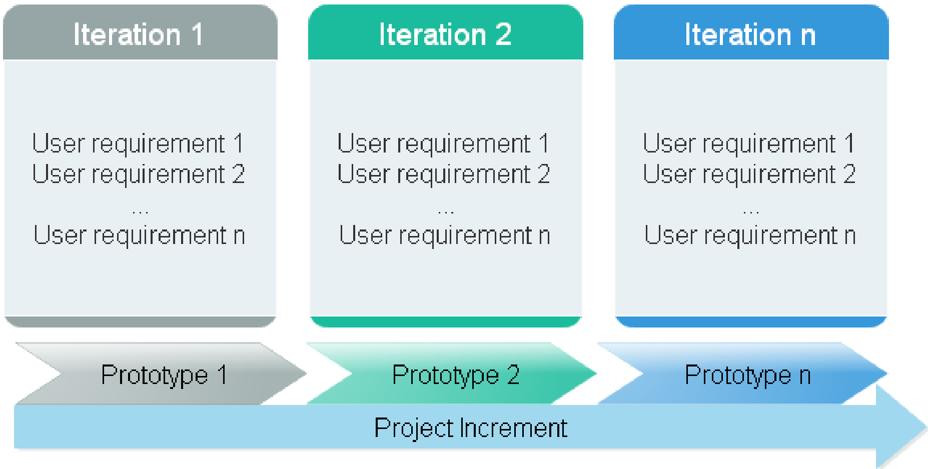
An essential requirement in physical rehabilitation is checking the patient progress and matching their levels. For example, the Wolf motor function test [36] or Fugl Meyer assessment [37] is used for upper limbs, the Berg Balance Scale [38] for balance and posture, and the Lower Extremity Motor Coordination Test for lower limbs [39]. These scales can be applied by the therapist or can be automated in the serious game. The latter must be added as a specific requirement for patient evaluation through the activities. The developed requirements are selected in each iteration. The requirements are adaptable to changes with the possibility of adding or removing them at any stage according to the system's needs. The iterations will conclude when the user requirements end. Measurement and compliance of the user requirements during development will result in a successful serious game that will improve patient safety, treatment effectiveness, and reduced rehabilitation time.

Planning

Several authors [17,18,40-42] have mentioned that planning is an essential part of project development. The work must be divided and assigned to the team members. Planning allows goal definition, objectives, and path to follow. The project size must be independently established with a project plan [43] containing at least the following elements [17,43]: team organization, risk analysis, requirements and estimation of resources, work division, and project schedule.

Once the user requirements are established, they must be divided by iterations to obtain a prototype in each cycle, as shown in Figure 6. Pressman [18] indicated that planning should be iterative and repeated at the end of each iteration based on therapist and patient feedbacks. Thus, planning is repeated in each iteration as user requirements are defined in the iteration, and time and resources are identified to successfully conclude the prototype. Planning should also be frequently monitored, and adjustments made as required. Assessing the progress daily will detect problem situations and adjust the plan accordingly.

Figure 6. Initial planning of user requirements.



Design Solutions

Designing is a creative activity where components and their relationships are identified based on user requirements. The team approaches designing through different solutions, and every idea must go through iterative development. The product meets the potential user needs through its development with some design elements such as mock-ups or interface screens for interaction, visualization, or comments. Another formal modeling such as UML [44] must be used by developers to represent the parts and communication of the system. Design changes can be made quickly in response to user feedback, and significant design issues can be identified before the system development begins. The solution is subsequently proposed through the prototypes. Hall [45] states the development of at least a low fidelity (for example, mock-ups) and a high fidelity (operational system, simulation) prototype. This will allow a usable product to satisfy the user requirements. Finally, tests must be run and possible errors must be corrected.

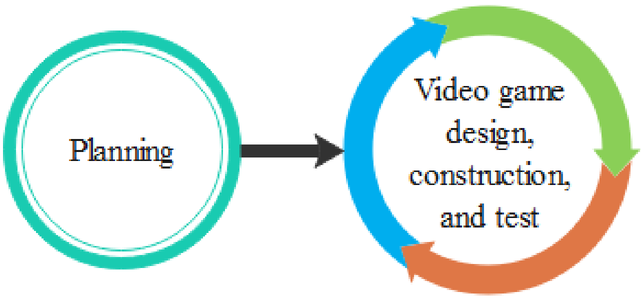
Device Design, Construction, and Test

Two possible scenarios are established in this phase, which are described below:

Scenario A: Use of an Existing or a Commercial Device

Devices previously created from an iteration or commercial devices such as Microsoft Kinect, Leap Motion, and Novint Falcon Game Controller are used in many serious game developments. The development team must ensure that it is safe and meets the patient's needs. Once the interaction device is selected, the creation phase of the video game is initiated. Figure 7 shows the transition phase.

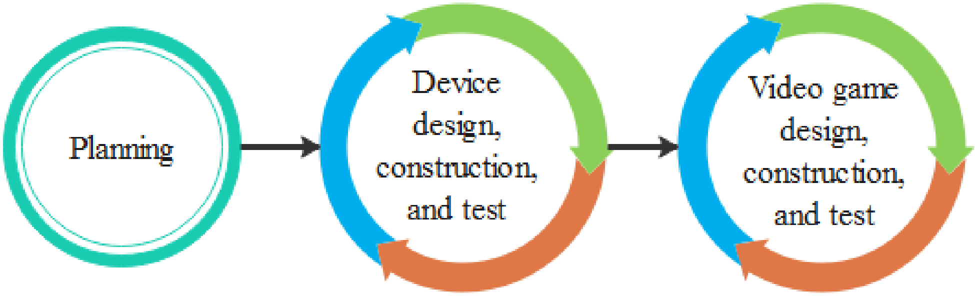
Figure 7. Scenario A: Use of an existing or commercial device.



Scenario B: New Device Development

This scenario occurs when the commercial device does not allow complete customization, and there are limitations in the data processing. This requires an additional phase to create personalized devices (eg, exoskeletons, gloves with inertial sensors) that match the motor skills of the patient for game movement and control. Figure 8 shows the phase to create a new device. Design in this phase creates models for the development team to understand the new device requirements. The models (for example, wireframes or mock-ups) help obtain comments and feedback for a better understanding of the operation. Devices are produced as a result of adequate designs in the construction phase [45]. The most appropriate components must be selected from the needs and limitations of the patients. Once the components are assembled, data processing is performed, including input, process, and output. Finally, the test activity allows testing and correcting possible errors to determine if the device meets the requirements.

Figure 8. Scenario B: New device development.



Design, Construction, and Testing of the Video Game

Design of the Video Game

In this phase, the video game prototype is developed, which is controlled by the interaction device of the previous phase. The previously analyzed and planned requirements are used in this phase. Regardless of the software's scope, size, or complexity, the software design must include at least four of the following models: data or class design, architecture design, interface design, and component-level design [18]. In data design, the data structures that will be required to implement the software are created, and data objects and relationships are defined. The architectural design defines the relationship between the major structural elements, and the "design patterns" can be used to achieve the requirements that have been defined for the system. The interface design describes the flow of information and how the software communicates within itself, with other systems that interoperate with it, and with humans who use it. The component-level design transforms structural elements of the software architecture into software components. Furthermore, game design elements are important to be included, which are described below.

Game Design Elements

In the Methods section, gamification was described and classified into 3 aspects: flow enhancement, immersive factors, and progress. Table 2 shows the benefits of gamification from game design elements. Therefore, it is of importance to include the following design elements:

Game Genre

Different game genres include action, adventure, music, puzzle, role-playing game, simulation, and strategy. The genre must be appropriate to the age of the patient in rehabilitation. For example, Chesham et al [46] indicated that puzzles are easy to understand, learn, and play for older adults.

Story or Narrative

According to Kuiper [47], a story is a series of events organized in a temporary order. According to Lu et al [48], the narrative influences the patient's cognition, affection, and potentially healthy behavior of the players. The story must be written according to the patient's average age and body part in rehabilitation.

Actors

According to Bergeron [30], primary actors in most games are player character(s), nonplayer character(s), vehicles, and effectors. Defining actors according to the rehabilitation process, history, and average age of the patient is important.

Effectors

Effectors are instruments that players have to interact with other game elements or to complete a mission. They are closely related with the interaction device (device that follows the patient's movements) and used in the rehabilitation process. They must agree with the story and avatar.

Screen and Menus

Bergeron [30] stated that an action or role-playing game must include startup, main menu, inventory, level, exit, and high-score screens. Screens and menus must be adapted to the physical limitations of the patient, for example, fine motor skill problems preventing the use of a mouse.

Levels

Baranowski et al [49] indicated that levels help players view their progress, thereby allowing the dominance of an action before moving to the next level. Levels must be associated with patient recovery in physical rehabilitation. Evaluation scales must be included to measure the patient's progress in a standardized way.

Help

Help shows game instructions to the user. It can be a small guide describing the movements to control the game or a document with frequently asked questions.

Sound and Music

Define different sounds according to the game's context. Bergeron [30] stated that music must be according to a particular situation, for example, when transmitting emotions such as happiness. Typical musical cues are needed for the introduction, level ending, high score, victory, and defeat.

Visual Enhancement

According to Schulz et al [28], games use visual cues to guide a player and provide options or interactive elements. Visualization helps patients become familiar with an environment, identify a real-life–like scenario, and intuitively select the effectors needed to accomplish a task. For example, if a player is presented with a dirty window on the screen, they will take a glass cleaner to clean it.

Table 2. Benefits of gamification through game design.

Game design	Gamification benefits
Game genre	Flow enhancement, immersive factors
Story or narrative	Flow enhancement, immersive factors, progress (challenge)
Actors	Flow enhancement, immersive factors (simulate real life)
Effectors	Immersion (simulate real life, accessibility), flow enhancement
Screens and menus	Flow enhancement (challenge, accessibility, rewards), progress
Levels	Flow enhancement, progress
Help	Flow enhancement (player skills), immersive factors (accessibility)
Sound and music	Immersive factors (adaptivity), feedback, rewards
Visual enhancement	Immersive factors (simulate real life, concentration), player skills, rewards, feedback

Construction of the Video Game

Software components, data, library, and other items are assembled at this stage to compile and link them to create an executable system.

Testing of the Video Game

Testing units may discover program defects before use. It has 2 distinct goals [17]: (1) show that the software meets the requirements of the development team and client and (2) find situations of wrong software behavior or not according to the specifications.

Evaluation

User-controlled testing is the most adequate method of assessment [50-52]. It consists of configuring system tests to perform a series of tasks by representative users. This can be configured in a controlled laboratory environment or with the developers. The objective is to collect information from the user's performance with the system, feedback, reactions, and observations. Another method is satisfaction questionnaires [53,54] with subjective impressions based on experiences with the system or a new prototype. Controlled clinical studies are recommended in the evaluation phase [55] to quantify the rehabilitation improvement with the exercises. The experiment, participants, and measurements must be defined according to the type of therapy [56]. If patients or therapists detect problems in the prototype created in an iteration, it must be solved in the previous iteration of the requirement or user context phase.

Discussion

Main Findings

The development of serious games for physical rehabilitation is a multidisciplinary process involving several elements: software development, design aspects, and direct involvement of health care specialists, patients, and other nonprofessional health care personnel. Although multiple developments have used UCD [57-60], they do not apply the structural activities required for a software system development [17,18,41,42]. Gamification allows obstacles to transform into positive and fun reinforcements in a physical rehabilitation process. The proposed framework considers gamification strategies and ensures their fulfillment with game design elements. This strategy is innovative since a similar proposal is not found in related literature. The reviewed studies described the concept but not the application of gamification in a development phase [26-28]. Serious games require a communication interface to control the video game. In physical rehabilitation, following a particular movement in a patient's limb or specific injury is required. Therefore, this conceptual framework includes scenarios to select the most appropriate device, including a commercial or existing device or the creation of a custom device. The frameworks of the reviewed studies did not consider the use or creation of interaction devices, as shown in Table 1.

Limitations

The authors acknowledge the limitations of this conceptual framework, such as validation, which has to be applied to patients requiring physical rehabilitation, and the generality in its description. However, the latter has the objective to provide a generic framework for physical rehabilitation with an understandable approach to development teams of serious games.

Opportunities for Further Research

This conceptual framework will be implemented in a serious game prototype application involving a health expert throughout the development process and validated by statistical analysis and clinical evaluation of patients.

Conclusions

Most serious games do not use a systematic process for their creation, thereby producing significant omissions in the rehabilitation process such as lack of rating scales to measure the patient's progress, no feedback, and exercises that do not adapt to the patient's disabilities. Therefore, this study provides a systematic process for the development of serious games for physical rehabilitation with the proposal of a conceptual framework. The framework applies 3 key concepts that increase the patient's adherence to rehabilitation therapy: UCD to understand the specific needs of patients, structural activities of software engineering for their development, and gamification elements, which aim to influence the behavior and motivation of users through the experiences obtained in the game. Access to this type of framework will assist development teams in the creation of safer, fun, motivating serious games, thereby improving the participation and commitment of patients. Finally, it would be essential that every serious game published in a journal be developed through a standardized process applying a framework, thereby ensuring that the game meets the minimum requirements necessary to satisfy user needs.

Acknowledgments

The authors would like to thank the National Council of Science and Technology (CONACyT) for their support of this study.

Conflicts of Interest

None declared.

References

1. Winstein CJ, Pohl PS, Lewthwaite R. Effects of physical guidance and knowledge of results on motor learning: support for the guidance hypothesis. Res Q Exerc Sport 1994 Dec;65(4):316-323. [doi: 10.1080/02701367.1994.10607635] [Medline: 7886280]
2. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228 [FREE Full text] [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
3. Das J. Post stroke depression: The sequelae of cerebral stroke. Neurosci Biobehav Rev 2018 Jul;90:104-114. [doi: 10.1016/j.neubiorev.2018.04.005] [Medline: 29656030]
4. Vieira ER, Brown E, Raue P. Depression in older adults: screening and referral. J Geriatr Phys Ther 2014;37(1):24-30. [doi: 10.1519/JPT.0b013e31828df26f] [Medline: 23619921]
5. Hatem SM, Saussez G, Della Faille M, Prist V, Zhang X, Dispa D, et al. Rehabilitation of Motor Function after Stroke: A Multiple Systematic Review Focused on Techniques to Stimulate Upper Extremity Recovery. Front Hum Neurosci 2016;10:442 [FREE Full text] [doi: 10.3389/fnhum.2016.00442] [Medline: 27679565]
6. Gotsis M. Games, Virtual Reality, and the Pursuit of Happiness. IEEE Comput. Grap. Appl 2009 Sep;29(5):14-19. [doi: 10.1109/MCG.2009.94]
7. Michael D, Chen S. Serius Games: Games That Educate, Train, and Inform. Mason, OH, USA: Course Technology PTR; Oct 05, 2005.
8. Eichenberg C, Schott M. Serious Games for Psychotherapy: A Systematic Review. Games Health J 2017 Jun;6(3):127-135. [doi: 10.1089/g4h.2016.0068] [Medline: 28628385]
9. Proença JP, Quaresma C, Vieira P. Serious games for upper limb rehabilitation: a systematic review. Disabil Rehabil Assist Technol 2018 Jan;13(1):95-100. [doi: 10.1080/17483107.2017.1290702] [Medline: 28359181]
10. Koutsiana E, Ladakis I, Fotopoulos D, Chytas A, Kilintzis V, Chouvarda I. Serious Gaming Technology in Upper Extremity Rehabilitation: Scoping Review. JMIR Serious Games 2020 Dec 11;8(4):e19071 [FREE Full text] [doi: 10.2196/19071] [Medline: 33306029]
11. Imenda S. Is There a Conceptual Difference between Theoretical and Conceptual Frameworks? Journal of Social Sciences 2017 Oct 09;38(2):185-195. [doi: 10.1080/09718923.2014.11893249]
12. Amengual Alcover E, Jaume-I-Capó A, Moyà-Alcover B. PROGame: A process framework for serious game development for motor rehabilitation therapy. PLoS One 2018;13(5):e0197383 [FREE Full text] [doi: 10.1371/journal.pone.0197383] [Medline: 29768472]
13. Ushaw G, Eyre J, Morgan G. A paradigm for the development of serious games for health as benefit delivery systems. 2017 Presented at: IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH); April 4; Perth, WA, Australia. [doi: 10.1109/SeGAH.2017.7939264]
14. Olszewski AE, Wolbrink TA. Serious Gaming in Medical Education: A Proposed Structured Framework for Game Development. Simul Healthc 2017 Aug;12(4):240-253. [doi: 10.1097/SIH.0000000000000212] [Medline: 28027076]
15. Pirovano M, Surer E, Mainetti R, Lanzi PL, Alberto Borghese N. Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames. Entertainment Computing 2016 May;14:55-65. [doi: 10.1016/j.entcom.2015.10.002]
16. Mnkandla E. About software engineering frameworks and methodologies. 2009 Presented at: AFRICON; 23-25 Sept; Nairobi, Kenya. [doi: 10.1109/AFRCON.2009.5308117]
17. Sommerville I. Software Engineering. United States of America: Pearson; 2016.
18. Pressman R. Software Engineering: A Practitioner's Approach. New York, USA: McGraw Hill Education; 2015.
19. O' Beirne R. Berkshire encyclopedia of human-computer interaction. Reference Reviews. URL: https://www.emerald.com/insight/content/doi/10.1108/09504120610685373/full/html [accessed 2006-07-01]
20. Karat J, Karat CM. The evolution of user-centered focus in the human-computer interaction field. IBM Syst J 2003;42(4):532-541. [doi: 10.1147/sj.424.0532]

21. Gulliksen J, Göransson B, Boivie I, Persson J, Blomkvist S, Cajander. Key Principles for User-Centred Systems Design. Behaviour and Information Technology 2005;22(6):397-409. [doi: 10.1080/01449290310001624329]
22. Helander M, Landauer T. Handbook of Human-Computer Interaction. North Holland: Elsevier; Aug 1, 1997.
23. Thimbleby H. Understanding User Centred Design (UCD) for People with Special Needs. Berlin, Heidelberg: Springer; 2008.
24. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. J Biomed Inform 2017 Jul;71:31-48 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.011] [Medline: 28536062]
25. Pereira P, Duarte E, Rebelo F, Noriega P. A review of gamification for health-related contexts. In: Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2014.
26. 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]
27. Zain N, Jaafar A, Razak F. SGameFlow framework: How to experience enjoyment in Serious Game (SG) for Motor Impaired Users (MIU). 2012 Presented at: International Conference on Computer & Information Science (ICCIS); June 12-14; Kuala Lumpur, Malaysia p. 1020. [doi: 10.1109/ICCISci.2012.6297175]
28. Schulz R, Martinez S, Hara T. Towards a game-design framework for evidence-based clinical procedure libraries. 2019 Presented at: IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); August 5-7; Kyoto, Japan. [doi: 10.1109/SeGAH.2019.8882474]
29. Vermeir J, White M, Johnson D, Crombez G, Ryckeghem L. The Effects of Gamification on Computerized Cognitive Trainingystematic Review and Meta-Analysis. JMIR Serious Games 2020 Aug 10;8:e-18644. [Medline: 32773374]
30. Bergeron B. Developing Serious Games. Hingham, Massachusetts: Charles River Media; 2006.
31. Taylor B, Catterall B, Galer M. The HUFIT planning, analysis and specification toolset: human factors as a normal part of the I.T. product design process. In: Taking Software Design Seriously. San Diego, CA, United States: Academic Press Professional, Inc; Jun 01, 1991:315-338.
32. Maguire M. Context of Use within usability activities. International Journal of Human-Computer Studies 2001 Oct;55(4):453-483. [doi: 10.1006/ijhc.2001.0486]
33. Thomas C, Bevan N. Usability context analysis: A practical guide. Loughborough University Repository. URL: https://repository.lboro.ac.uk/articles/report/Usability_context_analysis_a_practical_guide/9353600 [accessed 2021-06-16]
34. Saiedian H, Dale R. Requirements engineering: making the connection between the software developer and customer. Information and Software Technology 2000 Apr;42(6):419-428. [doi: 10.1016/S0950-5849(99)00101-9]
35. What is extreme programming (XP)? Agile Alliance. 2017. URL: https://www.agilealliance.org/glossary/xp/ [accessed 2021-02-28]
36. Wolf SL, Lecraw DE, Barton LA, Jann BB. Forced use of hemiplegic upper extremities to reverse the effect of learned nonuse among chronic stroke and head-injured patients. Exp Neurol 1989 May;104(2):125-132. [doi: 10.1016/s0014-4886(89)80005-6] [Medline: 2707361]
37. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehabil Med 1975;7(1):13-31. [Medline: 1135616]
38. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7-11. [Medline: 1468055]
39. Desrosiers J, Rochette A, Corriveau H. Validation of a new lower-extremity motor coordination test. Arch Phys Med Rehabil 2005 May;86(5):993-998. [doi: 10.1016/j.apmr.2004.11.007] [Medline: 15895347]
40. McDermid JA. Project planning and control. In: Software Engineer's Reference Book. Oxford, UK: Butterworth-Heinemann; 2014.
41. Rook P. Controlling software projects. Softw Eng J UK 1986;1(1):7. [doi: 10.1049/sej.1986.0003]
42. Schmidt R. Software Engineering Architecture-Driven Software Development. Waltham, USA: Morgan Kaufmann; 2013:978.
43. Vliet H. Software Engineering: Principles and Practice. USA: Wiley; 2010:978.
44. Rumbaugh J, Jacobson I, Booch G. The Unified Modeling Language Reference Manual. USA: Addison-Wesley Professional; Jul 2004:721.
45. Hall RR. Prototyping for usability of new technology. International Journal of Human-Computer Studies 2001 Oct;55(4):485-501. [doi: 10.1006/ijhc.2001.0478]
46. Chesham A, Wyss P, Müri RM, Mosimann UP, Nef T. What Older People Like to Play: Genre Preferences and Acceptance of Casual Games. JMIR Serious Games 2017 Apr 18;5(2):e8 [FREE Full text] [doi: 10.2196/games.7025] [Medline: 28420601]
47. Rimmong-Kenan K. Narrative Fiction: Contemporary Poetics. London, UK: Routledge; 2002.
48. 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]

49. Baranowski MT, Belchior PP, Chamberlin B, Mellecker R. Levels in Games for Health. Games Health J 2014 Apr;3(2):60-63. [doi: 10.1089/g4h.2014.0019] [Medline: 26196044]
50. Dumas J, Redish J. A Practical Guide to Usability Testing. United Kingdom: Intellect Books; 1999.
51. Nielsen J. Usability Engineering. Cambridge: Academic Press Professional, Inc; 1993.
52. Lindgaard G. Usability Testing and System Evaluation: A Guide for Designing Useful Computing Systems. UK: Nelson Thornes; 1994.
53. Chin J, Diehl V, Norman K. Development of an instrument measuring user satisfaction of the human-computer interface. 1988 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 15; Washington DC, USA. [doi: 10.1145/57167.57203]
54. Chin JP, Diehl VA, Norman KL. Development of an instrument measuring user satisfaction of the human-computer interface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1998 May 01 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 15 May, 1988-19 May, 1988; Washington p. 213-218. [doi: 10.1145/57167.57203]
55. Wright B, Shamley D. Introduction to clinical trials. In: A Comprehensive and Practical Guide to Clinical Trials. United Kingdom: Academic Press; 2017.
56. Martinez R. Clinical Trial Outcomes. JACC: Heart Failure 2019 Mar;7(3):272-273. [doi: 10.1016/j.jchf.2018.12.003]
57. Baranyi R, Willinger R, Lederer N, Grechenig T, Schramm W. Chances for serious games in rehabilitation of stroke patients on the example of utilizing the Wii Fit Balance Board. 2013 Presented at: IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH); November 21; Vilamoura, Portugal. [doi: 10.1109/SeGAH.2013.6665319]
58. Baranyi R, Czech P, Walcher F, Aigner C, Grechenig T. Reha@Stroke - A mobile application to support people suffering from a stroke through their rehabilitation. 2019 Oct 28 Presented at: IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); 2019; Kyoto, Japan. [doi: 10.1109/SeGAH.2019.8882447]
59. Dekker MR, Williams AD. The Use of User-Centered Participatory Design in Serious Games for Anxiety and Depression. Games Health J 2017 Dec;6(6):327-333. [doi: 10.1089/g4h.2017.0058] [Medline: 28956617]
60. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013 Jul 02;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]

Abbreviations

UCD: user-centered design

Edited by N Zary; submitted 18.11.20; peer-reviewed by U Juárez-Martínez, G Alor-Hernández, C Turcu; comments to author 08.01.21; revised version received 15.04.21; accepted 30.04.21; published 24.06.21.

Please cite as:
Beristain-Colorado MDP, Ambros-Antemate JF, Vargas-Treviño M, Gutiérrez-Gutiérrez J, Moreno-Rodriguez A, Hernández-Cruz PA, Gallegos-Velasco IB, Torres-Rosas R
Standardizing the Development of Serious Games for Physical Rehabilitation: Conceptual Framework Proposal
JMIR Serious Games 2021;9(2):e25854
URL: https://games.jmir.org/2021/2/e25854/
doi:10.2196/25854
PMID:34185003

Viewpoint

Gaming Your Mental Health: A Narrative Review on Mitigating Symptoms of Depression and Anxiety Using Commercial Video Games

Magdalena Kowal[1], BSc, MSc; Eoin Conroy[1], BSc, MSc; Niall Ramsbottom[1], BSc, MSc; Tim Smithies[1], BSc; Adam Toth[1], BSc, PhD; Mark Campbell[1], BA, GDip, PhD

Lero, The Science Foundation Ireland Research Centre for Software, Physical Education and Sport Sciences Department, University of Limerick, Castletroy, Limerick, Ireland

Corresponding Author:
Mark Campbell, BA, GDip, PhD
Lero, The Science Foundation Ireland Research Centre for Software
Physical Education and Sport Sciences Department
University of Limerick
p1041 PESS building
Castletroy, Limerick
Ireland
Phone: 353 061234944
Email: mark.campbell@ul.ie

Abstract

Globally, depression and anxiety are the two most prevalent mental health disorders. They occur both acutely and chronically, with various symptoms commonly expressed subclinically. The treatment gap and stigma associated with such mental health disorders are common issues encountered worldwide. Given the economic and health care service burden of mental illnesses, there is a heightened demand for accessible and cost-effective methods that prevent occurrence of mental health illnesses and facilitate coping with mental health illnesses. This demand has been exacerbated post the advent of the COVID-19 pandemic and the subsequent increase in incidence of mental health disorders. To address these demands, a growing body of research is exploring alternative solutions to traditional mental health treatment methods. Commercial video games have been shown to impart cognitive benefits to those playing regularly (ie, attention control, cognitive flexibility, and information processing). In this paper, we specifically focus on the mental health benefits associated with playing commercial video games to address symptoms of depression and anxiety. In light of the current research, we conclude that commercial video games show great promise as inexpensive, readily accessible, internationally available, effective, and stigma-free resources for the mitigation of some mental health issues in the absence of, or in addition to, traditional therapeutic treatments.

(JMIR Serious Games 2021;9(2):e26575) doi:10.2196/26575

KEYWORDS

commercial video games; mobile phone; clinical; mental health disorders; psychotherapy; pandemic; accessibility; health care

Introduction

Mental Health Disorders

Mental health disorders affect more than 14% of the global population [1] and are estimated to become 1 of the 3 major causes of morbidity and mortality by 2030 [2]. The current global pandemic and subsequent periods of economic uncertainty could increase the prevalence of symptoms of mental health disorders, thus increasing the ubiquitous and widespread requirement for mental health treatments [3]. This poses serious consequences for individuals and society by overburdening the current systems in place [4]. Depressive and anxiety disorders are the most prevalent mental health disorders in the general population, with nearly 264 million people (3.4% of global population) and 284 million people adversely affected by depression and anxiety disorders, respectively [1,5]. Depression and anxiety are often experienced simultaneously, with up to 81% of individuals having an anxiety disorder having a depressive disorder too in their lifetime [6]. Depression and anxiety are also associated with chronic physical comorbidities, such as somatoform disorders [7], cancer, stroke, acute coronary syndrome [8], cardiovascular diseases, diabetes, chronic pain, and visual and auditory impairments [9]. Notably, mental health problems are underreported worldwide [10,11]. Due to various types of stigma and beliefs associated with the causes of symptoms, individuals with a known mental disorder often fail to reach out for help [12-14]. Coupled with this, approximately 70% of those requiring treatment have limited access due to insufficient government funding [15]. These aforementioned factors contribute to the mental health *treatment gap* [15] that exists in high-income countries [16] and that is wider in low- to middle-income countries [17]. As of 2020, 63.2% of the world's population was reported to have access to the internet in 2020 [18], with a vast majority of households owning a computer (ie, 82% of households in the United Kingdom and 93% of households in Finland) [19]. This increase in internet access and computer ownership has facilitated accessibility to video gaming, with approximately 2.7 billion video gamers reported worldwide in 2020. Given the existing mental health care challenges (high costs, long waiting lists, limited technological support, and less alternatives to traditional mental health care), this paper examines recent research on the potential for commercial video games to ameliorate symptoms in the two most prevalent mental health disorders: depression and anxiety. We discuss the increase in prevalence and severity of these disorders globally and the potential benefits of commercial video games to meet some of the current mental health care challenges associated with these two disorders.

Background

From the advent of the current worldwide COVID-19 pandemic, subsequent social distancing, quarantining, and restrictive measures have been implemented. The increasing incidence of psychological risk factors such as loneliness [20], stress [21], and general poor mental health [22] have been highlighted globally and are a major cause of concern. This has been supported recently by Brooks et al [23], who identified that common psychological responses, such as poor mental health state, fear, posttraumatic symptoms, anxiety symptoms, anxiety-induced insomnia, and confusion, are increasingly occurring because of lockdowns and quarantines. Thus, not only is the development of these symptoms a problem but the risk of relapse and exacerbation is also elevated among such preexisting mental health issues [24]. Moreover, recent findings suggest that these issues are more likely to occur in developed countries [25]. Prolonged durations of isolation were also found to be associated with poor mental health, posttraumatic symptoms [23], and the so-called *lockdown loneliness* [26]. Unsurprisingly, statistics for anxiety and depressive disorders in April-May 2020 reveal that in comparison with the first half of 2019 [27], US adults were 3 times more likely to screen positive for one or both the disorders and 8 times more likely to screen positive for serious mental distress in 2020 (28%) than in 2018 (3.4%) [26]. Furthermore, it seems that young adults, aged between 18 and 44 years, are particularly affected by the pandemic [20,28,29]. In comparison with 2018, 32.4% of young individuals reported feeling downhearted or depressed with levels of satisfaction with life falling by 80% [29]. In addition, young adults were more affected by loneliness [20] and serious mental distress [28] than older age groups in 2020. Research around the world indicates that women are at a higher risk than men and are more prone to psychiatric disorders during the COVID-19 pandemic [3,20,21,23]. Finally, the increase in depressive and anxiety disorders has also been linked to an increased incidence of burnout, specifically due to financial losses, and interruption of professional activities and social life [30].

The Cost of Mental Health Care Globally

Although significant progress has been made toward developing empirically supported psychological and psychiatric treatments, a significant proportion of people with mental health problems do not receive these treatments [31]. Finding ways to effectively disseminate treatment is imperative. The most commonly implemented treatments for prevalent mental health issues (ie, depression and anxiety) are pharmacological, psychological, or a combination of the two. The monthly cost of these treatments can range drastically from country to country, with costs ranging from approximately US $370 in the United States [32] to between US $108 and US $340 on an average across Europe [33]. The high cost makes access to a broad number of socioeconomic groups across the globe difficult [34-36]. The cost and availability of mental health care could vary greatly among insurance companies, with health insurance policies not covering mental health care in all cases [4]. Finally, waiting lists for state health care around the globe are a major cause of concern. The median wait time in the European Union to receive an initial assessment of psychiatric needs is between 7 and 30 days [37]. For example, Ireland's waiting list for mental health services exceeds 10,000 individuals, and the corresponding waiting times can exceed 24 months for an initial assessment [38]. With such long waiting periods in public systems, many individuals might be forced to seek options in more expensive private health care settings. However, the cost barrier to private treatment is often a more salient issue. For example, private therapy sessions can cost US $30-$400 per hour in China, US $45-$224 in Canada, US $30-$60 in the United Kingdom, US $77-$132 in France [39], and US $70-$156 in Ireland [40]. With state-provided health care systems already overwhelmed, the cost barrier of private treatment is felt more acutely and contributes to increasing the socioeconomic divide in terms of access to mental health treatment [17]. With the current issues surrounding traditional mental health treatment highlighted above, there is a need for accessible, affordable, and empirically supported alternative treatment options.

Traditional Mental Health Care Versus Traditional Alternatives and Digital Technology

Traditional mental health care often includes cognitive behavioral therapy, psychosocial or self-care interventions, and the growing use of digital technologies, such as web-based and smartphone-based delivery [41]. Some therapists incorporate board games, such as tabletop role-playing games (RPGs) like *Dungeons and Dragons* [42] or games like *Dixit* [43], into their therapeutic sessions. There is a myriad of evidence for alternatives to traditional mental health care [44-47]. Some specialists include self-care lifestyle interventions (eg, diet, exercise, and meditation) [48,49]. Research documenting interventions with digital technologies has shown that these types of interventions can improve the quality and accessibility of mental health care in high- and low-income countries [50].

A multitude of mental health–focused smartphone apps have been developed to assist with the implementation of these interventions. Despite promising findings in this area, a recent meta-analysis by Lecomte et al [46] noted that to date, only approximately 3% of mental health apps have been thoroughly validated by research. Therefore, authors have called for high-quality research on the efficacy of mental health apps as a stand-alone self-management option and an additional treatment option for mental health issues. Extending the theme of technology, some researchers and clinical practitioners noted the promise of video games tailored to assist with the treatment and management of a wide variety of clinical issues [51]. These games have been found to be successful in improving mental health outcomes in a variety of domains [51-54]. Many bespoke video games (eg, the Pesky gNATs game) seem to be beneficial as alternatives to traditional treatments [55,56]. Nevertheless, only a few of these games are commercially available (ie, *Into the Woods* and *MindLight*), and unfortunately, most of these games cannot be used without the guidance of qualified personnel [57]. Given the dearth of research, lack of randomized controlled trials (RCTs), and meagre quality of literature on bespoke video games to date, Eichenberg and Schott [52] suggest that current findings (although promising) lack generalizability and that further research is required.

New Perspective: The Case for Commercial Video Games

Benefits of Commercial Video Games

In contrast to bespoke clinical video games being developed and customized to address clinical issues for a particular cohort or group, commercial video games are generally available video games primarily designed for entertainment and not therapeutic purposes. Commercial video games are widely accessible and have risen from a niche form of entertainment in the 1970s to become a ubiquitous part of modern society [58]. The ubiquity of the commercial video game industry can be attributed to several factors; however, none are more pertinent than both the enjoyment of play and widespread accessibility. The fact that commercial video games can be engaged with on a computer, mobile phone, or gaming console with internet access [58], and are affordable and readily available, ranging from free-to-play games to games worth US $71, they are commonplace in most households worldwide [59]. Initial studies by Granic et al [60] documented various potential benefits of playing commercial video games. Since then, a growing body of recent evidence suggests notable benefits related to the use of commercial video games for socialization [61], cognition [62-64], emotion regulation [65], and mental health [52,66]. With this in mind, this paper provides a state-of-the-art overview of the current research on the benefits of commercial video games specifically for mental health. With recent evidence and increased demands for mental health care, we aim to highlight the potential of commercial video games to be an inexpensive, readily accessible, worldwide, effective, and stigma-free resource for the prevention and mitigation of mental health issues.

Current Evidence for Commercial Video Games to Support Mental Health

The potential health benefits of commercial video games have been examined in the clinical population. For example, studies have investigated the advantages of using commercial video games among those with depression [67], anxiety [68], schizophrenia [69], cancer [70] and in older adults [71]. Steadman et al [66] demonstrated psychotherapeutic benefits of specific commercial video game genres (as listed by the authors: *first-person shooters and action games*, *RPGs*, *simulation games*, *miscellaneous and sports games*, *strategy*, and *other genres*) and have provided recommendations for mental health professionals (ie, rapport building and social co-operation). More recently, the potential of commercial video games to successfully combat symptoms of depression, anxiety, and stress is highlighted during the COVID-19 pandemic [72]. In the following sections, we present the latest research highlighting the potential for commercial video games to mitigate symptoms associated with the 2 most prevalent mental health disorders in the general population: anxiety and depression (refer to Table 1 for a summary).

Table 1. Selected studies of commercial video games and their reported positive mental health outcomes related to depression and anxiety.

Mental health aspect examined	Game or genre	Outcome	Reference
Prosocial behaviors and decreased loneliness	AVG[a], RPG[b], multiplayer games, and video game play	AVGs (specifically with multiplayer functions), RPGs, and video game play all showed benefits for socialization among clinical and nonclinical populations	[66,73,74]
Cognition	AVGs, strategy games, exergames, *Boson X* (CVG[c]), and *Rayman* (CVG)	All games identified a range of cognitive improvements (eg, high executive function and visuospatial perception) in clinical and nonclinical populations. In addition, games were found to mitigate symptoms of dyslexia	[66,74,75]
Goal achievement	*Portal 2* (CVG), *Team Fortress 2* (CVG), and RPGs	All games examined showed improvements in goal-setting behavior and motivation to attain said goals	[60]
Positive reappraisal and mood repair	*Portal 2* (CVG), *Mario Kart* (CVG), *Slenderman* (CVG), *Flappy Bird* (CVG), *Tap the Frog* (CVG), and RPGs	Related to goal achievement, all games showed benefits with mood repair, both in time and magnitude	[60,76,77]
Emotional regulation	*Portal 2* (CVG), *Slenderman* (CVG), *Flappy Bird* (CVG), *Tap the Frog* (CVG), RPGs, and video game play	All games facilitated coping with strong emotions and regulating strong emotive experiences	[60,65,73,74,77,78]
Depressive mood	*Candy Crush* (CVG), *Angry Birds* (CVG), *Limbo* (CVG), and casual games	Casual game interventions decreased negative affect by promoting enjoyment, flow states, and motivation	[73,79]
General anxiety	*MindLight* (strategy game), *Max and the Magic Marker* (CVG), *Rayman* (CVG), Nintendo Wii Exergames, and RPGs	Strategy games, CVG, and CVG exergames all showed significant decreases in general measures of anxiety both immediately after play and maintained with continual play	[74,80-83]
Anxiety prevention	*Rayman* (CVG)	*Rayman* reduced general anxiety to the same degree as a strategy game designed to prevent anxiety	[74,82]
State anxiety	*Plants vs. Zombies* (CVG), *Bejeweled II* (CVG), *Peggle* (CVG), *Bookworm Adventures* (CVG), and casual games	Casual games reduced state anxiety by promoting flow states and goal achievement	[68,84]
Trait anxiety	*Bejeweled II* (CVG), *Peggle* (CVG), *Bookworm Adventures* (CVG), and casual games	Casual games reduced trait anxiety by decreasing levels of general anxiety	[68]
Preoperative anxiety	*Angry Birds* (CVG)	*Angry Birds* reduced preoperative anxiety in children older than 36 months, while maintaining reduced levels of anxiety post operation	[85]
COVID-19 anxiety	CVG exergames	Helped combat anxiety related to the pandemic, lockdowns, and social isolation	[86]

[a]AVG: action video game.

[b]RPG: role-playing game.

[c]CVG: commercial video game.

Benefits of Commercial Video Games for Symptoms of Depression

The two core symptom criteria of depression are either chronic low mood or loss of interest and loss of pleasure (anhedonia). Depression is generally characterized by changes in affect, cognition, neurovegetative functions, and interepisode remissions [87]. Apart from physical manifestations, such as fatigue, insomnia or hypersomnia, weight loss, and digestive issues, individuals with depression can experience a diminished ability to think and concentrate, and experience greater indecisiveness, feelings of worthlessness, and excessive guilt with recurrent thoughts of death or suicide. Existing research has identified critical cognitive processes underlying depression, such as cognitive biases [88], deficits in cognitive control [89], difficulties with disengagement from negative stimuli [90], reduced inhibition [91], and increased rumination [92]. Such research has identified the use of distraction, acceptance, or reappraisal (flexible reframing of stimuli or situations to change their emotional valence or meaning) as effective strategies to fight depressive symptoms [93].

Regarding anhedonia experienced by individuals with depression, recent evidence has shown the utility of video games to evoke positive emotions such as joy and happiness [65], appreciation and competence [94], and social connectedness [73] in individuals. Referring to the latter, the literature shows that weak connectedness, loneliness, and social isolation [95]; a debilitated sense of psychological belonging [96,97]; and internalized stigma [98] are associated with depressive symptoms. Given the abovementioned links and accepting that the pandemic has entailed less movement and social restrictions, commercially available web-based multiplayer games might be a potentially viable tool to connect isolated individuals. For example, researchers have noted the efficacy in games such as *Minecraft* or *Animal Crossing: New Horizons* for social connectedness, fighting loneliness, maintenance of social interaction, and ultimately the alleviation of depressive symptoms [99-101].

An array of research alludes to the possible benefits of RPGs for individuals with depression. RPGs are most commonly recognized by users' immersion with an avatar or a played role and require achieving goals according to rules within a virtual world [102]. In early work, Steadman [66] noted that RPGs, through self-identification with played characters, could be used by psychotherapists to generate and test schemas on how individuals function. Moreover, it challenged their identified relations and promoted healthy actions. As such, Zayeni et al [74] highlighted that RPGs can serve as a therapeutic tool to challenge and even change ingrained patterns of thinking, promote generation of positive alternatives, and act as a means to question self-schemas among patients undergoing cognitive behavioral treatment. Such benefits are similar to those of the increasingly popular medium of play therapy [103]. Turning to cognitive behavioral outcomes, researchers have shown that games such as *Portal 2* or RPGs (ie, *World of Warcraft*, *Pokémon*, and *Final Fantasy*) can promote goal achievement and reappraisal and facilitate flexibility and efficient emotional regulation as an adaptive tool [60]. In addition, overall improvements in depressive mood with commercial video game use are seen with the use of the Wii racing game *Mario Kart* among adolescents [76]. In addition, this age group seems to benefit in general from gaming, as regular adolescent gamers showed superior emotional regulation abilities than nongamers [78].

In addition, an interesting 6-week EmotivaMente training intervention based on a range of 13 commercial video games has been recently developed and incorporated with school training for adolescents by Carissoli and Villani [77]. Although the study failed to show improvements in emotional intelligence outcomes and the use of adaptive emotion regulation strategies post training, participants showed improvements in cognitive reappraisal and the expression of emotions in relation to the self. Given these findings, it might be beneficial to thoroughly investigate a similar longitudinal intervention in individuals with depression. In addition, a very recent systematic review on casual commercial video games (ie, *Candy Crush* and *Angry Birds*), often played in short sessions on daily commutes, for example, reported that even 30-minute interactions are capable of reducing depressive symptoms and stress [79]. We also know that distraction (ie, reading a book, playing games, or music) serves as an effective mood-regulatory strategy when used in moderation. Some individuals with depression may find it difficult to initiate distraction spontaneously due to increased rumination and inhibition deficits [104]. A recent paper by Kühn et al [75] demonstrated that the fast-paced commercially available action video game *Boson X* not only improved cognitive outcomes but also showed promise in reducing rumination among individuals with depression.

Commercial video games can be helpful in evoking joy, happiness, and positive mood. Research conducted by Kneer et al [105] indicates that due to in-game success and challenge, the game *Team Fortress 2* can elicit positive postgame mood, and these mood benefits are conferred among players of all skill levels. Art therapy, with its use of visual artistic creation to focus on creativity, symbolism of colors or shapes, and the exploration and expression of personal emotions, could also be seen to permeate into commercial video games with positive outcomes for individuals with depression [106,107]. Distinctly designed gaming environments accompanied by tailored soundtracks serve as a safe setting for the exploration and interpretation of emotional states elicited by game graphics. Wolf [2] links these attributes with the gameplay of commercial video game titles, including *Limbo*, *Journey*, or *Flower*, noticing as well that *Limbo* players self-report improved depressive mood outcomes.

The Benefits of Commercial Video Games for Symptoms of Anxiety

Depression and anxiety are often comorbid disorders [6]. Consequently, exploring the potential of commercial video games to alleviate the symptoms of anxiety is also needed. Anxiety disorders (ie, excessive behavioral disturbances, worry, and fear) can cause significant distress and impairment in social, occupational, and other areas of an individual's functioning [87]. Separation anxiety, social anxiety, generalized anxiety disorders, anxiety-specific phobias, and panic attacks are some of the subclassifications of anxiety disorders [87]. Anxiety disorders are also associated with difficulty concentrating during work and play, and they can lead to excessive fear of negative evaluation. Different frameworks indicate anxious apprehension and automatic preparatory behavior as key components of anxiety [108,109].

In a recent systematic review, Zayeni et al [74] show that some commercially available video games can be effectively used to address some of the symptoms of anxiety disorders, help treat general symptoms of anxiety (ie, *MindLight* and *Max and the Magic Marker*), reduce measures of social anxiety (ie, *Adventures aboard the S.S. GRIN*), and promote anxiety prevention (ie, *Rayman*). Furthermore, Ohannessian [110] found that web-based video games lowered anxiety levels and increased levels of social connectedness in adolescent boys. Moreover, an intervention by Fish et al [84] compared a prescribed 30- to 45-minute session of *Plants vs. Zombies*, a tower defense game, 4 times per week with a traditional pharmacological selective serotonin reuptake inhibitors treatment for individuals with depression and comorbid anxiety. This study revealed a greater reduction in state anxiety severity (State-Trait Anxiety Inventory questionnaire) [111] in the gaming group than in the medication group. Similarly, prescribed commercial casual video games have been shown to significantly decrease both state and trait anxiety in patients diagnosed with depression and comorbid anxiety [68].

Use of commercial video games also appears to be beneficial in facilitating anxiety management in a clinical setting [112], showing effectiveness in improving the management of preoperative anxiety in pediatric patients [85]. Reduced levels of anxiety were also evident in patients with Parkinson disease who participated in an exergames intervention using the Nintendo Wii console [80]. Furthermore, commercially available exergames, which combine physical exercise and gaming, show promise for their anxiolytic effects. As presented in the systematic review and meta-analysis by Viana et al [81], it has been reported that these games provide similar results and high adherence when compared with rehabilitation or usual care, additionally bringing fun and enjoyment to anxious individuals using them. More recently, it has also been suggested that the use of commercial exergames might serve as a strategy to cope with anxiety during quarantine [86].

When comparing bespoke video games with commercially available video games, an RCT presented equal improvements in anxiety levels between groups of adolescents who played either the bespoke biofeedback game *Dojo* or the commercially available *Rayman 2: The Great Escape* 6 times over 3 weeks [82]. Another RCT study exhibited similar effects for 2 games believed to optimize emotional intensity, capturing the same motivation and engagement levels of players [83]. Initially designed for therapeutic purposes and now commercially available, *MindLight* was compared with another commercially available game, *Max and the Magic Marker*, sharing similar action mechanics. Both these games demonstrated prevention of elevated anxiety in children. The fact that commercial video games appear to be equally effective as bespoke video games in reducing symptoms of anxiety is extremely promising when considering their overall therapeutic potential, cost-effectiveness, and ease of use.

The Future of Commercial Video Games

The advent of virtual reality (VR) video games at home has begun, and with it, researchers and video game companies are beginning to find new solutions to address mental health support through VR gaming [113]. A recently published study by Pallavicini and Pepe [114] showed significant decreases in state anxiety and negative emotions, accompanied by increased levels of positive emotions, for young adults who played the commercially available VR games *Fruit Ninja VR* and *Audioshield*. Commercially available VR video games have great potential and are well suited for the implementation of cognitive behavioral techniques for the treatment of depressive and anxiety disorders. Given the immersive nature of VR technology and the controllability of the virtual environment, it could be particularly well suited for use in exposure therapy. Exposure therapy is often prescribed to individuals with anxiety disorders [115]. Although research is still preliminary, VR commercial video games show great promise for treating and coping with symptoms of mental health issues; however, further scientific research is required to validate these initial promising findings.

Mental health treatment options often emphasize the role of active involvement in the therapeutic process [116,117], which often results in reduced anxiety levels and stress-related symptoms. Pine et al [79] suggested that providing an option to choose among many casual video game options might be important in promoting autonomy and, thus, may increase the positive outcomes of interventions using commercial video games. Given the large variety (>1 million) [118] of existing commercial video games and varied means of their delivery (PCs, phones, tablets, consoles, and VR) on the market, there is a great opportunity for therapists and researchers to examine this wide range of commercially available games within the broad context of mental health. Therefore, further research should investigate the theoretical underpinnings of the effect of video game play on mitigating mental health symptoms. Although a growing body of research has examined the influence of commercial video gaming on mental health in diverse settings, the emphasis of future work should be on various populations and across different demographics. With the increased prevalence of mental health disorders, each age group could find a selection of commercial video games that are well suited for their needs, preferences, and age-related demands.

Some video gaming companies (ie, *thatgamecompany* and *Glitchers*) have noted the potential of video games to improve mental health outcomes and have based their game design on psychological theories and art and music therapies to favor the potential mental health benefits of users. Music engages several brain structures and neurotransmitters related to reward, motivation, pleasure, and stress [119], with music therapy purported to effectively reduce symptoms of depression and anxiety [120-122] and health promotion [123]. Many musical rhythm video games or games with singing inputs exist already (eg, *Guitar Hero, Let's Sing,* and *Beat Saber*). However, research examining singing games [124] and musical games [125,126] in general or the possible influence of music within video games on mental health benefits to date [127] is still in its infancy. Given that musical elements are essential to contemporary commercial video games (soundtracks), close collaborations between game developers and musicians already exist [128]. We argue that research could further this avenue of future multidisciplinary collaboration by considering and exploring the role played by such auditory and musical elements in improving mental health outcomes in gaming and digital contexts.

Conclusions

Given the aforementioned research on issues faced by mental health services and exacerbated by the recent worldwide pandemic, we are currently facing a serious threat to mental health globally. With many patients on waiting lists or unable to receive support, we refer to the purpose of this paper, which was to present the potential for commercial video games to ameliorate the symptoms of the 2 most prevalent mental health disorders: depression and anxiety. Overall, we have demonstrated the current evidence wherein commercial video games serve as a useful corollary and easily accessible tool for decreasing the severity of symptoms of depression and anxiety. Commercial video games are usually either freely available or available at a one-time, relatively low cost. In addition, commercial video games possess many important features needed (eg, schemas, controlled in-game scenarios related to specific mental health issues, design, user engagement, and immersive state of *flow*) to make them effective as a preventative tool or for supplementing traditional therapies in the treatment of anxiety and depressive disorders. Compared with most clinical treatments (eg, pharmacological and psychological), commercial video games do not require in-person supervision, can be accessed remotely, are low cost (often providing no financial burden), are readily available, are easy to implement, are portable, and carry a greatly reduced risk of adverse side effects. Moreover, gaming, as a form of entertainment, is already popular among those most affected by mental health issues these days, as the majority of gamers are 18 years or older (79% of gamers), with 41% being female [58]. An average gamer is 34 years old, with individuals aged between 18 and 34 years comprising 38% of video gamers, followed by individuals aged between 34 and 54 years (26%) [58].

In conclusion, the overall accessibility and pervasiveness of commercial video games within modern society positions them as an invaluable means of reaching individuals with mental health disorders, irrespective of age and sex, and individuals with limited access to mental health care, particularly relevant during the current pandemic. With mounting scientific evidence in support of the efficacy of commercial video games for improving mental health outcomes, commercial video games should be considered as a potential alternative for the improvement of various aspects of mental health globally.

Conflicts of Interest

None declared.

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858 [FREE Full text] [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Wolf L. An investigation of existing uses of therapy games. In: Video Games: New Directions in Game Research. Bielefeld: Verlag; 2019.
3. Gualano MR, Lo Moro G, Voglino G, Bert F, Siliquini R. Effects of covid-19 lockdown on mental health and sleep disturbances in Italy. Int J Environ Res Public Health 2020 Jul 2;17(13):- [FREE Full text] [doi: 10.3390/ijerph17134779] [Medline: 32630821]
4. Silva M, Resurrección DM, Antunes A, Frasquilho D, Cardoso G. Impact of economic crises on mental health care: a systematic review. Epidemiol Psychiatr Sci 2018 Dec 13;29:e7 [FREE Full text] [doi: 10.1017/S2045796018000641] [Medline: 30419997]
5. Ritchie H, Roser M. Mental Health. Our World in Data. 2018. URL: https://ourworldindata.org/mental-health#anxiety-disorders [accessed 2020-04-14]
6. Lamers F, van Oppen P, Comijs HC, Smit JH, Spinhoven P, van Balkom AJ, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). J Clin Psychiatry 2011 Mar;72(3):341-348. [doi: 10.4088/JCP.10m06176blu] [Medline: 21294994]
7. Haller H, Cramer H, Lauche R, Gass F, Dobos GJ. The prevalence and burden of subthreshold generalized anxiety disorder: a systematic review. BMC Psychiatry 2014 May 1;14:128 [FREE Full text] [doi: 10.1186/1471-244X-14-128] [Medline: 24886240]
8. Kang H, Kim S, Bae K, Kim S, Shin I, Yoon J, et al. Comorbidity of depression with physical disorders: research and clinical implications. Chonnam Med J 2015 May;51(1):8-18 [FREE Full text] [doi: 10.4068/cmj.2015.51.1.8] [Medline: 25914875]
9. Vancampfort D, Koyanagi A, Hallgren M, Probst M, Stubbs B. The relationship between chronic physical conditions, multimorbidity and anxiety in the general population: A global perspective across 42 countries. Gen Hosp Psychiatry 2017;45:1-6. [doi: 10.1016/j.genhosppsych.2016.11.002] [Medline: 28274332]
10. Ritchie H. Global Mental Health: Five Key Insights Which Emerge From the Data. Our World in Data. 2018. URL: https://ourworldindata.org/global-mental-health [accessed 2020-05-01]
11. Takayanagi Y, Spira AP, Roth KB, Gallo JJ, Eaton WW, Mojtabai R. Accuracy of reports of lifetime mental and physical disorders: results from the Baltimore Epidemiological Catchment Area study. JAMA Psychiatry 2014 Mar;71(3):273-280 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.3579] [Medline: 24402003]
12. Schnyder N, Panczak R, Groth N, Schultze-Lutter F. Association between mental health-related stigma and active help-seeking: systematic review and meta-analysis. Br J Psychiatry 2017 Apr;210(4):261-268. [doi: 10.1192/bjp.bp.116.189464] [Medline: 28153928]
13. Stolzenburg S, Freitag S, Evans-Lacko S, Speerforck S, Schmidt S, Schomerus G. Individuals with currently untreated mental illness: causal beliefs and readiness to seek help. Epidemiol Psychiatr Sci 2019 Aug;28(4):446-457 [FREE Full text] [doi: 10.1017/S2045796017000828] [Medline: 29335036]

14. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet 2007 Sep;370(9590):841-850. [doi: 10.1016/S0140-6736(07)61414-7] [Medline: 17826169]
15. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
16. Alonso J, Liu Z, Evans-Lacko S, Sadikova E, Sampson N, Chatterji S, WHO World Mental Health Survey Collaborators. Treatment gap for anxiety disorders is global: results of the World Mental Health Surveys in 21 countries. Depress Anxiety 2018 Mar;35(3):195-208 [FREE Full text] [doi: 10.1002/da.22711] [Medline: 29356216]
17. Wainberg ML, Scorza P, Shultz JM, Helpman L, Mootz JJ, Johnson KA, et al. Challenges and opportunities in global mental health: a research-to-practice perspective. Curr Psychiatry Rep 2017 May;19(5):28 [FREE Full text] [doi: 10.1007/s11920-017-0780-z] [Medline: 28425023]
18. World Internet Usage and Population Statistics. World Stats. URL: https://www.internetworldstats.com/stats.htm [accessed 2021-06-01]
19. Alsop T. Share of Individuals Who Have Access to a Computer in Their Household in 2020, By Country. Statista. URL: https://www.statista.com/statistics/1107826/access-to-computer-in-households-worldwide/ [accessed 2021-03-01]
20. Li LZ, Wang S. Prevalence and predictors of general psychiatric disorders and loneliness during COVID-19 in the United Kingdom. Psychiatry Res 2020 Sep;291:113267 [FREE Full text] [doi: 10.1016/j.psychres.2020.113267] [Medline: 32623266]
21. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Global Health 2020 Jul 6;16(1):57 [FREE Full text] [doi: 10.1186/s12992-020-00589-w] [Medline: 32631403]
22. Wang Z, Yang H, Yang Y, Liu D, Li Z, Zhang X, et al. Prevalence of anxiety and depression symptom, and the demands for psychological knowledge and interventions in college students during COVID-19 epidemic: A large cross-sectional study. J Affect Disord 2020 Oct 1;275:188-193 [FREE Full text] [doi: 10.1016/j.jad.2020.06.034] [Medline: 32734907]
23. 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]
24. Rajkumar RP. COVID-19 and mental health: a review of the existing literature. Asian J Psychiatr 2020 Apr 10;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
25. Jia R, Ayling K, Chalder T, Massey A, Broadbent E, Coupland C, et al. Mental health in the UK during the COVID-19 pandemic: cross-sectional analyses from a community cohort study. BMJ Open 2020 Sep 15;10(9):e040620 [FREE Full text] [doi: 10.1136/bmjopen-2020-040620] [Medline: 32933965]
26. Shah SG, Nogueras D, van Woerden HC, Kiparoglou V. The covid-19 pandemic: a pandemic of lockdown loneliness and the role of digital technology. J Med Internet Res 2020 Nov 5;22(11):e22287 [FREE Full text] [doi: 10.2196/22287] [Medline: 33108313]
27. Twenge JM, Joiner TE. US Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. Depress Anxiety 2020 Oct;37(10):954-956 [FREE Full text] [doi: 10.1002/da.23077] [Medline: 32667081]
28. Twenge JM, Joiner TE. Mental distress among US adults during the COVID-19 pandemic. J Clin Psychol 2020 Dec;76(12):2170-2182 [FREE Full text] [doi: 10.1002/jclp.23064] [Medline: 33037608]
29. Social Impact of COVID-19 Survey. Central Statistics Office. URL: https://www.cso.ie/en/releasesandpublications/ep/p-sic19/socialimpactofcovid-19surveyapril2020/introductionandsummaryofresults/ [accessed 2021-06-01]
30. de Lima CV, Cândido EL, da Silva JA, Albuquerque LV, Soares LD, do Nascimento MM, et al. Effects of quarantine on mental health of populations affected by Covid-19. J Affect Disord 2020 Oct 1;275:253-254 [FREE Full text] [doi: 10.1016/j.jad.2020.06.063] [Medline: 32734916]
31. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct 9;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
32. Carmona R, Cook BL, Baca-García E, Chavez L, Alvarez K, Iza M, et al. Comparison of mental health treatment adequacy and costs in public hospitals in Boston and Madrid. J Behav Health Serv Res 2019 Jul 7;46(3):464-474 [FREE Full text] [doi: 10.1007/s11414-018-9596-9] [Medline: 29516339]
33. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Oct;21(10):718-779. [doi: 10.1016/j.euroneuro.2011.08.008] [Medline: 21924589]
34. Kazdin AE. Annual research review: expanding mental health services through novel models of intervention delivery. J Child Psychol Psychiatry 2019 Apr 13;60(4):455-472. [doi: 10.1111/jcpp.12937] [Medline: 29900543]
35. Knapp M, Funk M, Curran C, Prince M, Grigg M, McDaid D. Economic barriers to better mental health practice and policy. Health Policy Plan 2006 May;21(3):157-170. [doi: 10.1093/heapol/czl003] [Medline: 16522714]
36. Marquez V, Saxena S. Making mental health a global priority. Cerebrum 2016;2016:- [FREE Full text] [Medline: 28058091]

37. Barbato A, Vallarino M, Rapisarda F, Lora A. Access to Mental Health Care in Europe. Lisbon Institute. 2016. URL: https://www.lisboninstitutegmh.org/assets/files/Access%20to%20mental%20health%20care%20in%20Europe.pdf [accessed 2021-06-07]
38. -. In: Psychological Society Ireland. Switzerland: Springer, Cham; 2020.
39. Pajer N. This Is What Therapy Is Like Around The World. Huffington Post UK. 2019. URL: https://www.huffpost.com/entry/what-therapy-is-like-around-the-world_l_5d2f2930e4b02fd71dddab1c [accessed 2021-06-07]
40. McBride L. Talk Isn't Cheap When It's With a Counsellor - Unless You Wait. Independent.ie. 2020. URL: https://www.independent.ie/business/personal-finance/talk-isnt-cheap-when-its-with-a-counsellor-unless-you-wait-38979826.html [accessed 2000-06-07]
41. Patel V, Chisholm D, Parikh R, Charlson FJ, Degenhardt L, Dua T, et al. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. Lancet 2016 Apr;387(10028):1672-1685. [doi: 10.1016/S0140-6736(15)00390-6] [Medline: 26454360]
42. Game to Grow. 2017. URL: https://gametogrow.org/ [accessed 2021-10-25]
43. Mousnier E, Knaff L, Es-Salmi A. The dixit' card game as a support for metaphorical representations: a medium for systemic therapy under mandate. Thérapie Familiale 2016;37(4):363-386. [doi: 10.3917/tf.164.0363]
44. Batra S, Baker RA, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. Med Devices (Auckl) 2017;10:237-251 [FREE Full text] [doi: 10.2147/MDER.S144158] [Medline: 29042823]
45. Langarizadeh M, Tabatabaei M, Tavakol K, Naghipour M, Rostami A, Moghbeli F. Telemental health care, an effective alternative to conventional mental care: a systematic review. Acta Inform Med 2017 Dec;25(4):240-246 [FREE Full text] [doi: 10.5455/aim.2017.25.240-246] [Medline: 29284913]
46. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. JMIR Mhealth Uhealth 2020 May 29;8(5):e17458. [doi: 10.2196/17458] [Medline: 32348289]
47. Spijkerman M, Pots W, Bohlmeijer E. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114. [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
48. Dale H, Brassington L, King K. The impact of healthy lifestyle interventions on mental health and wellbeing: a systematic review. Mental Health Rev J 2014 Mar 5;19(1):1-26. [doi: 10.1108/mhrj-05-2013-0016]
49. Walsh R. Lifestyle and mental health. Am Psychol 2011 Oct;66(7):579-592. [doi: 10.1037/a0021769] [Medline: 21244124]
50. Naslund JA, Shidhaye R, Patel V. Digital technology for building capacity of nonspecialist health workers for task sharing and scaling up mental health care globally. Harv Rev Psychiatry 2019;27(3):181-192. [doi: 10.1097/hrp.0000000000000217] [Medline: 30958400]
51. Fitzgerald M, Ratcliffe G. Serious games, gamification, and serious mental illness: a scoping review. Psychiatr Serv 2020 Feb 1;71(2):170-183. [doi: 10.1176/appi.ps.201800567] [Medline: 31640521]
52. Eichenberg C, Schott M. Serious games for psychotherapy: a systematic review. Games Health J 2017 Jun;6(3):127-135. [doi: 10.1089/g4h.2016.0068] [Medline: 28628385]
53. Khanna MS, Kendall PC. Computer-assisted cognitive behavioral therapy for child anxiety: results of a randomized clinical trial. J Consult Clin Psychol 2010 Oct;78(5):737-745. [doi: 10.1037/a0019739] [Medline: 20873909]
54. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. Br Med J 2012 May 18;344:e2598 [FREE Full text] [doi: 10.1136/bmj.e2598] [Medline: 22517917]
55. Chapman R, Loades M, O'Reilly G, Coyle D, Patterson M, Salkovskis P. 'Pesky gNATs': investigating the feasibility of a novel computerized CBT intervention for adolescents with anxiety and/or depression in a Tier 3 CAMHS setting. Cogn Behav Ther 2016 Dec 1;9:-. [doi: 10.1017/s1754470x16000222]
56. McCashin D, Coyle D, O'Reilly G. A qualitative evaluation of Pesky gNATs in primary care - the experiences of assistant psychologists providing computer-assisted CBT to children experiencing low mood and anxiety. Internet Interv 2020 Dec;22:100348 [FREE Full text] [doi: 10.1016/j.invent.2020.100348] [Medline: 32904717]
57. Göbel S. Serious Games. Cham, Switzerland: Springer; 2020:2016-2405.
58. Gough C. Share of Video Gamers Worldwide in 2020, by Region. Statista. 2020. URL: https://www.statista.com/statistics/297874/number-mobile-gamers-region/ [accessed 2021-06-01]
59. Yan M, Gilbert B. Here's the Reason Most New Console Video Games Cost $60. Business Insider. 2018. URL: https://www.businessinsider.com/why-video-games-always-cost-60-dollars-2018-10 [accessed 2000-06-07]
60. Granic I, Lobel A, Engels RC. The benefits of playing video games. Am Psychol 2014 Jan;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
61. Neville RD, Guo Y, Boreham CA, Lakes KD. Longitudinal association between participation in organized sport and psychosocial development in early childhood. J Pediatr 2021 Mar;230:152-60.e1. [doi: 10.1016/j.jpeds.2020.10.077] [Medline: 33157074]
62. Green C, Bavelier D. Action video game modifies visual selective attention. Nature 2003 May 29;423(6939):534-537 [FREE Full text] [doi: 10.1038/nature01647] [Medline: 12774121]

63. Toth A, Ramsbottom N, Kowal M, Campbell M. Converging evidence supporting the cognitive link between exercise and esport performance: a dual systematic review. Brain Sci 2020 Dec 15;10(11):- [FREE Full text] [doi: 10.3390/brainsci10110859] [Medline: 33203067]
64. Kowal M, Toth AJ, Exton C, Campbell MJ. Different cognitive abilities displayed by action video gamers and non-gamers. Comput Hum Behav 2018 Nov;88:255-262. [doi: 10.1016/j.chb.2018.07.010]
65. Hemenover SH, Bowman ND. Video games, emotion, and emotion regulation: expanding the scope. Ann Int Commun Assoc 2018 Feb 21;42(2):125-143. [doi: 10.1080/23808985.2018.1442239]
66. Steadman J, Boska C, Lee C, Lim X, Nichols N. Using popular commercial video games in therapy with children and adolescents. J Technol Hum Serv 2014 Aug 8;32(3):201-219. [doi: 10.1080/15228835.2014.930680]
67. Russoniello CV, Fish M, O'Brien K. The efficacy of casual videogame play in reducing clinical depression: a randomized controlled study. Games Health J 2013 Dec;2(6):341-346. [doi: 10.1089/g4h.2013.0010] [Medline: 26197075]
68. Fish MT, Russoniello CV, O'Brien K. The efficacy of prescribed casual videogame play in reducing symptoms of anxiety: a randomized controlled study. Games Health J 2014 Oct;3(5):291-295. [doi: 10.1089/g4h.2013.0092] [Medline: 26192483]
69. Suenderhauf C, Walter A, Lenz C, Lang UE, Borgwardt S. Counter striking psychosis: commercial video games as potential treatment in schizophrenia? A systematic review of neuroimaging studies. Neurosci Biobehav Rev 2016 Sep;68:20-36. [doi: 10.1016/j.neubiorev.2016.03.018] [Medline: 27090742]
70. Comello ML, Francis DB, Marshall LH, Puglia DR. Cancer survivors who play recreational computer games: motivations for playing and associations with beneficial psychological outcomes. Games Health J 2016 Aug;5(4):286-292. [doi: 10.1089/g4h.2016.0003] [Medline: 27487462]
71. Zeng N, Pope Z, Lee J, Gao Z. A systematic review of active video games on rehabilitative outcomes among older patients. J Sport Health Sci 2017 Mar;6(1):33-43 [FREE Full text] [doi: 10.1016/j.jshs.2016.12.002]
72. Vatansever D, Wang S, Sahakian BJ. Covid-19 and promising solutions to combat symptoms of stress, anxiety and depression. Neuropsychopharmacology 2021 Jan;46(1):217-218 [FREE Full text] [doi: 10.1038/s41386-020-00791-9] [Medline: 32792683]
73. Colder Carras M, Kalbarczyk A, Wells K, Banks J, Kowert R, Gillespie C, et al. Connection, meaning, and distraction: a qualitative study of video game play and mental health recovery in veterans treated for mental and/or behavioral health problems. Soc Sci Med 2018 Dec;216:124-132. [doi: 10.1016/j.socscimed.2018.08.044] [Medline: 30257787]
74. Zayeni D, Raynaud J, Revet A. Therapeutic and preventive use of video games in child and adolescent psychiatry: a systematic review. Front Psychiatry 2020;11:36 [FREE Full text] [doi: 10.3389/fpsyt.2020.00036] [Medline: 32116851]
75. Kühn S, Berna F, Lüdtke T, Gallinat J, Moritz S. Fighting depression: action video game play may reduce rumination and increase subjective and objective cognition in depressed patients. Front Psychol 2018;9:129 [FREE Full text] [doi: 10.3389/fpsyg.2018.00129] [Medline: 29483888]
76. Rieger D, Wulf T, Kneer J, Frischlich L, Bente G. The winner takes it all: the effect of in-game success and need satisfaction on mood repair and enjoyment. Comput Hum Behav 2014 Oct;39:281-286. [doi: 10.1016/j.chb.2014.07.037]
77. Carissoli C, Villani D. Can videogames be used to promote emotional intelligence in teenagers? Results from emotivamente, a school program. Games Health J 2019 Dec;8(6):407-413. [doi: 10.1089/g4h.2018.0148] [Medline: 31241356]
78. Gaetan S, Bréjard V, Bonnet A. Video games in adolescence and emotional functioning: emotion regulation, emotion intensity, emotion expression, and alexithymia. Comput Hum Behav 2016 Aug;61:344-349. [doi: 10.1016/j.chb.2016.03.027]
79. Pine R, Fleming T, McCallum S, Sutcliffe K. The effects of casual videogames on anxiety, depression, stress, and low mood: a systematic review. Games Health J 2020 Aug;9(4):255-264. [doi: 10.1089/g4h.2019.0132] [Medline: 32053021]
80. Alves MLM, Mesquita BS, Morais WS, Leal JC, Satler CE, dos santos Mendes FA. Nintendo wii versus xbox kinect for assisting people with parkinson's disease. Percept Mot Skills 2018 Jun;125(3):546-565. [doi: 10.1177/0031512518769204] [Medline: 29665760]
81. Viana RB, Dankel SJ, Loenneke JP, Gentil P, Vieira CA, Andrade MD, et al. The effects of exergames on anxiety levels: a systematic review and meta-analysis. Scand J Med Sci Sports 2020 Jul;30(7):1100-1116. [doi: 10.1111/sms.13654] [Medline: 32171032]
82. Scholten H, Malmberg M, Lobel A, Engels RC, Granic I. A randomized controlled trial to test the effectiveness of an immersive 3d video game for anxiety prevention among adolescents. PLoS One 2016;11(1):e0147763 [FREE Full text] [doi: 10.1371/journal.pone.0147763] [Medline: 26816292]
83. Schoneveld EA, Malmberg M, Lichtwarck-Aschoff A, Verheijen GP, Engels RC, Granic I. A neurofeedback video game (MindLight) to prevent anxiety in children: a randomized controlled trial. Comput Hum Behav 2016 Oct;63:321-333. [doi: 10.1016/j.chb.2016.05.005]
84. Fish MT, Russoniello CV, O’Brien K. Zombies vs anxiety: an augmentation study of prescribed video game play compared to medication in reducing anxiety symptoms. Simulation Gaming 2018 May 4;49(5):553-566. [doi: 10.1177/1046878118773126]
85. Lee J, Jung H, Lee G, Kim H, Park S, Woo S. Effect of behavioral intervention using smartphone application for preoperative anxiety in pediatric patients. Korean J Anesthesiol 2013 Dec;65(6):508-518. [doi: 10.4097/kjae.2013.65.6.508] [Medline: 24427456]

86. Viana RB, de Lira CA. Exergames as coping strategies for anxiety disorders during the covid-19 quarantine period. Games Health J 2020 Jun;9(3):147-149. [doi: 10.1089/g4h.2020.0060] [Medline: 32375011]
87. Diagnostic and Statistical Manual of Mental Disorders. New York, USA: American Psychiatric Association; 2013.
88. Peckham AD, McHugh RK, Otto MW. A meta-analysis of the magnitude of biased attention in depression. Depress Anxiety 2010 Dec;27(12):1135-1142. [doi: 10.1002/da.20755] [Medline: 21049527]
89. Joormann J, Quinn ME. Cognitive processes and emotion regulation in depression. Depress Anxiety 2014 Apr;31(4):308-315. [doi: 10.1002/da.22264] [Medline: 24668779]
90. Sanchez A, Vazquez C, Marker C, LeMoult J, Joormann J. Attentional disengagement predicts stress recovery in depression: an eye-tracking study. J Abnorm Psychol 2013 May;122(2):303-313. [doi: 10.1037/a0031529] [Medline: 23421524]
91. Joormann J, Gotlib IH. Emotion regulation in depression: relation to cognitive inhibition. Cogn Emot 2010 Feb 1;24(2):281-298 [FREE Full text] [doi: 10.1080/02699930903407948] [Medline: 20300538]
92. Nolen-Hoeksema S. Responses to depression and their effects on the duration of depressive episodes. J Abnorm Psychol 1991 Nov;100(4):569-582. [doi: 10.1037//0021-843x.100.4.569] [Medline: 1757671]
93. Joormann J, Stanton CH. Examining emotion regulation in depression: a review and future directions. Behav Res Ther 2016 Nov;86:35-49. [doi: 10.1016/j.brat.2016.07.007] [Medline: 27492851]
94. Oliver MB, Bowman ND, Woolley JK, Rogers R, Sherrick BI, Chung M. Video games as meaningful entertainment experiences. Psycholo Popular Media Cult 2016 Oct;5(4):390-405. [doi: 10.1037/ppm0000066]
95. Ge L, Yap CW, Ong R, Heng BH. Social isolation, loneliness and their relationships with depressive symptoms: a population-based study. PLoS One 2017 Aug 23;12(8):e0182145 [FREE Full text] [doi: 10.1371/journal.pone.0182145] [Medline: 28832594]
96. McLaren S, Gomez R, Bailey M, van der Horst RK. The association of depression and sense of belonging with suicidal ideation among older adults: applicability of resiliency models. Suicide Life Threat Behav 2007 Feb;37(1):89-102. [doi: 10.1521/suli.2007.37.1.89] [Medline: 17397283]
97. Sargent J, Williams RA, Hagerty B, Lynch-Sauer J, Hoyle K. Sense of belonging as a buffer against depressive symptoms. J Am Psychiatr Nurses Assoc 2016 Sep 2;8(4):120-129. [doi: 10.1067/mpn.2002.127290]
98. Wastler H, Lucksted A, Phalen P, Drapalski A. Internalized stigma, sense of belonging, and suicidal ideation among veterans with serious mental illness. Psychiatr Rehabil J 2020 Jun;43(2):91-96. [doi: 10.1037/prj0000386] [Medline: 31414842]
99. Sweeny K, Rankin K, Cheng X, Hou L, Long F, Meng Y. Flow in the time of COVID-19: findings from China. PLoS ONE 2020 Nov 11:1-12 [FREE Full text] [doi: 10.1371/journal.pone.0242043] [Medline: 33175915]
100. Zhu L. The psychology behind video games during COVID‐19 pandemic: a case study of animal crossing. Human Behav and Emerg Tech 2020 Sep 9;3(1):157-159. [doi: 10.1002/hbe2.221]
101. Zuo M. How Playing Video Games Can Help You Through Quarantine. South China Morning Post. 2020. URL: https://www.scmp.com/lifestyle/health-wellness/article/3110738/minecraft-over-mindfulness-how-playing-video-games-can [accessed 2020-05-01]
102. Zagal J, Deterding S. Definitions of 'Role-Playing Games'. New York, USA: Routledge; 2018.
103. Al Husni Al Keilani M, Delvenne V. Use of video games in a child and adolescent psychiatric unit. Psychiatr Danub 2020 Sep;32(Suppl 1):167-171 [FREE Full text] [Medline: 32890384]
104. Joormann J. Cognitive inhibition and emotion regulation in depression. Curr Dir Psychol Sci 2010 Jun 10;19(3):161-166. [doi: 10.1177/0963721410370293]
105. Kneer J, Elson M, Knapp F. Fight fire with rainbows: the effects of displayed violence, difficulty, and performance in digital games on affect, aggression, and physiological arousal. Comput Hum Behac 2016 Jan;54:142-148. [doi: 10.1016/j.chb.2015.07.034]
106. About Art Therapy. American Art Therapy Association. URL: https://arttherapy.org/about-art-therapy/ [accessed 2020-05-01]
107. Blomdahl C, Gunnarsson BA, Guregård S, Rusner M, Wijk H, Björklund A. Art therapy for patients with depression: expert opinions on its main aspects for clinical practice. J Ment Health 2016 Dec;25(6):527-535. [doi: 10.1080/09638237.2016.1207226] [Medline: 27469205]
108. Vytal K, Cornwell B, Arkin N, Grillon C. Describing the interplay between anxiety and cognition: from impaired performance under low cognitive load to reduced anxiety under high load. Psychophysiology 2012 Jun;49(6):842-852 [FREE Full text] [doi: 10.1111/j.1469-8986.2012.01358.x] [Medline: 22332819]
109. Mogg K, Bradley BP. Anxiety and attention to threat: cognitive mechanisms and treatment with attention bias modification. Behav Res Ther 2016 Dec;87:76-108 [FREE Full text] [doi: 10.1016/j.brat.2016.08.001] [Medline: 27616718]
110. Ohannessian CM. Video game play and anxiety during late adolescence: the moderating effects of gender and social context. J Affect Disord 2018 Jan 15;226:216-219 [FREE Full text] [doi: 10.1016/j.jad.2017.10.009] [Medline: 28992585]
111. Spielberger CD, Gorsuch RL, Lushene TE, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.
112. Horne-Moyer HL, Moyer BH, Messer DC, Messer ES. The use of electronic games in therapy: a review with clinical implications. Curr Psychiatry Rep 2014 Dec;16(12):520 [FREE Full text] [doi: 10.1007/s11920-014-0520-6] [Medline: 25312026]

113. Weerdmeester J, Rooij M, Maciejewski D, Engels R, Granic I. A randomized controlled trial assessing the efficacy of a virtual reality biofeedback video game. DANS EASY 2020:1-5. [doi: 10.17026/dans-zbd-bhbe]
114. Pallavicini F, Pepe A. Virtual reality games and the role of body involvement in enhancing positive emotions and decreasing anxiety: within-subjects pilot study. JMIR Serious Games 2020 Jun 17;8(2):e15635 [FREE Full text] [doi: 10.2196/15635] [Medline: 32554371]
115. 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 [FREE Full text] [doi: 10.1097/HRP.0000000000000138] [Medline: 28475502]
116. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013 Feb;32(2):207-214. [doi: 10.1377/hlthaff.2012.1061] [Medline: 23381511]
117. Werbart A, von Below C, Brun J, Gunnarsdottir H. 'Spinning one's wheels': nonimproved patients view their psychotherapy. Psychother Res 2015 Dec 17;25(5):546-564. [doi: 10.1080/10503307.2014.989291] [Medline: 25517205]
118. Sobolev J. How Many Video Game Exist? Gaming Shift. 2020 Nov 1. URL: https://gamingshift.com/how-many-video-games-exist/ [accessed 2021-06-07]
119. Chanda ML, Levitin DJ. The neurochemistry of music. Trends Cogn Sci 2013 Apr;17(4):179-193. [doi: 10.1016/j.tics.2013.02.007] [Medline: 23541122]
120. Shiranibidabadi S, Mehryar A. Music therapy as an adjunct to standard treatment for obsessive compulsive disorder and co-morbid anxiety and depression: A randomized clinical trial. J Affect Disord 2015 Sep 15;184:13-17. [doi: 10.1016/j.jad.2015.04.011] [Medline: 26066780]
121. Gold C, Solli HP, Krüger V, Lie SA. Dose-response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis. Clin Psychol Rev 2009 Apr;29(3):193-207. [doi: 10.1016/j.cpr.2009.01.001] [Medline: 19269725]
122. Kamioka H, Tsutani K, Yamada M, Park H, Okuizumi H, Tsuruoka K, et al. Effectiveness of music therapy: a summary of systematic reviews based on randomized controlled trials of music interventions. Patient Prefer Adherence 2014;8:727-754 [FREE Full text] [doi: 10.2147/PPA.S61340] [Medline: 24876768]
123. Silverman MJ. Music therapy for coping self-efficacy in an acute mental health setting: a randomized pilot study. Community Ment Health J 2019 May;55(4):615-623. [doi: 10.1007/s10597-018-0319-8] [Medline: 30099647]
124. Alam N, Agung A, Abdullah F, Paliyawan P, Thawonmas R. Singing with an Angry-Birds-like Game. In: Conference on Games. 2020 Presented at: CoG'20; August 20, 2020; Osaka, Japan. [doi: 10.1109/cog47356.2020.9231716]
125. Agres K, Lui S, Herremans D. A Novel Music-based Game With Motion Capture to Support Cognitive and Motor Function in the Elderly. In: Conference on Games. 2019 Presented at: CoG'19; September 26, 2019; London, UK. [doi: 10.1109/cig.2019.8847993]
126. Chao-Fernández R, Gisbert-Caudeli VR, Vázquez-Sánchez R. Emotional training and modification of disruptive behaviors through computer-game-based music therapy in secondary education. Applied Sciences 2020 Mar 5;10(5):1796 [FREE Full text] [doi: 10.3390/app10051796]
127. Unbehaun D, Taugerbeck S, Aal K, Vaziri DD, Lehmann J, Tolmie P, et al. Notes of memories: Fostering social interaction, activity and reminiscence through an interactive music exergame developed for people with dementia and their caregivers. Hum-Comput Interac 2020 Jun 24:1-34. [doi: 10.1080/07370024.2020.1746910]
128. Moormann P. Music and Game: Perspectives on a Popular Alliance. Berlin, Germany: Springer; 2013.

Abbreviations

RCT: randomized controlled trial
RPG: role-playing game
VR: virtual reality

Edited by N Zary; submitted 17.12.20; peer-reviewed by C Ball, R Pine; comments to author 22.02.21; revised version received 18.03.21; accepted 02.04.21; published 16.06.21.

Please cite as:
Kowal M, Conroy E, Ramsbottom N, Smithies T, Toth A, Campbell M
Gaming Your Mental Health: A Narrative Review on Mitigating Symptoms of Depression and Anxiety Using Commercial Video Games
JMIR Serious Games 2021;9(2):e26575
URL: https://games.jmir.org/2021/2/e26575
doi:10.2196/26575
PMID:34132648

Original Paper

Leaderboard Design Principles to Enhance Learning and Motivation in a Gamified Educational Environment: Development Study

Sungjin Park[1*], MEng; Sangkyun Kim[1*], PhD

Department of Industrial Engineering, Kangwon National University, Chuncheon, Republic of Korea

*all authors contributed equally

Corresponding Author:
Sangkyun Kim, PhD
Department of Industrial Engineering
Kangwon National University
417, Engineering Building
Kangwon Deahak Gil 1
Chuncheon, KS007
Republic of Korea
Phone: 82 033 250 6287
Email: saviour@kangwon.ac.kr

Abstract

Background: Gamification in education enhances learners' motivation, problem-solving abilities, decision-making abilities, and social skills such as communication. Numerous ongoing studies are examining the application of gamification design methodology and game mechanics to a learning environment. Leaderboards are a type of game mechanic that assist learners in goal setting and unleash the motivation for learning.

Objective: The aim of this study was to develop leaderboard design principles to assist learners in efficient goal setting, improve learning motivation, and promote learning in gamified learning environments.

Methods: This study implemented 2 different strategies. First, we analyzed previous research on leaderboards that focus on educational efficacy and influence on social interactions. Second, we collected and analyzed data related to cases of leaderboards being used in educational and sport environments.

Results: This study determined 4 leaderboard design objectives from previous studies. Based on these objectives, we developed 3 leaderboard design principles. First, macro leaderboards and micro leaderboards should be designed and used together. Second, all the elements used to measure learners' achievements in an educational environment should be incorporated into the micro leaderboard. Third, leaderboards should be designed and considered for application in contexts other than learning environments. This study further analyzes best practices considering the 3 leaderboard design principles.

Conclusions: This study contributes toward resolving problems associated with leaderboard design for the application of gamification in educational environments. Based upon our results, we strongly suggest that when teachers consider applying gamification in classrooms, the leaderboard design principles suggested in this research should be incorporated.

(JMIR Serious Games 2021;9(2):e14746) doi:10.2196/14746

KEYWORDS

leaderboard design; gamification; learning motivation; affordance

Introduction

Recent advances in technology have caused our methods of communication and our lifestyles to evolve at an unprecedented rate. However, this has not necessarily been the case in the field of education. Although there has been some technology advancement in the classroom, such as the introduction of digital devices to replace books, knowledge is mostly delivered to learners in one direction by the teacher. Currently, researchers are actively studying diverse educational methods such as problem-based learning and learner-centered educational environments [1]. Gamification, which means using gaming elements, structures, and principles in educational innovation, is attracting the attention of teachers and instructors.

Gamification uses game mechanics such as points, badges, levels, or avatars to motivate participants by providing flow and fun experiences while promoting social interaction among participants [2,3].

The application of gamification in the learning context is called "gamification in education" [4]. Al-Azawi et al [5] conducted a comparative analysis of gamification, game-based learning, and educational games, which are mechanics that represent educational innovations based on game thinking. They found that gamification is easier to achieve and more affordable than other methods and that it encourages learners to undertake new challenges without fear of failure. Considering the research of Al-Azawi et al, this study confirmed the trends of gamification, game-based learning, educational games, and serious games by using a Google Trends search (Figure 1). The results showed that gamification has been receiving worldwide attention and popularity since 2012. Park and Kim [6] collected 754 cases of gamification from pre-2010 to 2017, 270 of which were related to education.

Figure 1. Trends of gamification, game-based learning, educational games, and serious games by using the Google Trends search.

Gamification in education is beneficial to learners in several ways. Majuri et al [7] collected 807 studies related to gamification in education from June 2015 and performed an empirical study on 128 relevant studies. Their findings revealed that game mechanics unleash and improve affordance, immersion, and the social behavior of learners. Psychological benefits include improvements in cognitive and emotional ability, stimulation of the desire to be challenged, increased social interaction, projection of psychological states and traits, and the enhancement of personal traits. Behavioral outcomes include improved learning performance, engagement, and physical and social interactions. Kim et al [4] suggested the Lean Canvas-based 4F Process for more efficient gamification design. The Lean Canvas-based 4F Process is a methodology that analyzes the objects of gamification, player type, fun preferences, and game genre preferences to help users set game mechanics and develop prototypes. Meanwhile, Mora et al [8] analyzed 18 studies on gamification design methodology for the period 2011 to 2015 and identified 5 gamification design elements. These elements are as follows:

1. Economic: objectives, risk, return on investment, and stakeholders
2. Logic: loop, end game/epic win, on-boarding, rules
3. Measurement: metrics and analytics
4. Psychology: fun, motivation, socialness, desired behaviors, and ethics
5. Interaction: narrative, user interface/user experience, and technology

Octalysis [9] is a published gamification design framework for the development of gamified learning content that is based on human behavioral characteristics, cognitive structure, and the Gamified Environment and Learning Design model [10]. Game mechanics design is one of the most important elements shared by the published gamification design methodologies. Game mechanics represent the physical and environmental elements connecting users and games [11]. Through game mechanics, affordance is achieved and values nested in the content are transmitted to players [12].

Leaderboards, among other game mechanics, assist users with goal setting, boost competition, and provide feedback. Dicheva et al [13] conducted systematic mapping of 34 studies related to gamification in education published between 2011 and 2014. Their results showed that leaderboards were the most frequently used method—after badges—in the 34 studies. Goal setting is extremely important in educational environments [4]. A leaderboard is a device that guides learners to set specific goals and that represents the outcome in a visible way [7]. However, detailed leaderboard design methods are not addressed in existing gamification design methodology. Kim et al [4] state that a leaderboard should be designed at micro and macro levels but did not mention the structural arrangement of the elements. Despite all the existing research on the effects of leaderboards, their organization has been rarely addressed. This study, therefore, develops a set of leaderboard design principles based on a literature review and the best case analysis of leaderboards in educational gamification.

Methods

Learners' Self-feedback and Goal-Setting Via the Leaderboard

Leaderboards, as game mechanics, induce social behaviors and encourage interaction among participants through competition and cooperation [14]. Learners are self-motivated to check their positions on the leaderboard and stimulated to inform others of their accomplishments [15]. According to O'Donovan et al [16], leaderboards influence learners' motivation more than other game mechanics such as progress bars, end prizes, and awarded badges. However, leaderboards cannot be used alone; they provide information for learners in conjunction with other game mechanics. Teachers assess the performance of the learners and award points, levels, and badges accordingly. Additionally, teachers display who scored how many points and badges and who is at which level on the leaderboard. Learners can see their ranking on the leaderboard, compare themselves with other learners, and receive feedback on areas where they need to improve [17]. Gamification helps learners to immerse themselves in learning with enthusiasm and persistence [18]. As learners enter a state of flow in their learning, they set specific goals based on their level on the leaderboard [19]. McGonigal [20] presents goal setting, voluntary participation, rules, and feedback as game characteristics. These 4 characteristics can be applied to gamification in education. Game players gain a sense of purpose by setting goals. In gamified learning environments, learners receive feedback on their activities through the leaderboard and are rewarded for their achievements. Based on this dynamic, it is possible to induce voluntary participation and observance of classroom rules to leverage the utility of gamification in education [4]. However, a drawback of leaderboards is that they may dampen participant motivation or desire for learning because the mechanics are not applied to learners who are high on the leaderboard. Additionally, those in the lower ranks tend not to respond to the leaderboard and are likely to feel inadequate when comparing their achievements with those of the high-ranked learners [21]. Those in the upper ranks, those with higher than average ranks, or those who feel they are not significantly different from the higher-ranked learners may feel satisfied with their position or motivated to improve through upward counterfactual thinking [22]. Counterfactual thinking [23] is the act of considering past events that did not happen. Individuals with a positive mindset tend to employ upward counterfactual thinking. For example, a learner ranked seventh on the leaderboard may consider the learner ranked third and think, "If I had worked a little harder, I might have ranked third." Conversely, when people have negative experiences, they tend to employ downward counterfactual thinking. For example, a learner ranked 20th on the leaderboard might think, "I dropped this time. I will probably drop further [24]." If this negative experience continues, the individual may lose the confidence and motivation for learning [25]. Based on the literature review, this study suggests the following objectives that should be considered in the design of leaderboards:

Objective 1: Minimize relative deprivation.

Objective 2: Minimize learners' experiences of failure to minimize downward counterfactual thinking.

Objective 3: Maximize learners' experience of success to induce upward counterfactual thinking.

Structure of Leaderboards in a Gamified Learning Environment

Leaderboards can be divided into 2 types: macro leaderboards, which are associated with overall content, and micro leaderboards, which are associated with a subsection of content [4]. Figure 2 displays the player versus player ranking in the World of Warcraft. Figure 3 is a badge count leaderboard for Khan Academy—a gamified learning platform for mathematical education. This figure shows the number of badges granted to the learners who satisfied the criteria specified by the Khan Academy. According to the standards suggested by Kim et al [4], Figure 2 is a macro leaderboard while Figure 3 is a micro leaderboard. Although these leaderboards serve different purposes, the structural characteristics in these leaderboards are similar. The player profiles are provided. The leaderboards in games offer information, including ranking, ID or nickname, organization, occupation in the game, character information, and the player's win rate. The information includes rank, grade, experience, and earned badges. Game players meet and compete in the game and the information is reflected on the leaderboard. The same is true of the gamified learning environment. After conducting learning activities in the classroom, teachers assess the results and update them on the leaderboard. Based on the information registered on the leaderboard, learners maintain motivation for learning, compete with other learners, and set specific goals for themselves. Therefore, it is important that the leaderboard provides specific information. At the same time, the leaderboard should function smoothly to maintain the user-friendly and fun experience of the gamified learning content or environment.

Figure 2. World of Warcraft leaderboard.

US 2v2 Leaderboard

Showing 500 of 4601

Filter Show All

(or scroll down to view more)

Rank	Rating	Name	Realm	Faction	Race	Class	Spec	W - L	Win %
1	2871	Brain	Laughing Skull					477 - 31	93.9%
2	2797	Petkick	Sargeras					974 - 64	93.8%
3	2626	Øríginâl	Bleeding Hollow					139 - 58	70.6%
4	2619	Ceryln	Illidan					226 - 150	60.1%
4	2619	Golrathcr	Thrall					151 - 82	64.8%
6	2604	Zorian	Bleeding Hollow					116 - 57	67.1%
7	2602	Ssds	Sargeras					162 - 83	66.1%
8	2593	Doritotko	Drak'thul					128 - 29	81.5%
9	2591	Smx	Tichondrius					302 - 169	64.1%
10	2561	Noobender	Bleeding Hollow					284 - 153	65.0%
11	2559	Jolyrancher	Bleeding Hollow					250 - 202	55.3%

Figure 3. Badge count leaderboard of Khan Academy.

No.	User	Badges Earned	Active/Inactive
1	cr4k3d.3gg	**27,039**	Inactive
2	Anthon van der Neut	**20,083**	Inactive
3	Kati Susanna	**20,000+**	Active
4	TJ	**15,124**	Active
5	KEVIN	**13,684**	Active
6	Wudaifu	**13,453**	Inactive
7	Raymond Greenwood	**12,453**	Inactive
8	Alois	**11,238**	Active
9	Cyan Wind	**11,211**	Inactive
10	Blaze Runner™	**11,166**	Active
11	Quan Cheng Taian	**10,972**	Inactive
12	trek	**10,892**	Inactive
13	J Helston	**10,750**	Recently Active

Butler [26] examined the relationship between leaderboards and the fun experienced by players and 10 key findings were obtained. Those relevant to this research are as follows:

1. Players who have experienced competition play games at least once more than those who have not.
2. The level of fun experienced by players is not related to the number of plays.
3. Creating games that are too easy may downgrade the fun, depending on the player's opinion.
4. High scores are not necessarily related to fun.
5. When players consider a game too easy, it is difficult to change their opinion; however, alternative definitions can be considered when it is deemed too difficult.

Appropriate competition induces a state of flow. In learning environments, macro leaderboards are used frequently. However, macro leaderboards can only provide numeric data totaling the performance of participants and can therefore only induce fragmented competition. Thus, many presume that frequent competitions between participants cannot be induced by design. Additionally, participants in upper ranks may be exalted by high scores, while those in middle and lower ranks cannot even entertain the thought of competing with those in higher ranks because of the gap in achievement. Therefore, leaderboards that reflect standards or elements that encourage competition in other areas are just as important as leaderboards that promote high scores. Considering the necessity of the following structural conditions in designing leaderboards based on our literature review, this study suggests the following structural purpose of leaderboard design.

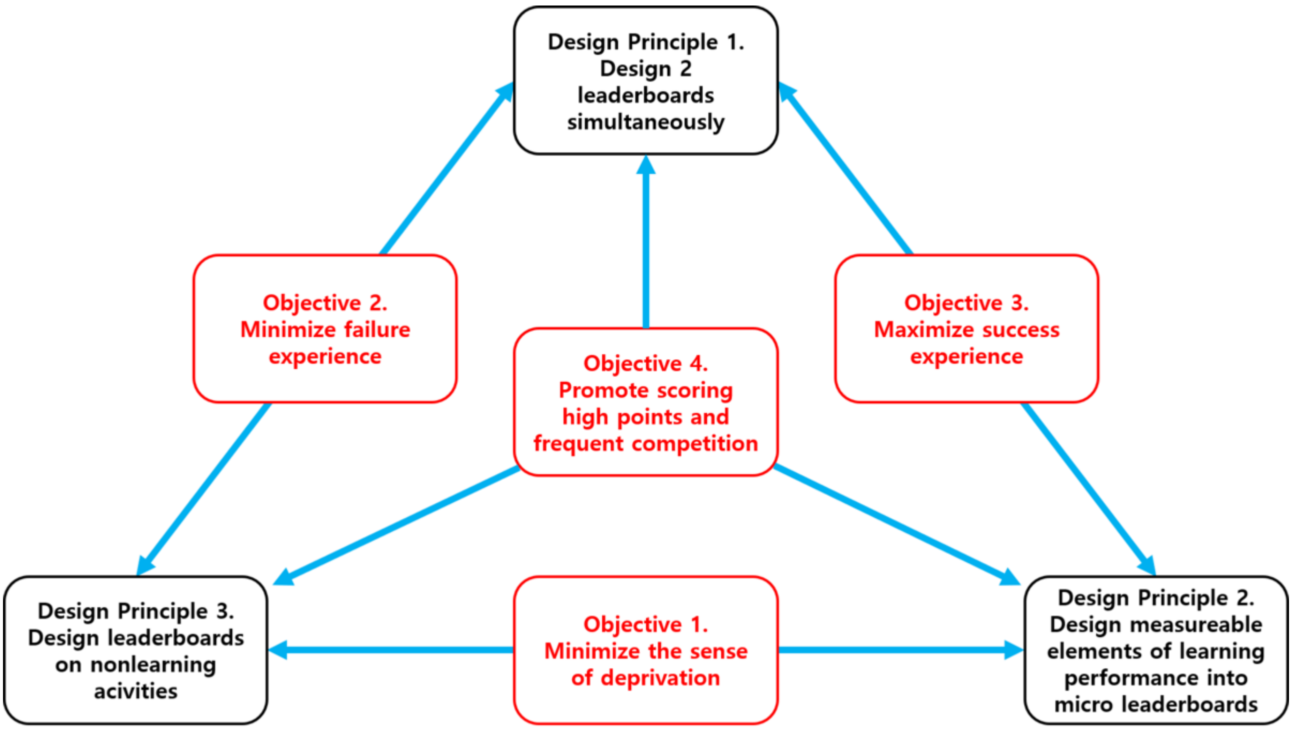
Objective 4: Design leaderboards with measurable learning performance to induce learners to obtain high scores and compete with one another.

Results

Leaderboard Design Principles for Effective Gamified Learning

To achieve the 4 objectives identified from the analysis of prior research, this study suggests 3 design principles, as summarized in Figure 4.

Figure 4. Suggested structure for leaderboard design.



First, both macro leaderboards and micro leaderboards should be designed. If only macro leaderboards are used, participants in the mid and upper rankings can maximize their experience of success. Conversely, those who belong to the lower rankings on a macro leaderboard experience a sense of failure. If they are continuously recorded in the lower rankings on a macro leaderboard, their failure experiences accumulate and negative perceptions and experiences of learning are maximized [27]. To address this problem, learners require feedback to assure them that they are learning properly in some areas. Micro leaderboards fill this need by providing more detailed feedback than macro boards. Using micro leaderboards makes it easier to encourage the scoring of high points and frequent competition and to maximize success while minimizing failure.

Second, all the elements used to measure learners' achievements in an educational environment should be incorporated into the micro leaderboard, as described in design principle 1. This allows participants to take part in learning activities in areas where they have confidence. Thus, learners realize an extended range of movement through leaderboard diversification and are consequently stimulated to achieve more in their learning [28]. Additionally, when the number of leaderboards increases, it is more likely that a learner will be listed in the upper ranks, resulting in increased motivation. The purpose of this design is to maximize success and minimize the sense of deprivation.

Third, leaderboards that reflect activities unrelated to learning performance should be designed to promote gamified learning. The goal of leaderboards is to encourage goal setting and engagement in learning activities. However, learners who are not interested in learning activities are not willing to engage in activities, no matter how much fun they are. This problem can be addressed by creating leaderboards for activities other than learning. For example, leaderboards for classroom tasks such as cleaning, counseling friends, library visits, and books borrowed from the library should be provided. Through leaderboards, participants can set a new and specific set of goals [29] and achieve success by moving up in the leaderboard ranks. In addition, activities other than learning can encourage frequent competition and easier point scoring. These features promote learning engagement, maximize success, and minimize a sense of deprivation.

Design Principle 1: Leaderboards Should be Designed on Both Macro and Micro Levels

Leaderboards frequently used in educational environments are macro leaderboards. In gamified learning environments, learner performance is incorporated into points or experience points that are shown on the leaderboard. Macro learning behavior is learning a theory named A in the gamified environment. The process of learning theory A is composed of microlevel learning behaviors such as A-1, A-2, and A-3 [30]. When using a macro leaderboard, it is possible to provide evaluation and feedback on macro learning behaviors but not on micro learning behaviors. It is therefore difficult to discern what activity a learner excels at from a macro leaderboard. This is because a macro leaderboard provides feedback on overall learning performance but does not give feedback on each learning activity. Kasworm and Blowers [31] suggested considering a variety of personal and environmental factors, as learners' performance is the result of a combination of both factors. It is not easy for teachers to understand personal factors; however, environmental factors can be addressed by designers. If environmental factors are addressed, they can influence participants' learning activities. Therefore, this study concludes that macro and micro leaderboards should be operated simultaneously. This provides learners with more experiences of success. Conversely, as the number of trials increases, the number of failures is expected to rise. Negative influences from increased failure should be controlled by gamification. If teachers can encourage learning activities through gamification [5], they can ensure more success than failure. Enhanced cooperation and more frequent competition between participants will create synergy for the achievement of individual goals [26,32].

Design Principle 2: Integrate Each Measurable Element in the Gamified Learning Environment With a Micro Leaderboard

In the same vein as Kasworm and Blowers [31], teachers who design gamified learning environments should integrate all elements into micro leaderboards. These elements include demographic factors. In learning environments, students are given grades based on their performance during a semester. Various evaluation items—from team activities to final exams—are designed to calculate grades. However, it is difficult for learners to obtain proper feedback because the items displayed on the leaderboard are only representative of activities or points that incorporate all activities. If a leaderboard is designed to incorporate each measurement applied in a gaming environment, learners will attempt to reach the upper ranking of that specific leaderboard. In the macro leaderboard, the greater the gap between the upper and lower ranks, the lower is the satisfaction with learning [33]. Meanwhile, learners' confidence and learning activities are positively influenced by those in lower ranks relative to themselves [34]. As learners participate in more leaderboards where they compete with other learners in terms of performance, they gain confidence and are stimulated to try harder to earn high points. It is also important to provide multiple micro leaderboards to expose learners to objective evaluation and the achievements and competence of their peers. Using leaderboards facilitates the comparison of learning performance [35]. For example, if gamification is applied to a class during a semester and team activity scores, task scores, test scores, and badge acquisition status are elements applied to the leaderboard, other elements such as attendance points, the rankings of specific earned badges gained, the number of questions asked, the number of presentations, and other learning activities should be integrated into micro leaderboards. In the case of a web-based learning platform, all activities related to learning such as the time spent learning, the number of badges earned, or the number of assignments submitted (points) should be integrated with offline learning environment elements. Moreover, if the gender and age of the learner can be checked, leaderboards should be designed according to these factors. Actual web-based games provide a leaderboard based on most elements of the game such as server (region), occupation, and the gender of the game character.

Design Principle 3: Leaderboards Should Incorporate Activities Other Than Learning

The social characteristics of leaderboards have a positive influence on the learning effect through synergy with other game mechanics [35]. Leaderboards that focus on learning effects are effective for learners who rank highly; however, they can cause stress to newcomers or those who rank lower [27]. As a result, participants may have a negative perception of leaderboards. To encourage these disillusioned learners to focus on leaderboards again, a different approach is required. This requires a strategy to induce participants to engage in full-scale learning activities after eliminating the negative perceptions of leaderboards. This is done by introducing nonlearning activities. The leaderboard provides learners with their roles, responsibilities, and feedback on their status. [36]. The leaderboard causes participants to consider the influence of the leaderboard on other participants for activities other than learning. Nonlearning activities can weaken the sense of inadequacy caused by the leaderboard. Additionally, it will become easier for the participant to rank higher on the leaderboard for nonlearning activities than it is for learning-related leaderboards. Participants can maximize their success and find a state of flow through the process. At the same time, the needs of the participants in the mid and lower ranks will be met to minimize the sense of deprivation. Zhao and Tang

[37] conducted an analysis of gamification cases currently in service based on the 8 core drives of octalysis [9]. They showed that one of the sources of motivational affordance was scarcity. Scarcity is explained as the core driver of wanting something simply because it is unattainable. The less scarce the user perceives an object or activity, the more patient the user becomes. Lower-ranked learners may have already given up learning. However, if they experience a higher ranking on nonlearning leaderboards, they may develop a different perspective on learning leaderboards. Therefore, designing micro leaderboards gives participants new learning experiences and motivates them to achieve higher rankings on learning leaderboards.

Analysis of Leaderboard Cases Using the 3 Design Principles

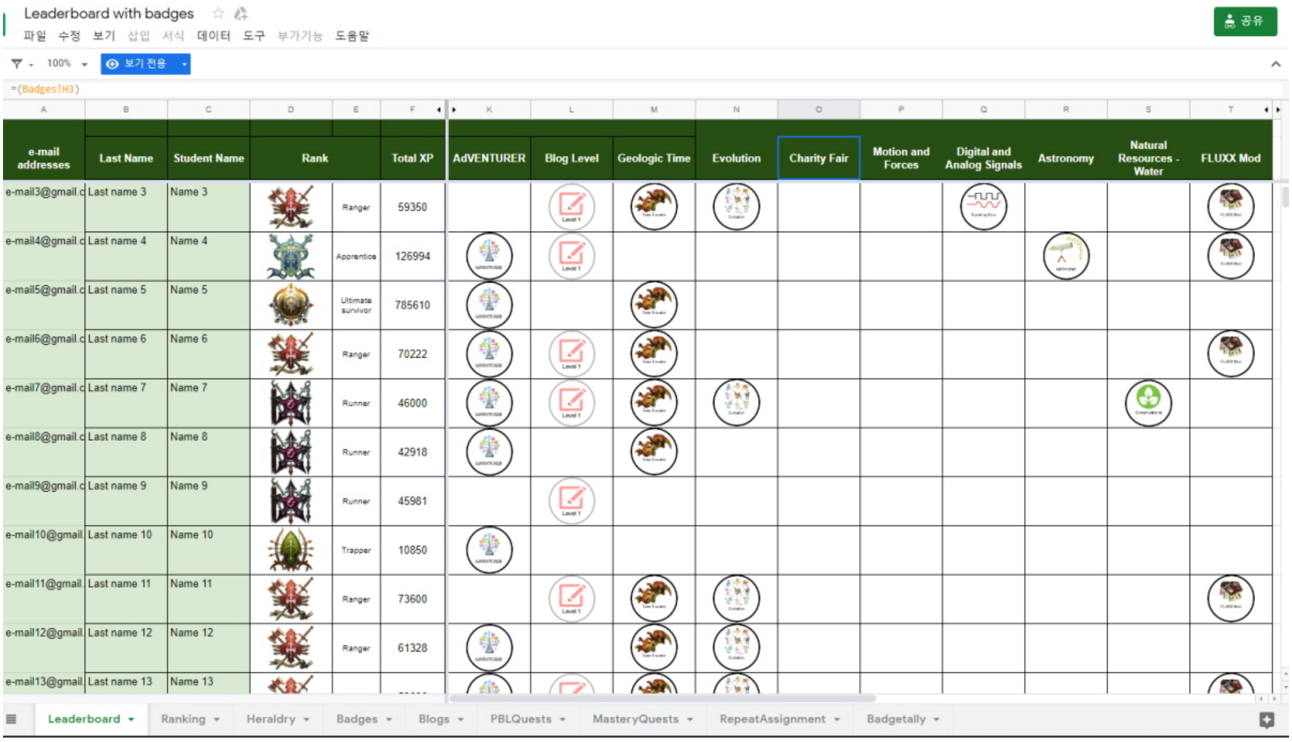
Figure 5 [38-45] shows the leaderboard cases collected for this research. We conducted a Google search to discover leaderboard cases. During our search for leaderboard cases, we aimed to find cases of educational gamification. During the leaderboard case collection stage, it was confirmed that the use of leaderboards in the sports field was frequent. Therefore, we expanded our search range to the sports field. The keywords used for our Google search were as follows: (1) leaderboard case(s), (2) gamified leaderboard case(s), (3) gamification leaderboard case(s), (4) leaderboard in sport, and (5) gamified leaderboard cases in sport. A total of 10 cases were collected and analyzed with the design principles developed in this study. The degree of application of each design principle is represented as ● (completely applied), ◐ (partially applied), and ○ (not applied) in Figure 5.

Figure 5. Collected leaderboard cases. ●: Completely applied; ◐: partially applied; and ○: not applied.

Case	Applied domain	Reference	Design principle 1	Design principle 2	Design principle 3
Duolingo	Education	[39]	●	○	○
FitBit	Health Care	[40]	●	◐	○
Khan Academy	Education	[41]	◐	●	●
Memrise	Education	[42]	●	◐	○
Nike Plus	Health Care		●	◐	◐
Imgur's Pokemon Go	Health Care	[43]	●	●	●
Reebok Crossfit	Health Care	[44]	●	●	●
Science Level Up	Education	[45]	○	○	○
Sololearn	Education		●	◐	○
Task Ville	Human Resource Manag ement		●	●	◐
Teaching Above the test	Education	[46]	●	●	◐

Duolingo, Khan Academy, Memrise, Science Level Up, and Sololearn are web-based learning platforms. Fitbit, NikePlus, Imgur's Pokemon Go, and Reebok Crossfit's "The Open" Challenge are cases related to health care. Pokemon Go is analyzed from the perspective of gamification because it is possible to monitor health through the game [46]. Teaching Above the Test is a leaderboard freely available through Google Sheets in gamified learning environments (Figure 6).

Figure 6. Leaderboard provided by Teaching Above the Test.



According to the analysis, the leaderboards on Reebok Crossfit's "The Open," Khan Academy, and Imgur's Pokemon Go were found to be the most efficient. In contrast, Science Level Up's leaderboard was the least efficient. Imgur's Pokemon Go leaderboard lists all the badges obtainable in Pokemon Go and offers leaderboards according to region with relevant statistics. In addition, leaderboards are renewed in the form of monthly reports and published. The first leaderboard on the page is a macro leaderboard, and micro leaderboards are displayed by content underneath. Design principles 1, 2, and 3 have all been applied to this leaderboard. Figure 7 shows the leaderboard used in Reebok Crossfit's "The Open" challenge. "The Open" is an event that has been ongoing since 2011. Anyone can participate and the event is held around the world simultaneously. The management designs a specific exercise as 1 set. Participants record 1 set during the period set by the management and the result is reflected on the leaderboard [47,48]. The main leaderboard webpage of "The Open" is a macro leaderboard. It shows the total points and ranks for each participant during the entire period. Users can set options to see micro leaderboards. This corresponds to principle 1 developed in this study. It is possible to see micro leaderboards by competition type (open, online qualifier, regionals, sectionals, games, team series, and liftoff), gender, and age group (divided into 5-year age groups from 10 years to 60 years), region, and year. This corresponds to principles 2 and 3 developed in this study. The strategies and physical competence of the players can be analyzed through the leaderboard. The Reebok Crossfit games webpage provides information on the types of exercises for each season. Based on the information, additional information about each player, such as his/her physical competence, characteristics, and strategies, can be obtained to compare with others who are ranked on the leaderboard. Thus, the exercises are designed to enable the leaderboard to be used from various perspectives and to provide the player with the ability to use the leaderboard as a means of improving performance.

Figure 7. Leaderboard provided by Teaching Above the Test.

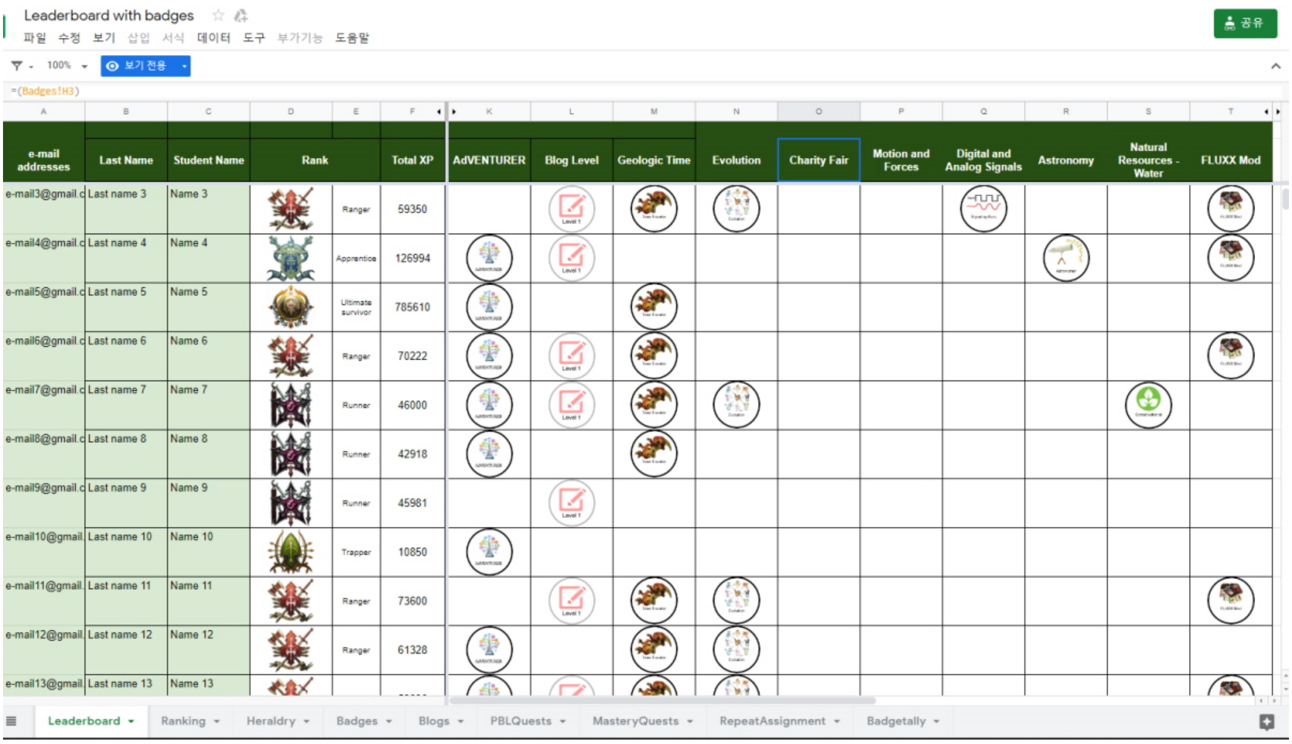
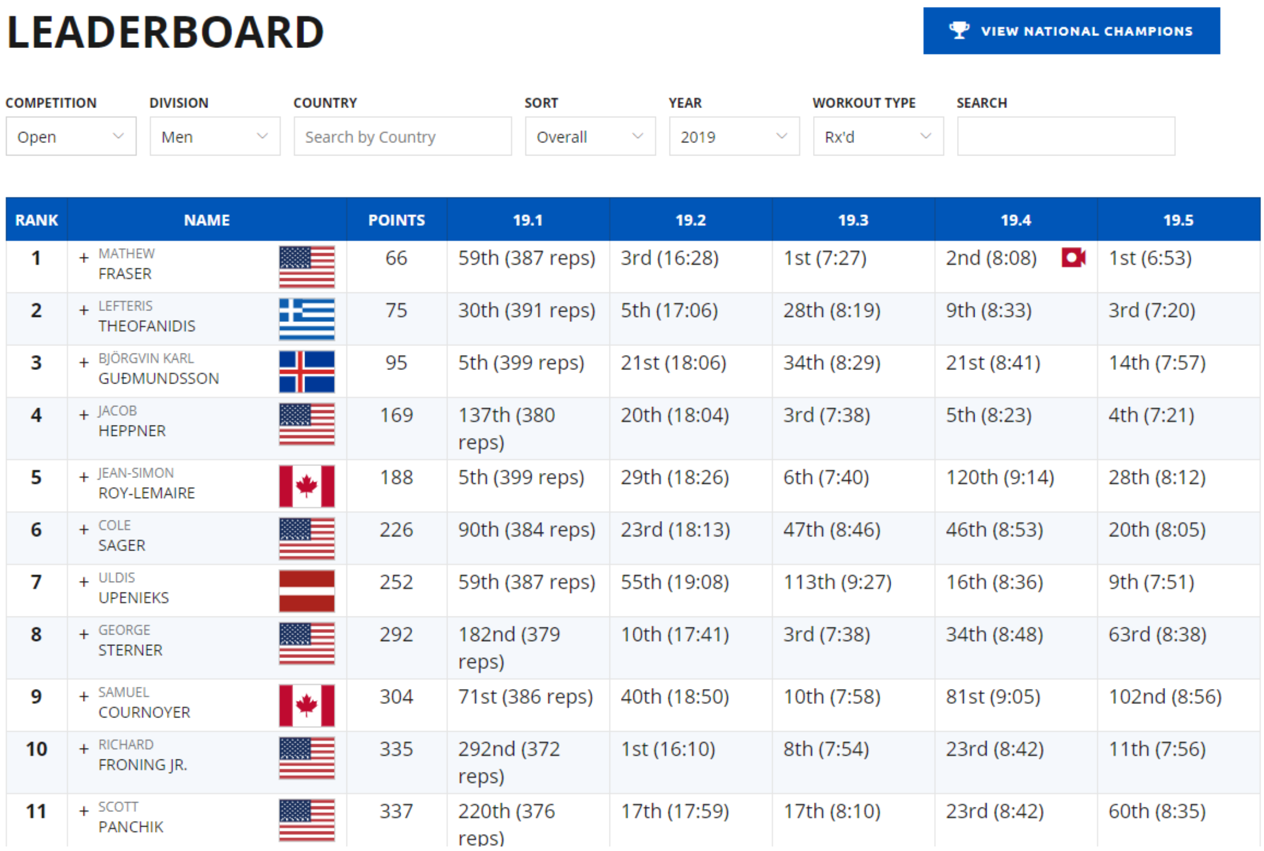


Figure 8 shows the leaderboard of Khan Academy, which offers gamified programming learning content. The leaderboard provides micro leaderboards instead of a macro leaderboard. Leaderboards are provided for each of the elements identifiable as an indicator of learning activity in Khan Academy (challenge patches, energy points, video watching count, badge counts, streak stacks, answer counts, and project evaluation). This corresponds to principles 2 and 3 developed in this study. Web-based learning activities are presumed not to provide leaderboards reflecting precise gender and age groups. When users click on the identification of learners registered on the leaderboard, they can move to each player's profile webpage to see detailed information. Khan Academy's micro leaderboards reflect the number of badges acquired and social interactions such as the number of comments sent to other learners and programming training. This represents the principle 3 developed in this study. In contrast, Science Level Up was found to be the worst leaderboard. A macro leaderboard is not provided and micro boards are only offered by content. Additionally, only 1 micro leaderboard is provided for each content type, without differentiation by grade or demographic characteristics. None of the principles developed in this research are applied.

Figure 8. Leaderboard used in Reebok Crossfit's "The Open" challenge.



LEADERBOARD

VIEW NATIONAL CHAMPIONS

COMPETITION	DIVISION	COUNTRY	SORT	YEAR	WORKOUT TYPE	SEARCH
Open	Men	Search by Country	Overall	2019	Rx'd	

RANK	NAME	POINTS	19.1	19.2	19.3	19.4	19.5
1	+ MATHEW FRASER	66	59th (387 reps)	3rd (16:28)	1st (7:27)	2nd (8:08)	1st (6:53)
2	+ LEFTERIS THEOFANIDIS	75	30th (391 reps)	5th (17:06)	28th (8:19)	9th (8:33)	3rd (7:20)
3	+ BJÖRGVIN KARL GUÐMUNDSSON	95	5th (399 reps)	21st (18:06)	34th (8:29)	21st (8:41)	14th (7:57)
4	+ JACOB HEPPNER	169	137th (380 reps)	20th (18:04)	3rd (7:38)	5th (8:23)	4th (7:21)
5	+ JEAN-SIMON ROY-LEMAIRE	188	5th (399 reps)	29th (18:26)	6th (7:40)	120th (9:14)	28th (8:12)
6	+ COLE SAGER	226	90th (384 reps)	23rd (18:13)	47th (8:46)	46th (8:53)	20th (8:05)
7	+ ULDIS UPENIEKS	252	59th (387 reps)	55th (19:08)	113th (9:27)	16th (8:36)	9th (7:51)
8	+ GEORGE STERNER	292	182nd (379 reps)	10th (17:41)	3rd (7:38)	34th (8:48)	63rd (8:38)
9	+ SAMUEL COURNOYER	304	71st (386 reps)	40th (18:50)	10th (7:58)	81st (9:05)	102nd (8:56)
10	+ RICHARD FRONING JR.	335	292nd (372 reps)	1st (16:10)	8th (7:54)	23rd (8:42)	11th (7:56)
11	+ SCOTT PANCHIK	337	220th (376 reps)	17th (17:59)	17th (8:10)	23rd (8:42)	60th (8:35)

Discussion

This study designed efficient leaderboards to motivate learners in gamified learning environments. Leaderboards are an efficient tool for competition and cooperation and can help learners set specific goals, boost learning motivation, and unleash affordances in the desired direction. However, as the gap between learners in terms of learning performance widens, learning motivation weakens. This study analyzed the negative influence of leaderboards on participants through a literature review and set the following 4 objectives for the design principles.

1. Objective 1: Minimize the sense of inadequacy.
2. Objective 2: Minimize learners' experience of failure to minimize downward counterfactual thinking.
3. Objective 3: Maximize learners' experience of success to induce upward counterfactual thinking.
4. Objective 4: Design leaderboards that measure learning performance to induce learners to obtain high scores and compete with each other frequently.

This study expects the following 3 leaderboard design principles to minimize the negative influence of leaderboards on participants and to improve leaderboard effectiveness. The design principles are as follows:

1. Design principle 1: Macro leaderboard and micro leaderboards should be designed and operated together.
2. Design principle 2: All the elements used to measure learners' achievements in an educational environment should be incorporated into micro leaderboards.
3. Design principle 3: A "geeks leaderboard," a type of micro leaderboard for activities other than learning, should be designed.

The negative influence of leaderboards should be controlled by gamification and teachers should promote affordances to guide learners in the right direction. Among game mechanics, leaderboards that encourage competition and cooperation based on social competence provide direct feedback to learners. Leaderboards should be designed and operated following appropriate design standards. Efficient leaderboard design principles are suggested in this study based on a literature review. In offline classrooms, a few instructors must control many learners and there are many items to manage. However, the introduction of gamification facilitates efficient classroom operation. There are cases of gamification that assist classroom operations such as class craft and class dojo. If leaderboards are designed and operated according to the leaderboard design principles suggested in this research, learning satisfaction and performance are expected to improve. When using leaderboards in web-based learning platforms as well as offline classrooms, this study recommends applying the principles developed in this study. Learners who experience leaderboards in other web-based environments perceive them as content rather than as a tool for competition or ranking [49]. Motivational affordances can be promoted through leaderboards unlike points or levels [50]. Individuals who have experienced leaderboards in other domains accept the competitive environment of the leaderboard in an educational context. We expect that this dynamic will motivate learners more naturally than other game mechanics. Thus, this study suggests that the leaderboard design principles developed here will enhance web-based gamified learning environments. Designers should refer to Reebok Crossfit's "The Open" Challenge, Khan Academy, and Imgur's Pokemon Go leaderboard.

This research has the following limitations. The suggested leaderboard design principles should be applied in education contexts and be verified for effectiveness. Basic leaderboard designs use leaderboards provided by the gamification system or formats shared on the internet. However, gamified leaderboards do not always produce positive effects. Therefore, the leaderboard design principles suggested in this research should be applied to leaderboard design and their effectiveness should be verified by learners. Glynn et al [51] developed the science motivation questionnaire 2 with reliability and validity guaranteed by statistical validation. The questionnaire is composed of intrinsic motivation, career motivation, self-determination, self-efficacy, and grade motivation factors. Future studies should design leaderboards following the leaderboard design principles developed in this study and use the survey tool to analyze the efficiency of the leaderboard design principles of this study in the education field. Furthermore, the leaderboards for use in other fields can be designed based on the results of this study, and the possibility of field expansion will be suggested through future studies.

Acknowledgments

This study was supported by the National Research Foundation of Korea grant and funded by the Korean Government (MSIT No. 2017R1A2B2002798 and MSIT No. 2020R1A2B501001801).

Conflicts of Interest

None declared.

References

1. Lakritz T. How schools and classrooms have changed over the last 80 years. Insider. URL: https://www.thisisinsider.com/old-school-vintage-classroom-photos-evolution-2018-5 [accessed 2019-04-20]
2. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification using game-design elements in non-gaming contexts. ACM Digital Library 2011:2425-2428. [doi: 10.1145/1979742.1979575]
3. Gamification 101: An introduction to the use of game dynamics to influence behavior. 2010. URL: http://jndglobal.com/wp-content/uploads/2011/05/gamification1011.pdf [accessed 2021-04-06]
4. Kim S, Song K, Lockee B, Burton J. Gamification in Learning and Education. Cham: Springer; 2018.
5. Al-Azawi R, Al-Faliti F, Al-Blushi M. Educational Gamification Vs. Game Based Learning: Comparative Study. IJIMT 2016:131-136. [doi: 10.18178/ijimt.2016.7.4.659]
6. Park S, Kim S. Patterns Among 754 Gamification Cases: Content Analysis for Gamification Development. JMIR Serious Games 2018 Dec 06;6(4):e11336 [FREE Full text] [doi: 10.2196/11336] [Medline: 30401666]
7. Majuri J, Koivisto J, Hamari J. Gamification of education and learning: A review of empirical literature. 2018. URL: http://ceur-ws.org/Vol-2186/paper2.pdf [accessed 2021-04-05]
8. Mora A, Riera D, Gonzalez C, Arnedo-Moreno J. A literature review of gamification design frameworks. In: IEEE. 2015 Presented at: 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games); 16-18 Sept; Skovde, Sweden p. 1-8. [doi: 10.1109/vs-games.2015.7295760]
9. Chou Y. Octalysis. Yu-kai Chou: Gamification & Behavioral Design. URL: https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/ [accessed 2019-04-02]
10. Aldemir T, Ata A, Celik B. A systematic design model for gamified learning environments: GELD model. In: Design, Motivation, and Frameworks in Game-Based Learning. Turkey: IGI global; 2019:30-56.
11. Sicart M. Defining game mechanics. Game Studies. 2008. URL: http://gamestudies.org/0802/articles/sicart [accessed 2021-04-05]
12. Hamari J. Game design as marketing: How game mechanics create demand for virtual goods. International Journal of Business Science & Applied Management. 2010. URL: https://www.researchgate.net/publication/26644787_Game_design_as_marketing_How_game_mechanics_create_demand_for_virtual_goods [accessed 2021-03-31]
13. Dicheva D, Dichev C, Agre G, Angelova G. Gamification in education: A systematic mapping study. Educational Technology & Society. 2015. URL: https://www.jstor.org/stable/jeductechsoci.18.3.75 [accessed 2021-04-05]
14. Huang WH, Soman D. Gamification of education. In: Report Series: Behavioural Economics in Action. Toronto, Canada: Rotman School of Management, University of Toronto; 2013:1-29.
15. Nah F, Zeng Q, Telaprolu V, Ayyappa A, Eschenbrenner B. Gamification of education: A review of literature. In: HCI in Business. Cham: Springer; 2014.
16. O'Donovan S, Gain J, Marais P. A case study in the gamification of a university-level games development course. 2013 Presented at: SAICSIT '13: Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference; October; South Africa p. 242-251. [doi: 10.1145/2513456.2513469]

17. Jung JH, Schneider C, Valacich J. Enhancing the Motivational Affordance of Information Systems: The Effects of Real-Time Performance Feedback and Goal Setting in Group Collaboration Environments. Management Science 2010 Apr;56(4):724-742. [doi: 10.1287/mnsc.1090.1129]
18. Nicholson S. A recipe for meaningful gamification. In: Gamification in Education and Business. Cham: Springer; 2015:1-20.
19. Locke EA, Latham GP. New Directions in Goal-Setting Theory. Curr Dir Psychol Sci 2016 Jun 23;15(5):265-268. [doi: 10.1111/j.1467-8721.2006.00449.x]
20. McGonigal J. Reality is broken: Why games make us better and how they can change the world. Westminster, UK: Penguin Press HC; 2011.
21. Walker I, Pettigrew T. Relative deprivation theory: An overview and conceptual critique. British Journal of Social Psychology 1984;23(4):A. [doi: 10.1111/j.2044-8309.1984.tb00645.x]
22. Sun E, Jones B, Traca S, Bos M. Leaderboard position psychology: Counterfactual thinking. In: Conference on Human Factors in Computing Systems. 2015 Presented at: CHI EA '15: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems; April; Japan p. 1217-1222. [doi: 10.1145/2702613.2732732]
23. Roese NJ. Counterfactual thinking. Psychol Bull 1997 Jan;121(1):133-148. [doi: 10.1037/0033-2909.121.1.133] [Medline: 9000895]
24. Roese NJ. The functional basis of counterfactual thinking. Journal of Personality and Social Psychology 1994;66(5):805-818. [doi: 10.1037/0022-3514.66.5.805]
25. Nicholson S. Exploring gamification techniques for classroom management. Games Learning Society. 2013. URL: https://www.researchgate.net/publication/264129658_Exploring_Gamification_Techniques_for_Classroom_Management [accessed 2021-04-05]
26. Butler C. The effect of leaderboard ranking on players' perception of gaming fun. 2013 Presented at: International Conference on Online Communities and Social Computing; 21-26 July; Las Vegas, USA p. 129-136. [doi: 10.1007/978-3-642-39371-6_15]
27. Kumar J. Gamification at work: Designing engaging business software. 2013 Presented at: International Conference of Design, User Experience, and Usability; 15-20 July; Heidelberg, Germany p. 528-537. [doi: 10.1007/978-3-642-39241-2_58]
28. Matallaoui A, Zarnekow R. Introduction to gamification: Foundation and underlying theories. In: Gamification Using Game Elements in Serious Contexts. Berlin, Germany: Springer; 2017:3-18.
29. Mekler E, Bruhlmann F, Opwis F, Tuch A. Do points, levels and leaderboards harm intrinsic motivation? an empirical analysis of common gamification elements. In: Association for Computing Machinery. 2013 Presented at: Gamification '13: Proceedings of the First International Conference on Gameful Design, Research, and Applications; October; Toronto, Canada p. 66-73. [doi: 10.1145/2583008.2583017]
30. Korn O, Funk M, Schmidt A. Design approaches for the gamification of production environments: a study focusing on acceptance. 2015 Presented at: Proceedings of the 8th ACM international conference on Pervasive Technologies Related to Assitive Environment; July; Corfu, Greece p. 1-7. [doi: 10.1145/2769493.2769549]
31. Kasworm C, Blowers S. Adult Undergraduate Students: Patterns of Learning Involvement. In: Final Research Report. Knoxville, US: College of Education, The University of Tennessee; 1994:233-254.
32. Vegt N, Visch V, de RH, Vermeeren A. Designing gamification to guide competitive and coopertavie behavior in teamwork. In: Gamification in Education and Business. Cham: Springer; 2015:513-533.
33. Buunk B, Gibbons F. Toward an enlightenment in social comparison theory. In: Handbook of Social Comparison. Boston, MA: Springer; 2000:487-499.
34. Major B, Testa M, Blysma W. Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. 1991. URL: https://www.researchgate.net/publication/232504594_Responses_to_upward_and_downward_social_comparisons_The_impact_of_esteem-relevance_and_perceived_control [accessed 2021-04-05]
35. Hanus M, Fox J. Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. Computers & Education 2015 Jan;80:152-161. [doi: 10.1016/j.compedu.2014.08.019]
36. Adnan E, Volkan S. The role of gamfication in online learning management system. 2017 Presented at: 7th International Conference of Strategic Research on Social Science and Education; 13-15 October; Turkey p. 239-247.
37. Zhao Y, Tang J. Exploring the motivational affordances of Danmaku wideo sharing websites: evidence from gamification design. 2016 Presented at: International Conference on Human-Computer Interaction; 17-22 July; Toronto, Canada p. 467-479. [doi: 10.1007/978-3-319-39513-5_44]
38. Duolingo. Learn a language for free. Forever. URL: https://www.duolingo.com/ [accessed 2019-04-02]
39. Olivier S. Leaderboard for Fitbit. 2019. URL: https://play.google.com/store/apps/details?id=org.bitbucket.osalasc.fitleaders&hl=en [accessed 2019-04-02]
40. Fandom. Khan Academy WiKi Leaderboard. URL: https://khanacademy.fandom.com/wiki/Energy_Point_Leaderboard [accessed 2019-04-02]
41. Memrise. URL: https://www.memrise.com/home/leaderboard/ [accessed 2019-04-02]
42. Imgur's DTC Pokemon Go Rankings- November 2018. URL: https://imgur.com/a/Bp575SV [accessed 2019-04-02]

43. Reebok. Reebok Crossfit Games 'The Open' Challenge Games 2019. URL: https://games.crossfit.com/leaderboard/open/2019?country_champions=0&division=1&sort=0&scaled=0&page=1 [accessed 2019-04-02]
44. Science Level Up. KOFAC. 2019. URL: https://sciencelevelup.kofac.re.kr/# [accessed 2019-04-02]
45. Leaderboard and badging with Google sheet. Teaching Above the Test. URL: http://www.teachingabovethetest.com/2017/12/leaderboard-and-badging-with-google.html [accessed 2019-04-02]
46. Althoff T, White RW, Horvitz E. Influence of Pokémon Go on Physical Activity: Study and Implications. J Med Internet Res 2016 Dec 06;18(12):e315 [FREE Full text] [doi: 10.2196/jmir.6759] [Medline: 27923778]
47. Crossfit games. Wikipedia. URL: https://en.wikipedia.org/wiki/CrossFit_Games [accessed 2019-04-02]
48. Reebok crossfit games. URL: https://games.crossfit.com/sanctionals [accessed 2019-04-02]
49. Jia Y, Lui Y, Yu X, Voida S. Designing leaderboards for gamification: Perceived differences based on user ranking, application domain, and personality traits. 2017 Presented at: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; May; Denver Colorado USA p. 1949-1960. [doi: 10.1145/3025453.3025826]
50. Werbach K, Hunter D. For the Win: How Game Thinking Can Revolutionize Your Business. Pennsylvania, US: Wharton School Press; 2012.
51. Glynn SM, Brickman P, Armstrong N, Taasoobshirazi G. Science motivation questionnaire II: Validation with science majors and nonscience majors. J Res Sci Teach 2011 Sep 20;48(10):1159-1176. [doi: 10.1002/tea.20442]

Edited by N Zary; submitted 19.05.19; peer-reviewed by L Gutierrez-Puertas, V Márquez-Hernández, K Blondon, P Roman Lopez; comments to author 23.09.20; revised version received 12.10.20; accepted 12.02.21; published 20.04.21.

Please cite as:
Park S, Kim S
Leaderboard Design Principles to Enhance Learning and Motivation in a Gamified Educational Environment: Development Study
JMIR Serious Games 2021;9(2):e14746
URL: https://games.jmir.org/2021/2/e14746
doi:10.2196/14746
PMID:33877049

Original Paper

Leisure and Problem Gaming Behaviors Among Children and Adolescents During School Closures Caused by COVID-19 in Hong Kong: Quantitative Cross-sectional Survey Study

Shimin Zhu[1], PhD; Yanqiong Zhuang[1], MA; Paul Lee[2], PhD; Jessica Chi-Mei Li[1], PhD; Paul W C Wong[3], DPsyc

[1]Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, Hong Kong

[2]Department of Health Sciences, University of Leicester, Leicester, United Kingdom

[3]Department of Social Work and Social Administration, The University of Hong Kong, Pok Fu Lam, Hong Kong

Corresponding Author:
Paul W C Wong, DPsyc
Department of Social Work and Social Administration
The University of Hong Kong
511, JC Teaching Tower
Pok Fu Lam, 000
Hong Kong
Phone: 852 39175029
Fax: 852 28587604
Email: paulw@hku.hk

Abstract

Background: School closures during the COVID-19 pandemic may have exacerbated students' loneliness, addictive gaming behaviors, and poor mental health. These mental health issues confronting young people are of public concern.

Objective: This study aimed to examine the associations between loneliness and gaming addiction behaviors among young people in Hong Kong and to investigate how familial factors, psychological distress, and gender differences moderate these relationships.

Methods: This cross-sectional study was conducted in June 2020 when schools reopened after 6 months of school closures. Participants included 2863 children and adolescents in primary (Grades 4 to 6) and secondary (Grades 7 and 8) schools (female participants: 1502/2863, 52.5%). Chi-square tests, one-way analyses of variance, and independent-samples *t* tests were performed to compare the differences of distribution in gaming addiction behaviors across gender, age, and other sociodemographic characteristics. Multinomial logistic regression analyses were conducted to identify factors that relate to excessive or pathological gaming behaviors separately, in comparison with leisure gaming.

Results: A total of 83.0% (2377/2863) of the participants played video games during the COVID-19 pandemic. The prevalence of excessive and pathological game addiction behaviors was 20.9% (597/2863) and 5.3% (153/2863), respectively. More male students had gaming addiction symptoms than female students. The multinomial logistic regressions showed that feeling lonely was associated with more problematic gaming behaviors, and the association was stronger for older female students. Low socioeconomic status, less parental support and less supervision, and poor mental health were risk factors for gaming addiction behaviors, especially among primary school students.

Conclusions: Loneliness was associated with gaming addiction behaviors; the findings from this study suggested that this association was similar across gender and age groups among young people. Familial support and supervision during school closures can protect young people from developing problematic gaming behaviors. Results of this study have implications for prevention and early intervention on behalf of policy makers and game developers.

(JMIR Serious Games 2021;9(2):e26808) doi:10.2196/26808

KEYWORDS

COVID-19; leisure gaming; excessive gaming; pathological gaming; familial factors; loneliness; COVID-19 lockdown; school closure

Introduction

It is estimated that over 80% of adolescents play video games [1,2]. The prevalence of gaming disorders among adolescents is estimated to be between 1.3% and 19.9% [3], with 8.1% to 10.6% of them involved in excessive gaming and 2.3% to 2.7% involved in pathological gaming [4,5]. Young people are more likely to get involved in pathological gaming than people in other age groups [6]. Social distancing and daily routine changes due to the COVID-19 pandemic may have intensified young people's gaming behaviors. Gaming disorders were found to be associated with poor academic performance [7,8], loneliness [9-11], and a number of health and mental health issues.

Problem gaming behaviors among physically and socially isolated young people during school closures under the pandemic have drawn the concerns of policy makers, researchers, and helping professionals. Elevated loneliness may lead to depression and suicidal ideation, and the chronic impact may last even after the pandemic. According to the self-determination theory, playing video games may provide a sense of autonomy, competence, and relatedness for most people [12]; however, uncontrolled excessive gaming may lead to gaming disorders. Notably, the linear relationship between loneliness and gaming addiction is debatable. Some studies have suggested that problem gaming behaviors were not associated with loneliness [13,14], while others provided support for their association [15-20]. Moreover, gender differences in gaming disorders, with more males having them [3,21], and inconsistent gender differences in perceived loneliness [22-24] have been observed in previous studies. Yet, few studies have examined whether there are sex differences in the association of game addiction with loneliness. As young people stay at home longer during school closures caused by COVID-19, family structure, socioeconomic status (SES) [25-28], and parenting characteristics (ie, parental support and supervision) [29,30] play a strong role in the impact of gaming behaviors on young people's well-being and development. It is clear from the above literature that the association between gaming disorders and a sense of loneliness among young people remains uncertain.

To fill the void, this study aims to examine the prevalence of gaming behaviors during the school closure period in Hong Kong and identify factors that contribute to the association between gaming behaviors and loneliness among young people. We hypothesized that young people who were lonelier during the isolation period were more likely to be addicted to gaming and that there were age and gender differences after adjustment by familial and mental health factors.

Methods

Data Collection and Sample

This study adopted a cross-sectional survey design with a questionnaire that adopted a variety of validated scales from previous studies. Ethical approvals were granted by the Human Subjects Ethics Sub-Committee of the Hong Kong Polytechnic University (No. HSEARS20161222006). This study drew a stratified random sample of adolescent students from four primary schools and 13 secondary schools in Hong Kong. Students in all classes from grades 4 to 6 in primary schools and grades 7 to 8 in secondary schools were approached with an invitation letter detailing the research objectives and possible risks and benefits of participation. Parental consent was obtained and each respondent gave formal consent for his or her participation in the study by signing a consent form. We aimed to have a sample that could yield a margin of error of $0.98/\sqrt{n} = 1.8\%$. In total, 4635 students and their parents were approached and 2863 students filled in the survey, giving a response rate of 61.8%. Students completed the questionnaires in class in the presence of the research assistant and received souvenirs (worth about US $5) after the completion of the survey.

Measures

Gaming Addiction

Gaming behavior was measured using the Chinese children's version of the 7-item Game Addiction Scale (GAS) [4,31]. The GAS was conceptually developed based on the criteria for pathological gaming in the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition). Some items were adapted for young adolescents (eg, "I played a game to forget my real life" was changed to "I have played games and missed planned work") [31]. Responses were measured on a 6-point Likert scale, with scores ranging from 1 (strongly disagree) to 6 (strongly agree); scores of 4 (slightly agree) to 6 (strongly agree) were taken as endorsement of the described addiction symptom. Endorsement of zero to three items was categorized as leisure gaming, endorsement of four to six items was categorized as excessive gaming, and endorsement of seven items was categorized as pathological gaming [32]. The internal consistency of this scale, as measured by Cronbach α, is .87.

Gaming Time and Mode

Participants were asked to report their average hours (ie, none, half hour, 1 hour, 2 hours, 3 hours, 4 hours, 5 hours, and more than 5 hours) spent playing video games each day during school closures. Participants were also asked whether their main mode of electronic gaming was multiplayer or single player [8] and to report their most-used device for gaming: (1) smartphone, tablet, or iPad; (2) computer; or (3) other.

Loneliness

Loneliness was measured by a single-item question: "How often do you feel lonely?" [33]. Responses were based on frequency and ranged from 0 (not at all) to 3 (nearly every day). This is a commonly used, single self-reporting format to measure loneliness in academic research studies [34]. This short question regarding loneliness was asked directly and made it easy to administer and score [35]. Higher scores indicated greater loneliness.

Parental Support

Parental support was measured using the 4-item family subscale of the Multidimensional Scale of Perceived Social Support (MSPSS) [36]. The word "family" in the original MSPSS was replaced with the word "parents" to assess the social support from parents only. An example item was "My parent is willing to help me make decisions." Students responded using a 6-point Likert scale, ranging from 1 (strongly disagree) to 6 (strongly agree). Higher scores indicated higher levels of parental support. The internal consistency of this scale, as measured by Cronbach α, is .91.

Parental Supervision

Parental supervision was measured by three items: (1) "My parents know what I do after school," (2) "My parents know how I spend my pocket money," and (3) "My parents know how I spend my free time." These items were adapted from the Parental Monitoring Scale [37]. Students responded using a 6-point Likert scale, ranging from 1 (strongly disagree) to 6 (strongly agree). Higher scores represented higher levels of parental supervision. The internal consistency of this scale, as measured by Cronbach α, is .81.

Depression

Depression was measured using the Patient Health Questionnaire-9 [38,39]. It consists of nine items for assessing depression symptoms in the past 2 weeks (eg, little interest or pleasure in doing things). Response options, with respect to frequency, range from 0 (not at all) to 3 (nearly every day). Higher scores represents more depressive symptoms. The internal consistency of this scale, as measured by Cronbach α, is .85.

Anxiety

Anxiety was measured using the Generalized Anxiety Disorder-7 scale [40,41], which is a popular tool that screens for, and measures the severity of, generalized anxiety disorder. It consists of seven items that assess general anxiety symptoms in the past 2 weeks (eg, not being able to stop or control worrying). Response options, with respect to frequency, range from 0 (not at all) to 3 (nearly every day). Higher scores represents more anxiety symptoms. The internal consistency of this scale, as measured by Cronbach α, is .92.

Sociodemographic Characteristics

Information collected on the sociodemographic characteristics of the participants included gender, age, grade, family structure (ie, living with both parents, living with father or mother, or living with no parent), and SES. SES was measured by collecting information about parental education, parents having a paid job or not, the number of household appliances, the number of e-learning devices, home internet accessibility, and the family having a car or not. These questions were more tangible for young children to answer in order to collect SES information [42]. Household appliances included televisions, refrigerators, microwaves, laundry washing machines, dishwashers, air conditioners, and audio-visual equipment [43]. The number of household appliances was summed up and this indicator ranged from 1 to 7. E-learning devices included desktop computers, notebook computers, iPads or other tablets, smartphones, and e-book readers. The number of devices was summed up and this indicator ranged from 1 to 5. Home internet accessibility was measured by a single-item question: "Are internet conditions at home satisfying your needs in your daily lives and in learning?"; responses ranged from 0 to 4 [44-46]. Possible responses regarding car ownership by the student's family were 0 (no) and 1 (yes). The four variables above reflected participants' SES; higher values indicated higher family SES.

Data Analysis

Data were entered into SPSS Statistics software, version 24 (IBM Corp), for analyses. First, the correlations and reliabilities for major variables were established. Second, frequencies and proportions for categorical variables or mean and standard deviation for continuous variables were used to describe the characteristics of the participants and to describe excessive or pathological gaming addiction among the participants by different characteristics. Third, chi-square tests, one-way analyses of variance, and independent-samples *t* tests were used to compare the differences of distribution in gaming addictions between genders, age groups (ie, primary or secondary school students), and other sociodemographic characteristics. Missing data were handled by multiple imputation, imputing 10 sets of complete data sets with 25 iterations per imputation. Fourth, multinomial logistic regression analyses were conducted to estimate the odds ratios (ORs) with 95% CIs of different characteristics for participants engaged in excessive or pathological gaming separately from leisure gaming; different levels of adjustment were controlled to examine the robustness of the associations. In the crude model, we examined the association between loneliness and gaming. In Model 1, we included age and gender (ie, male or female). In Model 2, we included family and parental factors of participants, including family structure (ie, living with no parent, living with father or mother, and living with both parents), father's job status (ie, with or without a paid job), mother's job status (ie, with or without a paid job), household appliances, e-learning devices, home internet accessibility, parental support, and parental supervision. In Model 3, we additionally adjusted for covariates of students' mental health: depression and anxiety. We conducted subgroup analyses to examine whether the associations of gaming addiction with loneliness were different among primary school students and secondary school students.

Two subgroup analyses were conducted to examine whether the associations of gaming addiction with loneliness differed by gender and were conducted by three-step multinomial logistic regressions. The regression analyses for the age subgroups were similar to the full sample analysis. In the gender subgroup analyses in the crude model, we examined the association between loneliness and gaming. In Model 1, we adjusted for family and parental factors of participants, including age, family structure, father's job status, mother's job status, household appliances, e-learning devices, home internet accessibility, parental support, and parental supervision. In Model 2, we additionally adjusted for covariates of students' mental health: depression and anxiety.

Multiple imputation was used to minimize the impact of missing data of the sociodemographic factors. These factors were father's and mother's educational level and father's and mother's job status, which had more than 25% missing data due to children indicating they did not know the answers. A total of 10 imputations were generated. The pooled standardized regression coefficients were calculated across the 10-imputation data sets [47].

Results

Overview

Among the 2863 students participating in this study, 181 (6.3%) were in Grade 4, 208 (7.3%) were in Grade 5, 210 (7.3%) were in Grade 6, 1209 (42.2%) were in Grade 7, and 1055 (36.8%) were in Grade 8. The participants' ages ranged from 8 to 17 years (mean 12.6, SD 1.32). The gender ratio of the participants was close to 1:1 (male: 1346/2848, 47.3%). Most participants lived with both parents (2282/2818, 81.0%), and a small proportion of them lived with either their father (109/2818, 3.9%) or their mother (324/2818, 11.5%). Their parents' educational levels were primary school level (father: 112/1723, 6.5%; mother: 122/1766, 6.9%), middle school level (father: 1034/1723, 60.0%; mother: 1075/1766, 60.9%), and university level or above (father: 577/1723, 33.5%; mother: 569/1766, 32.2%). Many students did not know their parents' educational levels and/or did not answer the questions (1140/2863, 39.8% missing data regarding father's educational level; 1097/2863, 38.3% missing data regarding mother's educational level). Around half of their parents had a paid job (father: 1905/2037, 93.5%; mother: 1429/2088, 68.4%); many students did not know their parents' job statuses and did not answer the questions (826/2863, 28.9% missing data regarding father's job status; 775/2863, 27.1% missing data regarding mother's job status). A total of 693 participants out of 2541 (27.3%) indicated that their family had a car. The mean numbers of household appliances and e-learning devices were 4.80 (SD 1.48) and 2.57 (SD 1.00), respectively. The mean response score given by participants for satisfactory home internet access was 3.04 (SD 0.93).

The Gaming Behaviors of Children and Adolescents

A majority of students in the study (2377/2863, 83.0%) admitted that they played video games every day during the school closure period for 30 minutes (305/2863, 10.7%), 1 hour (340/2863, 11.9%), 2 hours (451/2863, 15.8%), 3 hours (516/2863, 18.0%), 4 hours (376/2863, 13.1%), 5 hours (258/2863, 9.0%), and for more than 5 hours (131/2863, 4.6%). Most of them (1927/2863, 67.3%) were involved in multiplayer games while the rest (582/2863, 20.3%) played individually. Many of them used smartphones, tablets, or iPads (1996/2863, 69.7%); computers (274/2863, 9.6%); and other devices (108/2863, 3.8%) for game playing.

The findings of the GAS are shown in Table 1. In this study, 266 participants out of 2863 (9.3%) reported that they did not play any games during the school closure period, whereas 1432 students (55.2%) reported to have played games longer than intended. The most frequently reported gaming addiction symptom was *tolerance* (1432/2595, 55.2%), which was followed by *preoccupation* (1189/2597, 45.8%), *escape* (913/2587, 35.3%), and *unsuccessful attempts to stop or reduce playing* (885/2588, 34.2%). Male participants were more likely to report gaming addiction symptoms than their female counterparts, except on the third item, "I have played games and missed planned work."

Overall, 2518 out of 2863 (87.9%) participants reported having at least one gaming addiction symptom, with 1768 (61.8%) students playing games as a leisure activity (1 to 3 symptoms), 597 (20.9%) students reporting excessive gaming addiction behaviors (4 to 6 symptoms), and 153 (5.3%) students reporting pathological gaming behaviors (7 symptoms) (Table 2; for more detailed information about participant characteristics, see Table S1 in Multimedia Appendix 1). Male students showed a significantly higher prevalence of excessive gaming and pathological gaming behaviors than females, while more females reported no gaming behaviors (χ^2_3=103.1; P<.001). Primary school students showed a significantly higher prevalence of excessive gaming behaviors than secondary school students. Secondary school students showed a significantly higher prevalence of pathological gaming behaviors than primary school students (χ^2_3=8.5; P=.04). Those who reported excessive gaming behaviors were significantly younger than those who reported leisure and pathological gaming behaviors (F_3=4.03; P=.007).

With respect to the influence of familial and mental health factors, students reporting pathological gaming behaviors had lower home internet access rates, less parental support and supervision, and higher loneliness, depression, and anxiety levels than students reporting leisure and excessive gaming behaviors (all had P values <.05).

Table 1. Game Addiction Scale responses and DSM-5[a] criteria.

Statement	Gave valid response (N=2863), n (%)	Disagree[b], n (%)	Agree[c], n (%)	DSM-5 criterion
I have thought all day long about playing a game.				Preoccupation
All participants	2597 (90.7)	1408 (54.2)	1189 (45.8)	
Male	1314/2583 (50.9)	620 (47.2)	694 (52.8)	
Female	1269/2583 (49.1)	783 (61.7)	486 (38.3)	
I have played longer than intended.				Tolerance
All participants	2595 (90.6)	1163 (44.8)	1432 (55.2)	
Male	1313/2581 (50.9)	547 (41.7)	766 (58.3)	
Female	1268/2581 (49.1)	612 (48.3)	656 (51.7)	
I have played games and missed planned work.				Escape
All participants	2587 (90.4)	1674 (64.7)	913 (35.3)	
Male	1306/2573 (50.8)	848 (64.9)	458 (35.1)	
Female	1267/2573 (49.2)	820 (64.7)	447 (35.3)	
Others have unsuccessfully tried to reduce my time spent on games.				Unsuccessful attempts to stop or reduce playing
All participants	2588 (90.4)	1703 (65.8)	885 (34.2)	
Male	1307/2574 (50.8)	823 (63.0)	484 (37.0)	
Female	1267/2574 (49.2)	870 (68.7)	397 (31.3)	
I have felt upset when I was unable to play.				Withdrawal
All participants	2586 (90.3)	1976 (76.4)	610 (23.6)	
Male	1304/2573 (50.7)	911 (69.9)	393 (30.1)	
Female	1269/2573 (49.3)	1055 (83.1)	214 (16.9)	
I have had arguments with others (eg, family and friends) over my time spent on games.				Deceiving others
All participants	2575 (89.9)	1913 (74.3)	662 (25.7)	
Male	1303/2561 (50.9)	896 (68.8)	407 (31.2)	
Female	1258/2561 (49.1)	1006 (80.0)	252 (20.0)	
I have neglected important activities (eg, school, work, and sports) to play games.				Loss of interest
All participants	2583 (90.2)	1958 (75.8)	625 (24.2)	
Male	1304/2569 (50.8)	975 (74.8)	329 (25.2)	
Female	1265/2569 (49.2)	972 (76.8)	293 (23.2)	

[a]DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

[b]Answers of *disagree* were taken to mean there were no addiction symptoms.

[c]Answers of *agree* were taken to mean there were addiction symptoms.

Table 2. Participant characteristics in different gaming groups.

Variable	Gaming behavior (N=2863)					*P* value
	None	Leisure	Excessive	Pathological	Overall	
Gender[a], n (%)						<.001
Male (n=1346)	85 (6.3)	833 (61.9)	332 (24.7)	96 (7.1)	1346 (100)	
Female (n=1502)	258 (17.2)	927 (61.7)	261 (17.4)	56 (3.7)	1502 (100)	
Total (n=2848)	343 (12.0)	1760 (61.8)	593 (20.8)	152 (5.3)	2848 (100)	
School group[a], n (%)						.04
Primary (n=599)	62 (10.4)	368 (61.4)	145 (24.2)	24 (4.0)	599 (100)	
Secondary (n=2264)	283 (12.5)	1400 (61.8)	452 (20.0)	129 (5.7)	2264 (100)	
Total (N=2863)	345 (12.1)	1768 (61.8)	597 (20.9)	153 (5.3)	2863 (100)	
Live with parent[a], n (%)						.003
Both parents (n=2282)	283 (12.4)	1436 (62.9)	457 (20.0)	106 (4.6)	2282 (100)	
Father (n=109)	7 (6.4)	62 (56.9)	30 (27.5)	10 (9.2)	109 (100)	
Mother (n=324)	38 (11.7)	180 (55.6)	77 (23.8)	29 (9.0)	324 (100)	
Neither parent (n=103)	11 (10.7)	70 (68.0)	18 (17.5)	4 (3.9)	103 (100)	
Total (n=2818)	339 (12.0)	1748 (62.0)	582 (20.7)	149 (5.3)	2818 (100)	
Father's education level[a], n (%)						<.001
Primary school (n=112)	13 (11.6)	67 (59.8)	24 (21.4)	8 (7.1)	112 (100)	
Middle school (n=1034)	106 (10.3)	652 (63.1)	221 (21.4)	55 (5.3)	1034 (100)	
University or above (n=577)	114 (19.8)	358 (62.0)	83 (14.4)	22 (3.8)	577 (100)	
Total (n=1723)	233 (13.5)	1077 (62.5)	328 (19.0)	85 (4.9)	1723 (100)	
Mother's education level[a], n (%)						<.001
Primary school (n=122)	17 (13.9)	71 (58.2)	27 (22.1)	7 (5.7)	122 (100)	
Middle school (n=1075)	124 (11.5)	658 (61.2)	229 (21.3)	64 (6.0)	1075 (100)	
University or above (n=569)	93 (16.3)	379 (66.6)	75 (13.2)	22 (3.9)	569 (100)	
Total (n=1766)	234 (13.3)	1108 (62.7)	331 (18.7)	93 (5.3)	1766 (100)	
Father has a paid job[a], n (%)						.32
No (n=132)	17 (12.9)	75 (56.8)	31 (23.5)	9 (6.8)	132 (100)	
Yes (n=1905)	238 (12.5)	1207 (63.4)	379 (19.9)	81 (4.3)	1905 (100)	
Total (n=2037)	255 (12.5)	1282 (62.9)	410 (20.1)	90 (4.4)	2037 (100)	
Mother has a paid job[a], n (%)						.26
No (n=659)	98 (14.9)	397 (60.2)	135 (20.5)	29 (4.4)	659 (100)	
Yes (n=1429)	170 (11.9)	905 (63.3)	286 (20.0)	68 (4.8)	1429 (100)	
Total (n=2088)	268 (12.8)	1302 (62.4)	421 (20.2)	97 (4.6)	2088 (100)	
Family owns a car[a], n (%)						.41
No (n=1848)	240 (13.0)	1128 (61.0)	391 (21.2)	89 (4.8)	1848 (100)	
Yes (n=693)	81 (11.7)	446 (64.4)	131 (18.9)	35 (5.1)	693 (100)	
Total (n=2541)	321 (12.6)	1574 (61.9)	522 (20.5)	124 (4.9)	2541 (100)	
Number of household appliances[b], mean (SD)	4.90 (1.34)	4.82 (1.48)	4.71 (1.51)	4.59 (1.61)	4.80 (1.48)	.07
Number of e-learning devices[b], mean (SD)	2.60 (1.00)	2.64 (1.01)	2.43 (0.93)	2.33 (1.02)	2.57 (1.00)	<.001

Variable	Gaming behavior (N=2863)					*P* value
	None	Leisure	Excessive	Pathological	Overall	
Home internet access satisfaction score[b], mean (SD)	3.05 (0.99)	3.10 (0.89)	2.93 (0.97)	2.76 (1.01)	3.04 (0.93)	<.001
Age (years)[b], mean (SD)	12.64 (1.31)	12.63 (1.30)	12.45 (1.39)	12.78 (1.23)	12.61 (1.32)	.01
Parental support (MSPSS[c] family subscale score)[b], mean (SD)	4.00 (1.20)	3.94 (1.27)	3.79 (1.20)	3.63 (1.13)	3.90 (1.24)	.002
Parental supervision score[b], mean (SD)	4.36 (1.23)	4.22 (1.25)	4.09 (1.24)	3.87 (1.12)	4.19 (1.24)	<.001
Loneliness score[b], mean (SD)	0.53 (0.92)	0.47 (0.84)	0.60 (0.88)	0.75 (1.07)	0.52 (0.87)	<.001
Depression (PHQ-9[d] score)[b], mean (SD)	5.52 (5.19)	4.75 (4.84)	6.73 (5.29)	8.63 (6.51)	5.44 (5.18)	<.001
Anxiety (GAD-7[e] scale score)[b], mean (SD)	4.18 (4.81)	3.38 (4.66)	4.76 (5.16)	6.15 (6.29)	3.91 (4.94)	<.001

[a]Chi-square tests were performed to compare differences in game addiction with respect to different characteristics.

[b]One-way analyses of variance were performed to compare differences in game addiction with respect to different characteristics.

[c]MSPSS: Multidimensional Scale of Perceived Social Support.

[d]PHQ-9: Patient Health Questionnaire-9.

[e]GAD-7: Generalized Anxiety Disorder-7.

Loneliness and Gaming Addiction

In the previous 2 weeks, 176 out of 2863 participants (6.1%) felt lonely almost every day, 206 participants (7.2%) felt lonely nearly half of the days, and 538 participants (18.8%) felt lonely several days. Among those who reported loneliness, 58.4% (536/918) were females (χ^2_3=21.4; *P*<.001). A significant difference was found between genders (t_{2842}=4.40; *P*<.001) but not between age groups.

Loneliness was positively associated with both excessive and pathological gaming behaviors. After adjusting for age and gender, the adjusted ORs (aORs) of loneliness increased to 1.23 (95% CI 1.19-1.26) and 1.45 (95% CI 1.38-1.53), respectively, for excessive (Table 3) and pathological gaming behaviors (Table 4). Compared to leisure gaming, being lonely increased the likelihood of excessive and pathological gaming behaviors by 23% and 45%, respectively. The aORs did not change substantially from Model 1 to Model 2, which adjusted for parenting. However, after adjusting for depression and anxiety variables, the aORs of loneliness decreased substantially, from 1.19 (95% CI 1.15-1.23) to 0.91 (95% CI 0.87-0.95) for excessive gaming behaviors and from 1.39 (95% CI 1.33-1.47) to 0.89 (95% CI 0.82-0.96) for pathological gaming behaviors. Significant associations between loneliness and gaming disorders were also found across different age and sex groups (Tables S2-S5 in Multimedia Appendix 1). Generally, being lonely was associated with a higher risk of excessive and pathological gaming behaviors, except for primary school students where there was a higher risk of excessive gaming behaviors. The magnitude of the association was relatively higher for pathological gaming behaviors, especially for girls and primary school students.

Table 3. Multinomial logistic regression of excessive gaming addiction (N=2863).

Variable	Odds ratio (95% CI)[a]			
	Crude model[b]	Model 1[c]	Model 2[d]	Model 3[e]
Loneliness	1.19*** (1.15-1.23)	1.23*** (1.19-1.26)	1.15*** (1.11-1.19)	0.91*** (0.87-0.95)
Age	N/A[f]	0.91*** (0.89-0.92)	0.90*** (0.88-0.92)	0.89*** (0.87-0.91)
Gender				
Male	N/A	1.48*** (1.39-1.56)	1.43*** (1.35-1.52)	1.67*** (1.57-1.78)
Female	N/A	Reference	Reference	Reference
Live with parent				
Neither parent	N/A	N/A	0.72*** (0.61-0.85)	0.74** (0.62-0.89)
Either father or mother	N/A	N/A	1.27*** (1.17-1.38)	1.24*** (1.14-1.35)
Both father and mother	N/A	N/A	Reference	Reference
Father has a paid job				
No	N/A	N/A	1.19** (1.07-1.33)	1.19** (1.05-1.34)
Yes	N/A	N/A	Reference	Reference
Mother has a paid job				
No	N/A	N/A	1.06 (1.00-1.13)	1.10** (1.03-1.18)
Yes	N/A	N/A	Reference	Reference
Household appliances	N/A	N/A	1.01 (0.99-1.04)	1.01 (0.99-1.03)
e-Learning devices	N/A	N/A	0.84*** (0.81-0.86)	0.86*** (0.83-0.89)
Home internet access	N/A	N/A	0.87*** (0.85-0.90)	0.87*** (0.84-0.90)
Parental support	N/A	N/A	0.96** (0.93-0.98)	1.01 (0.98-1.04)
Parental supervision	N/A	N/A	0.99 (0.97-1.02)	0.98 (0.96-1.01)
Depression	N/A	N/A	N/A	1.09*** (1.08-1.11)
Anxiety	N/A	N/A	N/A	1.00 (0.99-1.01)

[a]The odds ratios and 95% CIs for excessive and pathological gaming addiction behaviors were calculated separately, with leisure gaming as the reference sample.

[b]In the crude model, we examined the association between loneliness (continual) and gaming.

[c]In Model 1, we additionally adjusted for age and gender.

[d]In Model 2, we additionally adjusted for family structure, the father's job status, the mother's job status, household appliances (continual), e-learning devices (continual), home internet access (continual), parental support (continual), and parental supervision (continual).

[e]In Model 3, we additionally adjusted for depression (continual) and anxiety (continual).

[f]N/A: not applicable; this variable was adjusted for in subsequent models.

**P<.01.

***P<.001.

Table 4. Multinomial logistic regression of pathological gaming addiction (N=2863).

Variable	Odds ratio (95% CI)[a]			
	Crude model[b]	Model 1[c]	Model 2[d]	Model 3[e]
Loneliness	1.39*** (1.33-1.47)	1.45*** (1.38-1.53)	1.35*** (1.28-1.43)	0.89** (0.82-0.96)
Age	N/A[f]	1.07*** (1.04-1.11)	1.07*** (1.03-1.11)	1.04 (1.00-1.08)
Gender				
Male	N/A	2.09*** (1.88-2.32)	1.90*** (1.71-2.12)	2.37*** (2.09-2.69)
Female	N/A	Reference	Reference	Reference
Live with parent				
Neither parent	N/A	N/A	0.60** (0.43-0.82)	0.69* (0.49-0.96)
Either father or mother	N/A	N/A	1.68*** (1.48-1.91)	1.85*** (1.61-2.13)
Both father and mother	N/A	N/A	Reference	Reference
Father has a paid job				
No	N/A	N/A	1.21* (1.01-1.46)	1.33** (1.09-1.62)
Yes	N/A	N/A	Reference	Reference
Mother has a paid job				
No	N/A	N/A	0.98 (0.87-1.10)	1.01 (0.89-1.14)
Yes	N/A	N/A	Reference	Reference
Household appliances	N/A	N/A	1.00 (0.96-1.03)	0.97 (0.93-1.01)
e-Learning devices	N/A	N/A	0.76*** (0.71-0.80)	0.78*** (0.73-0.83)
Home internet access	N/A	N/A	0.80*** (0.76-0.84)	0.79*** (0.75-0.84)
Parental support	N/A	N/A	0.97 (0.92-1.02)	1.05 (0.99-1.11)
Parental supervision	N/A	N/A	0.94** (0.89-0.98)	0.97 (0.92-1.02)
Depression	N/A	N/A	N/A	1.14*** (1.12-1.16)
Anxiety	N/A	N/A	N/A	1.02 (1.00-1.04)

[a]The odds ratios and 95% CIs for excessive and pathological gaming addiction behaviors were calculated separately, with leisure gaming as the reference sample.

[b]In the crude model, we examined the association between loneliness (continual) and gaming.

[c]In Model 1, we additionally adjusted for age and gender.

[d]In Model 2, we additionally adjusted for family structure, the father's job status, the mother's job status, household appliances (continual), e-learning devices (continual), home internet access (continual), parental support (continual), and parental supervision (continual).

[e]In Model 3, we additionally adjusted for depression (continual) and anxiety (continual).

[f]N/A: not applicable; this variable was adjusted for in subsequent models.

*P<.05.

**P<.01.

***P<.001.

Familial Factors and Gaming Addiction

Family structure and socioeconomic factors were found to be significantly associated with gaming behaviors. Students from single-parent families (aOR 1.85, 95% CI 1.61-2.13; *P*<.001) and with unemployed fathers (aOR 1.33, 95% CI 1.09-1.62; *P*=.006) were at a higher risk of developing pathological gaming behaviors (Table 4). Higher SES reflected by more e-learning devices (aOR 0.78, 95% CI 0.73-0.83; *P*<.001) and satisfactory internet access (aOR 0.79, 95% CI 0.75-0.84; *P*<.001) were associated with fewer gaming addiction symptoms. The magnitude of the associations was higher among primary school students compared to secondary school students (Tables S2 and S3 in Multimedia Appendix 1).

Parental support was a significant protective factor for excessive (aOR 0.87, 95% CI 0.82-0.92; *P*<.001) and pathological (aOR 0.82, 95% CI 0.73-0.93; *P*=.002) gaming behaviors only among primary school students (Table S2 in Multimedia Appendix 1) but not among secondary school students. Higher perceived parental support, however, was associated with more excessive and pathological gaming behaviors among secondary school students after adjusting for depression and anxiety (*P*<.001) (Table S3 in Multimedia Appendix 1).

Perceived parental supervision was not significantly associated with gaming disorders among the whole sample, except for pathological gaming behaviors in Model 2 (aOR 0.94, 95% CI 0.89-0.98; *P*=.007) (Table 4). However, in the stratified analysis, the significant protective effect of higher parental supervision for pathological gaming behaviors was found among secondary school students (aOR 0.91, 95% CI 0.86-0.96; *P*=.001), especially for female students (aOR 0.81, 95% CI 0.75-0.89; *P*<.001) (Tables S3 and S5 in Multimedia Appendix 1), but not among primary school students. In terms of gender differences, parental supervision was a significant protective factor for excessive gaming behaviors among male students only (Table S4 in Multimedia Appendix 1). On the contrary, parental supervision was a significant risk factor for excessive gaming behaviors among female students only (Table S5 in Multimedia Appendix 1).

Discussion

Principal Findings

It is alarming that excessive and potential pathological gaming behaviors were prevalent among adolescent students in Hong Kong when schools were closed owing to the COVID-19 pandemic. Female students in this study were more likely to report loneliness than male students, while those with excessive and pathological gaming behaviors reported more loneliness. The results showed that loneliness was positively associated with gaming addiction when adjusted for sociodemographic characteristics, but the association became negative when depression and anxiety were present. It is noteworthy that familial factors played protective roles in the development of excessive and pathological gaming behaviors only for the younger students.

A recent systematic review on addictive gaming behaviors suggests that the prevalence rates for excessive gaming range from 1.2% to 8.1% and those for pathological gaming range from 1.4% to 2.7% [3]. Empirical studies on gaming behaviors among young people in Hong Kong have been scant. One study conducted in 2014 concluded that 93.2% of 503 high school students had played video or internet games, and 15.7% had a gaming addiction [48]. Our findings indicate that the rates of excessive and pathological gaming behaviors among students are much higher than the local and international rates, and they seem to suggest that there are positive relationships between school closures during the COVID-19 pandemic and addictive gaming behaviors among adolescents. The COVID-19 pandemic definitely disrupted students' daily routines, limited their outdoor activities, and, hence, easily enabled students to play games for leisure to reduce loneliness and to connect with others [49]. This study has affirmed the positive relationship between loneliness and gaming behaviors that was concluded by previous studies [16,17] and highlighted the situation during a pandemic when many schools are closed down accordingly. However, the conclusion that excessive gaming behaviors increased among students in Hong Kong was not substantiated as we did not have data collected from before the outbreak of COVID-19 for comparison. Nevertheless, when comparing with the findings from previous studies based on school students in Hong Kong, the prevalence of excessive gaming behaviors among the sample of this study is prominent.

It is interesting to find that although male students in this study exhibited significantly higher game addiction behaviors than their female counterparts [7,50], female secondary students were more likely to have a higher risk of pathological gaming behaviors. It is speculated that older female students are more likely to have a desire for closeness with others and dependency on peer relationships, worrying more about abandonment, loneliness, and loss of relationships [51,52] during the school closure period. They were, therefore, more likely to use online gaming as a coping strategy for maintaining social connection in dealing with adverse emotional experiences [53-57] during the isolation period. Truly, any development of prevention and intervention for addictive gaming behaviors must take the relationship between gaming and gender into account [56].

Regarding the mental health of students with gaming behaviors, students with excessive and pathological gaming behaviors were more likely to be depressed, anxious, and lonely [17,58]. However, when these variables were analyzed within multilevel models, loneliness became negatively associated with gaming addiction behaviors. We suspect that feeling depressed may decrease the motivation for socializing and decrease the motivation for gaming among both male and female students [59,60].

Familial factors play a crucial role in the association between loneliness and gaming addiction behaviors, especially for primary school students. It was found that students from single-parent families, with unemployed fathers, or with lower SES were more likely to exhibit gaming addiction behaviors. Considering how SES relates to gaming behaviors, we suspect that those with lower SES had fewer options for leisure activities (eg, music, cooking, and traveling). For primary school students but not secondary students, if their mothers had no paid job and stayed at home, they may have had more parental support and supervision that led to a lesser chance for developing gaming addiction behaviors. Younger children who are supervised and supported by their mothers have been found to be less vulnerable to loneliness and problem gaming behaviors [29,30]. In other words, age also plays an important role in the development of gaming addiction behaviors.

Strengths and Limitations

To our knowledge, this is the first study that examines the association between loneliness and gaming addiction behaviors among young people in Hong Kong during the pandemic. Stratified analysis of age and gender subgroups provided in-depth and specific information that may aid in prevention and treatment of gaming addiction. In addition, incorporating familial and mental health factors as extensive and significant confounders in adjusted multivariable analysis may make the associations of gaming addiction and loneliness valid and robust. Also, the classroom surveys with rigorous, systematic, and consistent instructions by trained research assistants could yield more reliable data than online surveys.

Some limitations associated with this study should be noted when interpreting the findings reported above. First, the cross-sectional data generated from this study may make causal inferences difficult. Loneliness may be the cause or consequence of gaming addiction behaviors. The negative association between loneliness and gaming addiction behaviors adjusted for depression and anxiety may also be explained reversely. For instance, depressed students who play games may have reduced their feelings of loneliness. Besides, cross-sectional data make it uncertain whether gaming addiction behaviors increased during the COVID-19 pandemic. A longitudinal or panel study collecting information on gaming involvement and levels of loneliness of school students over time will assist in ascertaining the impact of various factors. Also, collecting qualitative data from target groups and their parents and teachers who are directly in touch with them can enrich the interpretations of the results. In addition, future studies may also include pre-COVID-19 gaming experience, types and frequencies of nongaming leisure activities, and levels of accessibility to outdoor facilities in the survey because they are important confounding factors that may provide contextual information of one's gaming behaviors. Second, the self-reported responses were less able to capture multifaceted information about loneliness. However, the single-item measure of loneliness, which is well adapted to measure loneliness, has been shown to be valid in previous studies [34]. Third, the generalizability of this study's findings is limited to upper primary school and lower secondary school students as well as context (ie, based on one city, Hong Kong). A diversified sample that includes in-school and out-of-school adolescents across cities may address this inadequacy.

Implications and Conclusions

Despite these limitations, these findings expand our knowledge regarding gaming behaviors and loneliness among primary and secondary school students in Hong Kong during the pandemic. Findings of this study have potential public health and clinical implications. The results suggest that problematic gaming behaviors are common among young people during the pandemic, and their association with loneliness varies across age groups and genders. Policy makers and designers of prevention programs must consider age and gender as significant factors that will impact the outcomes of such programs. Parents and teachers with young children can organize more nondigital leisure activities to reduce the potential overuse of digital gadgets used for gaming. Findings from this study may also help parents and teachers to plan family and school activities for children and adolescents. If choices for alternative leisure activities are limited due to public health concerns during the pandemic, parental supervision is very important. Parents need to establish ground rules that limit the time for accessing digital gadgets, help younger ones to set priorities for daily routines, and supervise children so they will not play games during their online classes. If parents are also interested in gaming, they can play together with their children to strengthen parent-child interactions, family relationships, and understanding of their children's interests and habits.

Since gaming can enhance one's sense of autonomy, competence, and connectedness with others [12], it should not just be considered as a harmful and addictive behavior. If conducted for the right amount of time and in the appropriate contexts, gaming can be a healthy behavior. It is exciting to learn that some *functional* games have been developed specifically to promote mental well-being during the pandemic [61]. For instance, a game named Shadow's Edge was developed based on empirical and theory-based foundations, such as narrative therapy and artistic expression, to help players cope with isolation and fear during school closures [62]. Interestingly, it is great to learn that major game publishers around the world and the World Health Organization are supporting the #PlayApartTogether campaign to encourage staying home and maintaining social distance by gaming at home during the lockdown in many countries [63]. In conclusion, gaming can be helpful to the well-being of young people, but can also be harmful, especially to those with lower self-control tendencies. Mature caregivers can help promote positive gaming by providing parental supervision. In addition, more game developers can take steps to design more functional games with social scientists and young people that promote mental well-being and protect players' health by including planned designs to minimize addictive gaming behaviors.

Acknowledgments

This work was supported by a grant funded by the Hong Kong Research Grant Council (Reference No. 25605418) that was awarded to SZ. We would like to thank the participants of the study and the schools for their assistance with participant recruitment.

Authors' Contributions

SZ and PW were responsible for the conception of the research question and drafting of the manuscript. SZ was responsible for the study design, data collection, and data cleaning. YZ and PL were responsible for data analysis. JL proofread the manuscript. All authors have approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary files.
[DOCX File , 78 KB - games_v9i2e26808_app1.docx]

References

1. Video games. Boston Children's Hospital Digital Wellness Lab. 2020. URL: https://cmch.tv/parents/video-games/ [accessed 2021-04-21]
2. Clement J. Percentage of teenagers who play video games in the United States as of April 2018, by gender. Statista. 2021 Jan 29. URL: https://www.statista.com/statistics/454309/teen-video-game-players-gender-usa/ [accessed 2021-04-21]
3. 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]
4. 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 [FREE Full text] [doi: 10.1186/s12888-016-0836-3] [Medline: 27160387]
5. Seok S, DaCosta B. The world's most intense online gaming culture: Addiction and high-engagement prevalence rates among South Korean adolescents and young adults. Comput Human Behav 2012 Nov;28(6):2143-2151. [doi: 10.1016/j.chb.2012.06.019]
6. Griffiths M, Wood R. Risk factors in adolescence: The case of gambling, videogame playing, and the internet. J Gambl Stud 2000;16:199-225. [doi: 10.1023/a:1009433014881]
7. Gentile D. Pathological video-game use among youth ages 8 to 18: A national study. Psychol Sci 2009 May;20(5):594-602. [doi: 10.1111/j.1467-9280.2009.02340.x] [Medline: 19476590]
8. van Rooij AJ, Kuss DJ, Griffiths MD, Shorter GW, Schoenmakers TM, van de Mheen D. The (co-)occurrence of problematic video gaming, substance use, and psychosocial problems in adolescents. J Behav Addict 2014 Sep;3(3):157-165. [doi: 10.1556/jba.3.2014.013]
9. Groarke JM, Berry E, Graham-Wisener L, McKenna-Plumley PE, McGlinchey E, Armour C. Loneliness in the UK during the COVID-19 pandemic: Cross-sectional results from the COVID-19 Psychological Wellbeing Study. PLoS One 2020 Sep 24;15(9):e0239698 [FREE Full text] [doi: 10.1371/journal.pone.0239698] [Medline: 32970764]
10. Luchetti M, Lee JH, Aschwanden D, Sesker A, Strickhouser JE, Terracciano A, et al. The trajectory of loneliness in response to COVID-19. Am Psychol 2020 Oct;75(7):897-908 [FREE Full text] [doi: 10.1037/amp0000690] [Medline: 32567879]
11. Loneliness during coronavirus. Mental Health Foundation. 2020. URL: https://www.mentalhealth.org.uk/coronavirus/loneliness-during-coronavirus [accessed 2021-04-21]
12. Ryan R, Rigby C, Przybylski A. The motivational pull of video games: A self-determination theory approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
13. von der Heiden JM, Braun B, Müller KW, Egloff B. The association between video gaming and psychological functioning. Front Psychol 2019 Jul 26;10:1731 [FREE Full text] [doi: 10.3389/fpsyg.2019.01731] [Medline: 31402891]
14. Cudo A, Kopiś N, Zabielska-Mendyk E. Personal distress as a mediator between self-esteem, self-efficacy, loneliness and problematic video gaming in female and male emerging adult gamers. PLoS One 2019 Dec 10;14(12):e0226213 [FREE Full text] [doi: 10.1371/journal.pone.0226213] [Medline: 31821356]
15. Cheng C, Cheung MW, Wang H. Multinational comparison of internet gaming disorder and psychosocial problems versus well-being: Meta-analysis of 20 countries. Comput Human Behav 2018 Nov;88:153-167. [doi: 10.1016/j.chb.2018.06.033]
16. Kim J, LaRose R, Peng W. Loneliness as the cause and the effect of problematic internet use: The relationship between internet use and psychological well-being. Cyberpsychol Behav 2009 Aug;12(4):451-455. [doi: 10.1089/cpb.2008.0327] [Medline: 19514821]
17. 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]
18. Pontes HM. Investigating the differential effects of social networking site addiction and internet gaming disorder on psychological health. J Behav Addict 2017 Dec 01;6(4):601-610 [FREE Full text] [doi: 10.1556/2006.6.2017.075] [Medline: 29130329]
19. Shaw LH, Gant LM. In defense of the internet: The relationship between internet communication and depression, loneliness, self-esteem, and perceived social support. Cyberpsychol Behav 2002 Apr;5(2):157-171. [doi: 10.1089/109493102753770552] [Medline: 12025883]
20. Zhou S, Leung L. Gratification, loneliness, leisure boredom, and self-esteem as predictors of SNS-game addiction and usage pattern among Chinese college students. Int J Cyber Behav Psychol Learn 2014;2(4):34-48. [doi: 10.4018/ijcbpl.2012100103]
21. Su W, Han X, Jin C, Yan Y, Potenza MN. Are males more likely to be addicted to the internet than females? A meta-analysis involving 34 global jurisdictions. Comput Human Behav 2019 Oct;99:86-100. [doi: 10.1016/j.chb.2019.04.021]
22. Koenig LJ, Isaacs AM, Schwartz JA. Sex differences in adolescent depression and loneliness: Why are boys lonelier if girls are more depressed? J Res Pers 1994 Mar;28(1):27-43. [doi: 10.1006/jrpe.1994.1004]
23. Maes M, Qualter P, Vanhalst J, Van den Noortgate W, Goossens L. Gender differences in loneliness across the lifespan: A meta–analysis. Eur J Pers 2019 Nov 01;33(6):642-654. [doi: 10.1002/per.2220]
24. Sadia Qureshi H, Khan M, Masroor U. Increased aggression and loneliness as potential effects of pathological video-gaming among adolescents. Pak J Soc Clin Psychol 2013;11(1):66-71 [FREE Full text]

25. Toker S, Baturay MH. Antecedents and consequences of game addiction. Comput Human Behav 2016 Feb;55:668-679. [doi: 10.1016/j.chb.2015.10.002]
26. Lee JY, Shin KM, Cho S, Shin YM. Psychosocial risk factors associated with internet addiction in Korea. Psychiatry Investig 2014 Oct;11(4):380-386 [FREE Full text] [doi: 10.4306/pi.2014.11.4.380] [Medline: 25395968]
27. Lee C, McKenzie K. Socioeconomic and geographic inequalities of internet addiction in Korean adolescents. Psychiatry Investig 2015 Oct;12(4):559-562 [FREE Full text] [doi: 10.4306/pi.2015.12.4.559] [Medline: 26508969]
28. Schneider LA, King DL, Delfabbro PH. Family factors in adolescent problematic internet gaming: A systematic review. J Behav Addict 2017 Sep 01;6(3):321-333 [FREE Full text] [doi: 10.1556/2006.6.2017.035] [Medline: 28762279]
29. Su B, Yu C, Zhang W, Su Q, Zhu J, Jiang Y. Father-child longitudinal relationship: Parental monitoring and internet gaming disorder in Chinese adolescents. Front Psychol 2018 Feb 6;9:95 [FREE Full text] [doi: 10.3389/fpsyg.2018.00095] [Medline: 29467704]
30. Choo H, Sim T, Liau AKF, Gentile DA, Khoo A. Parental influences on pathological symptoms of video-gaming among children and adolescents: A prospective study. J Child Fam Stud 2014 Mar 23;24(5):1429-1441. [doi: 10.1007/s10826-014-9949-9]
31. Hsiao YH. The relationship between the mobile game using behaviors and positive mental health among Taiwanese adolescents. DSpace at National Taiwan Normal University. Taipei, Taiwan: National Taiwan Normal University; 2016. URL: http://rportal.lib.ntnu.edu.tw/bitstream/20.500.12235/90527/1/060101011e01.pdf [accessed 2021-04-30]
32. 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]
33. Banks J, Batty GD, Coughlin K, Dangerfield P, Marmot M, Nazroo J, et al. English Longitudinal Study of Ageing: Waves 0-9, 1998-2019. [data collection]. 33rd Edition. Colchester, UK: UK Data Service; 2019. URL: https://www.elsa-project.ac.uk/ [accessed 2021-04-30]
34. Pinquart M, Sorensen S. Influences on loneliness in older adults: A meta-analysis. Basic Appl Soc Psych 2010 Jun 07;23(4):245-266. [doi: 10.1207/s15324834basp2304_2]
35. Measuring Your Impact on Loneliness in Later Life. London, UK: Campaign to End Loneliness; 2015. URL: https://www.campaigntoendloneliness.org/wp-content/uploads/Loneliness-Measurement-Guidance1-1.pdf [accessed 2021-04-21]
36. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988 Mar;52(1):30-41. [doi: 10.1207/s15327752jpa5201_2]
37. Stattin H, Kerr M. Parental monitoring: A reinterpretation. Child Dev 2000 Jul;71(4):1072-1085. [doi: 10.1111/1467-8624.00210] [Medline: 11016567]
38. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
39. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2014 Sep;36(5):539-544. [doi: 10.1016/j.genhosppsych.2014.05.021] [Medline: 25023953]
40. He X, Li C, Qian J, Cui H, Wu W. Reliability and validity of a generalized anxiety disorder scale in general hospital outpatients. Shanghai Arch Psychiatry 2010;22(4):200-203. [doi: 10.3969/j.issn.1002-0829.2010.04.002]
41. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
42. Oakes M. Measuring Socioeconomic Status. Bethesda, MD: Office of Behavioral and Social Sciences Research (OBSSR); 2016 May. URL: https://obssr.od.nih.gov/wp-content/uploads/2016/05/Measuring-Socioeconomic-Status.pdf [accessed 2021-04-21]
43. Karatasou S, Santamouris M. Socio-economic status and residential energy consumption: A latent variable approach. Energy Build 2019 Sep;198:100-105. [doi: 10.1016/j.enbuild.2019.06.013]
44. Bucy EP. Social access to the internet. Int J Press Polit 2016 Sep 14;5(1):50-61. [doi: 10.1177/1081180x00005001005]
45. US Department of Commerce, Economic and Statistics Administration, National Telecommunications and Information Administration. Falling Through the Net: Toward Digital Inclusion. A Report on Americans' Access to Technology Tools. Washington, DC: US Department of Commerce; 2000 Oct. URL: https://www.ntia.doc.gov/files/ntia/publications/fttn00.pdf [accessed 2021-04-30]
46. Sun P, Unger JB, Palmer PH, Gallaher P, Chou C, Baezconde-Garbanati L, et al. Internet accessibility and usage among urban adolescents in Southern California: Implications for web-based health research. Cyberpsychol Behav 2005 Oct;8(5):441-453. [doi: 10.1089/cpb.2005.8.441] [Medline: 16232037]
47. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 3rd edition. Hoboken, NJ: John Wiley & Sons, Inc; 2020.
48. Wang C, Chan C, Mak K, Ho S, Wong P, Ho R. Prevalence and correlates of video and internet gaming addiction among Hong Kong adolescents: A pilot study. ScientificWorldJournal 2014;2014:874648 [FREE Full text] [doi: 10.1155/2014/874648] [Medline: 25032242]
49. Trepte S, Reinecke L, Juechems K. The social side of gaming: How playing online computer games creates online and offline social support. Comput Human Behav 2012 May;28(3):832-839. [doi: 10.1016/j.chb.2011.12.003]

50. Colder Carras M, Van Rooij AJ, Van de Mheen D, Musci R, Xue Q, Mendelson T. Video gaming in a hyperconnected world: A cross-sectional study of heavy gaming, problematic gaming symptoms, and online socializing in adolescents. Comput Human Behav 2017 Mar;68:472-479 [FREE Full text] [doi: 10.1016/j.chb.2016.11.060] [Medline: 28260834]
51. Henrich CC, Blatt SJ, Kuperminc GP, Zohar A, Leadbeater BJ. Levels of interpersonal concerns and social functioning in early adolescent boys and girls. J Pers Assess 2001 Feb;76(1):48-67. [doi: 10.1207/s15327752jpa7601_3]
52. Rose AJ, Rudolph KD. A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. Psychol Bull 2006 Jan;132(1):98-131 [FREE Full text] [doi: 10.1037/0033-2909.132.1.98] [Medline: 16435959]
53. Blasi MD, Giardina A, Giordano C, Coco GL, Tosto C, Billieux J, et al. Problematic video game use as an emotional coping strategy: Evidence from a sample of MMORPG gamers. J Behav Addict 2019 Mar 01;8(1):25-34 [FREE Full text] [doi: 10.1556/2006.8.2019.02] [Medline: 30739460]
54. Kardefelt-Winther D. The moderating role of psychosocial well-being on the relationship between escapism and excessive online gaming. Comput Human Behav 2014 Sep;38:68-74. [doi: 10.1016/j.chb.2014.05.020]
55. Király O, Potenza MN, Stein DJ, King DL, Hodgins DC, Saunders JB, et al. Preventing problematic internet use during the COVID-19 pandemic: Consensus guidance. Compr Psychiatry 2020 Jul;100:152180 [FREE Full text] [doi: 10.1016/j.comppsych.2020.152180] [Medline: 32422427]
56. Király O, Urbán R, Griffiths MD, Ágoston C, Nagygyörgy K, Kökönyei G, et al. The mediating effect of gaming motivation between psychiatric symptoms and problematic online gaming: An online survey. J Med Internet Res 2015 Apr 07;17(4):e88 [FREE Full text] [doi: 10.2196/jmir.3515] [Medline: 25855558]
57. Stockdale L, Coyne SM. Video game addiction in emerging adulthood: Cross-sectional evidence of pathology in video game addicts as compared to matched healthy controls. J Affect Disord 2018 Jan 01;225:265-272. [doi: 10.1016/j.jad.2017.08.045] [Medline: 28841491]
58. Männikkö N, Ruotsalainen H, Miettunen J, Pontes HM, Kääriäinen M. Problematic gaming behaviour and health-related outcomes: A systematic review and meta-analysis. J Health Psychol 2020 Jan 01;25(1):67-81. [doi: 10.1177/1359105317740414] [Medline: 29192524]
59. Smith B. Depression and motivation. Phenomenol Cogn Sci 2012 Jun 3;12(4):615-635. [doi: 10.1007/s11097-012-9264-0]
60. Blinka L, Mikuška J. The role of social motivation and sociability of gamers in online game addiction. Cyberpsychology 2014 Jul 01;8(2):6. [doi: 10.5817/cp2014-2-6]
61. Dos Santos IK, da Silva Cunha de Medeiros RC, Azevedo de Medeiros J, de Almeida-Neto PF, Souza de Sena DC, Cobucci RN, et al. Active video games for improving mental health and physical fitness-an alternative for children and adolescents during social isolation: An overview. Int J Environ Res Public Health 2021 Feb 09;18(4):1641 [FREE Full text] [doi: 10.3390/ijerph18041641] [Medline: 33572159]
62. Miller J. Video game helps teens cope with COVID-19. Healio. 2020 Jun 10. URL: https://www.healio.com/news/primary-care/20200610/video-game-helps-teens-cope-with-covid19 [accessed 2021-04-21]
63. KLab Inc. KLab unites with the game industry and WHO #PlayApartTogether campaign to prevent the spread of COVID-19. PR Newswire. 2020 Apr 17. URL: https://en.prnasia.com/releases/global/klab-unites-with-the-game-industry-and-who-playaparttogether-campaign-to-prevent-the-spread-of-covid-19-277698.shtml [accessed 2021-04-21]

Abbreviations

aOR: adjusted odds ratio
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GAS: Game Addiction Scale
MSPSS: Multidimensional Scale of Perceived Social Support
OR: odds ratio
SES: socioeconomic status

Edited by N Zary; submitted 29.12.20; peer-reviewed by H Namba, Z Aghaei; comments to author 19.03.21; revised version received 23.03.21; accepted 02.04.21; published 07.05.21.

Please cite as:
Zhu S, Zhuang Y, Lee P, Li JCM, 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;9(2):e26808
URL: https://games.jmir.org/2021/2/e26808
doi:10.2196/26808
PMID:33960954

Corrigenda and Addenda

Correction: Virtual Reality App for Treating Eating Behavior in Eating Disorders: Development and Usability Study

Billy Sundström Langlet[1], PhD; Dorothy Odegi[1], MSc; Modjtaba Zandian[1], PhD; Jenny Nolstam[2]; Per Södersten[1], PhD; Cecilia Bergh[1,2], PhD

[1]Division of Clinical Geriatrics, Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[2]Mandometer Clinic, Stockholm, Sweden

Corresponding Author:
Billy Sundström Langlet, PhD
Division of Clinical Geriatrics, Center for Alzheimer Research
Department of Neurobiology, Care Sciences and Society
Karolinska Institutet
Hälsovägen 7, 3rd floor
Stockholm, 141 52
Sweden
Phone: 46 852488034
Email: billy.langlet@ki.se

Related Article:

Correction of: https://games.jmir.org/2021/2/e24998

(JMIR Serious Games 2021;9(2):e29686) doi:10.2196/29686

In “Virtual Reality App for Treating Eating Behavior in Eating Disorders: Development and Usability Study” (JMIR Serious Games 2021;9(2):e24998) the authors noted two errors.

Due to a system error, the name of one author, Dorothy Odegi, was replaced with the name of another author on the paper, Jenny Nolstam. In the originally published paper, the order of authors was listed as follows:

> *Billy Sundström Langlet; Jenny Nolstam; Modjtaba Zandian; Jenny Nolstam; Per Södersten; Cecilia Bergh*

This has been corrected to:

> *Billy Sundström Langlet; Dorothy Odegi; Modjtaba Zandian; Jenny Nolstam; Per Södersten; Cecilia Bergh*

In the originally published paper, the ORCID of author Jenny Nolstam was incorrectly published as follows:

> *0000-0001-6120-9412*

This has been corrected to:

> *0000-0002-9928-2778*

The correction will appear in the online version of the paper on the JMIR Publications website on April 20, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 16.04.21; this is a non–peer-reviewed article;accepted 16.04.21; published 20.04.21.

Please cite as:
Langlet BS, Odegi D, Zandian M, Nolstam J, Södersten P, Bergh C
Correction: Virtual Reality App for Treating Eating Behavior in Eating Disorders: Development and Usability Study
JMIR Serious Games 2021;9(2):e29686
URL: https://games.jmir.org/2021/2/e29686
doi:10.2196/29686
PMID:33878026

Corrigenda and Addenda

Correction: Design and Evaluation of User-Centered Exergames for Patients With Multiple Sclerosis: Multilevel Usability and Feasibility Studies

Alexandra Schättin[1], PhD; Stephan Häfliger[1], MSc; Alain Meyer[1], MSc; Barbara Früh[1], MSc; Sonja Böckler[2], MA; Yannic Hungerbühler[2], BA; Eling D de Bruin[1,3], PhD; Sebastian Frese[4], PhD; Regula Steinlin Egli[5], BSc; Ulrich Götz[2], Dipl; René Bauer[2], Lic Phil; Anna Lisa Martin-Niedecken[2], PhD

[1]Department of Health Sciences and Technology, Institute of Human Movement Sciences and Sport, ETH Zurich, Zurich, Switzerland

[2]Department of Design, Subject Area in Game Design, Zurich University of the Arts, Zurich, Switzerland

[3]Division of Physiotherapy, Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden

[4]Technology and Innovation Unit and Department of Research, ZURZACH Care, Bad Zurzach, Switzerland

[5]Physiotherapy Langmatten, Binningen, Switzerland

Corresponding Author:
Alexandra Schättin, PhD
Department of Health Sciences and Technology, Institute of Human Movement Sciences and Sport
ETH Zurich
IBWS, HCP H, Leopold Ruzicka Weg 4
Zurich, 8093
Switzerland
Phone: 41 44 633 90 57
Email: alexandra.schaettin@hest.ethz.ch

Related Article:

Correction of: https://games.jmir.org/2021/2/e22826

(JMIR Serious Games 2021;9(2):e30326) doi:10.2196/30326

In “Design and Evaluation of User-Centered Exergames for Patients With Multiple Sclerosis: Multilevel Usability and Feasibility Studies” (JMIR Serious Games 2021;9(2):e22826) the authors noted two errors.

In the originally published manuscript, the images displayed as Figures 9-13 were associated with the incorrect figure numbers. Figure captions and numbering were otherwise unaffected. The following changes have been made in the corrected version of the manuscript:

- The image displayed as Figure 9 has been corrected to appear as Figure 10.
- The image displayed as Figure 10 has been corrected to appear as Figure 11.
- The image displayed as Figure 11 has been corrected to appear as Figure 9.
- The image displayed as Figure 12 has been corrected to appear as Figure 13.
- The image displayed as Figure 13 has been corrected to appear as Figure 12.

In addition, under the section “Qualitative Data,” the following sentence appeared in the originally published version:

> *In summary, all participants reported an enjoyable, motivating, varied, and fun experience with the exergames, which was a completely new thing for most of them (Figures 9, 11, and 12).*

This has been corrected as follows to properly format in-text references to figures according to journal style:

> *In summary, all participants reported an enjoyable, motivating, varied, and fun experience with the exergames, which was a completely new thing for most of them (Figure 9, Figure 11, and Figure 12).*

The correction will appear in the online version of the paper on the JMIR Publications website on May 12, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 10.05.21; this is a non–peer-reviewed article;accepted 10.05.21; published 13.05.21.

Please cite as:
Schättin A, Häfliger S, Meyer A, Früh B, Böckler S, Hungerbühler Y, de Bruin ED, Frese S, Steinlin Egli R, Götz U, Bauer R, Martin-Niedecken AL
Correction: Design and Evaluation of User-Centered Exergames for Patients With Multiple Sclerosis: Multilevel Usability and Feasibility Studies
JMIR Serious Games 2021;9(2):e30326
URL: https://games.jmir.org/2021/2/e30326/
doi:10.2196/30326
PMID:33983893

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org